Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

Your Name	Date

Flexible and Robust Methods for Evaluating Covariate Effects on Biomedical Outcomes

By

Ying Cui

Limin Peng, Ph.D.

Amita K. Manatunga, Ph.D.

Yijian (Eugene) Huang, Ph.D.

Rabindra M. Tirouvanziam, Ph.D.

Accepted:

Kimberly Jacob Arriola, Ph.D., MPH
Dean of the James T. Laney School of Graduate Studies

Date

Flexible and Robust Methods for Evaluating Covariate Effects on Biomedical Outcomes

By

Ying Cui
BS, Nankai University, 2017

Advisor: Limin Peng, Ph.D.

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Biostatistics
2023

Abstract

Flexible and Robust Methods for Evaluating Covariate Effects on Biomedical Outcomes

In practice, it is often of scientific interest to evaluate covariate effects on biomedical outcomes. This task can be complicated by the presence of dynamic (or varying) variable effects that often manifest meaningful scientific mechanisms. Appropriately accounting for possible dynamic effects is crucial to avoid depreciating some important variables. Moreover, with technology advancement, modern biomedical studies often collect a huge number of variables, posing ultra-high dimensional data settings. Furthermore, important contributors of biomedical outcomes may evolve over time, posing time-dependent covariates. The overall objective of my dissertation is to develop statistical methods that can provide robust and flexible assessment of covariate effects that can address the limitations of existing approaches while leading to meaningful scientific discovery.

In the first dissertation project, we adopt the device of globally concerned quantile regression, and propose a flexible testing framework suited to assess either constant or dynamic covariate effects on outcomes subject to random censoring. We study the powerful Kolmogorov-Smirnov (K-S) and Cramér-Von-Mises (C-V) type test statistics and develop a resampling procedure to tackle their complicated limit distributions. We provide rigorous theoretical results, including the limit null distributions, consistency under a general class of alternative hypotheses of the proposed tests, and the justifications for the presented resampling procedure. Extensive numerical studies demonstrate the utility of the new testing procedures and their advantages over existing approaches.

In the second dissertation project, we propose a model-free testing and screening framework by adopting a global view pertaining to the concept of interval quantile independence. The new framework not only permits robust identification of variables dynamically associated with an outcome, but also offers the flexibility to evaluate multiple covariates simultaneously, where the covariates under consideration can be either continuous or discrete. The key testing strategy naturally evolves into unconditional and conditional screening procedures for ultra-high dimensional settings that enjoys the desirable sure screening property. We demonstrate good practical utility of the proposed methods via extensive simulation studies and a real application to a microarray data set.

In the third dissertation project, utilizing the interval-quantile index, we propose a new model-free globally-concerned test statistic for evaluating the impact of time-dependent covariates on time-to-event outcomes. Additionally, we develop a resampling procedure based on perturbation resampling. We establish the limit null distributions and consistency under a general class of alternative hypotheses of the proposed tests and provide justification for the resampling procedure. The proposed methods are demonstrated through extensive simulation studies, as well as an application to the *Feeding Infants Right... from the STart* (FIRST) study.

Flexible and Robust Methods for Evaluating Covariate Effects on Biomedical Outcomes

By

Ying Cui
BS, Nankai University, 2017

Advisor: Limin Peng, Ph.D.

Acknowledgement

I want to express my sincere gratitude to Dr. Limin Peng, my advisor, for her invaluable supervision, unwavering support, and endless patience during my PhD study. Her insightful feedback motivate me to sharpen my thinking and polish my work to a higher level. Her immense knowledge, enormous enthusiasm, rigorousness in research, and meticulousness make her a role model who has already had a positive impact on my career and will continue to do so in the future. Without her invaluable help and guidance, completing this dissertation would have been impossible.

I also want to express my heartfelt appreciation to Dr. Amita Manatunga for her useful suggestions not limited to research and encouragement for all these years.

Moreover, I would like to express my deepest appreciation to my other committee members: Dr. Yijian Eugene Huang, and Dr. Rabin M. Tirouvanziam for generously offering their time and suggestions. I would also like to thank Dr. Jeong Hoon Jang, Dr. Zhaohui Steve Qin, Dr. Huichuan J. Lai, Dr. Yijuan Hu and other faculty members that I have been fortunate to collaborated with. I learned a lot from all of them.

I thank to my best friend, Dr. Zilu Liu, for her unwavering support and encouragement. I am glad to have her in my life. My sincere thanks also go to my Emory colleagues and friends, for accompanying me throughout these years.

Finally, I owe my deepest gratitude to my parents whose constant love and support kept me motivated and confident throughout this journey. My accomplishments and success are because they believed in me. I also extend my heartfelt appreciation to my husband, Dr. Bo Wei, who is the love of my life. I am forever grateful for his unconditional love and support, especially throughout the pandemic, and for the privilege of sharing my life with him.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 A Brief Overview

A general question arising from many research studies is to determine whether some covariates are relevant to a study outcome. For instance, in chronic disease studies, researchers may seek to identify prognostic factors that are predictive for the time-to-event disease outcome. As another example, in a genetics study, it is often of interest to identify a group of genes that contribute to the variations of a known disease marker or symptom (Subramanian et al., 2005; Efron and Tibshirani, 2007; Newton et al., 2007, for example). Addressing such an interest, however, may be complicated by the presence of dynamic (or varying) covariate effects.

Moreover, revolutionized with technological advances, genetics studies are experiencing a significant shift from studying a few genes of interest to exploring the entire genome all at once, posing additional challenges due to ultra-high dimensionality. Firstly, when $p > n$, there are more likely to have a singular fitting. Secondly, it becomes harder to eliminate unimportant variables as p gets larger. This is due to the fact that as p increases, the unimportant variables are more likely to be correlated with the important variables.

Furthermore, time-dependent covariates are common in biomedical studies where the outcome of interest evolves over time. For example, in the *Feeding Infants Right.. from the STart* (FIRST) study, the covariate is the feeding type (e.g. breast milk, formula) that can change over time, and the outcome is the time to infection. We want to investigate whether feeding can affect the risk of infection. Ignoring time-dependent feature of the covariates in statistical analysis can lead to biased and inaccurate results, reducing the statistical power of the analysis and potentially impacting the study's conclusions. Therefore, it is crucial to employ appropriate methods to handle these time-dependent covariates.

In the next section, we present some literature reviews on existing approaches to evaluate covariate effects on time-to-event outcome, existing work on variable screen-

ing for ultra-high dimensional data, as well as the existing methods to assess the effect of a time-dependent covariates on a survival outcome. Following these, we briefly describe the proposed methods. At the end of this chapter, we will give an outline of the rest parts of this dissertation.

1.2 Literature Review

1.2.1 Existing Approaches to Evaluate Covariate Effects on Time-to-event Outcome

When the disease outcome is captured by a time-to-event, a commonly used approach is to model the mechanism of a prognostic factor influencing the time-to-event outcome via a standard survival regression model and then test the corresponding covariate effects (see a review in Kleinbaum and Klein (2010) and Cox and Oakes (2018)). These standard survival regression models, such as the Cox proportional hazard (PH) regression model and the accelerated failure time (AFT) model, impose assumptions like the proportional hazards and the location-shift effects, which implicitly confine the prognostic factor of interest to be a static portent of disease progression.

There has been growing awareness that a prognostic factor may follow a dynamic association with a time-to-event disease outcome. Many reports in literature (Dickson et al., 1989; Thorogood et al., 1990; Verweij and van Houwelingen, 1995; Bellera et al., 2010, for example) have suggested that postulating constant covariate effects, sometimes, is not adequate to reflect underlying physiological disease mechanisms, leading to distorted assessment of the prognostic factor. For example, an analysis of a dialysis dataset reported by Peng and Huang (2008) suggested that the severity of restless leg syndrome (RLS) symptoms may be prognostic of mortality for short-term dialysis survivors but not for long-term dialysis survivors. The standard tests based

on the Cox PH model and the AFT model failed to detect such a dynamic effect.

Quantile regression, as a useful alternative, directly formulates covariate effects on quantile(s) of a response, thus confers a seminal venue to characterize a dynamic effect of a prognostic factor. Many authors have studied linear quantile regression with a time-to-event outcome (Powell, 1986; Ying et al., 1995; Portnoy, 2003; Zhou, 2006; Peng and Huang, 2008; Wang and Fygenson, 2009; Huang, 2010, for example). Most of the existing methods concern covariate effects on a single or multiple pre-specified quantile levels (e.g. Δ is a singleton set $\{0.5\}$), and, following the terminology of Zheng et al. (2015), are locally concerned. As discussed in Zheng et al. (2015), locally concerned quantile regression cannot inform of the covariate effect on quantiles other than the specifically targeted ones (e.g. median), and thus may miss important prognostic factors. Adopting the perspective of globally concerned quantile regression, one can simultaneously examine covariate effects over a continuum of quantile levels (e.g. Δ is an interval $[0.1, 0.9]$), and thus confer a more comprehensive assessment of a prognostic factor. However, powerful tests tailored to evaluate covariate effects under the perspective of globally concerned quantile regression have not been formally studied, partly owing to the associated inferential complexity.

1.2.2 Existing Work on Variable Screening for Ultra-high Dimensional Data

In the context of variable screening, the importance of a covariate was ranked by marginal correlation (Fan and Lv, 2008), maximum marginal likelihood estimate of a generalized linear model (Fan et al., 2010) or a generalized marginal utility function (Fan et al., 2009), and generalized correlation (Hall and Miller, 2009). These approaches involve an assumed linear or generalized linear relationship between the outcome and covariates or transformation thereof, which implicitly asserts a location-shift (or constant) effect for each covariate. Such a restriction was relaxed in model-

free screening procedures through adopting nonparametric regression modeling (Fan et al., 2011; He et al., 2013, for example). However, there was a subtle limitation that the adopted nonparametric modeling only examines the local influence of a covariate on the mean or a pre-specified quantile of the outcome. A relevant covariate can be missed if its impact on the outcome is not manifested on the mean or the targeted quantile level.

Addressing these caveats, a viable option is to measure a covariate's outcome-relevance pertaining to the concept of interval quantile independence (Zhu et al., 2018). Several authors (Zhu et al., 2011; Mai and Zou, 2015, for example) considered a similar general framework for defining covariate relevance, which utilizes the functional dependence of the conditional cumulative distribution of the outcome upon the covariate. Compared to this alternative, employing the conditional quantile function allows one to naturally pinpoint one part of the outcome distribution (for covariate effect assessment) with a proper choice of Δ to align with particular scientific interests, for example, in average or abnormal outcomes. The flexibility in specifying Δ may also help mitigate potential identifiability concern. For example, when data are limited, say due to censoring, simply setting $\Delta = (0, 1)$ may necessitate extrapolation with additional model assumptions.

To examine covariate effects over Δ, one available approach is to utilizing the novel interval quantile index proposed by Zhu et al. (2018), which is designed to measure the departure from the interval quantile independence between a pair of continuous variables. Zhu et al. (2018)'s nonparametric index estimator and the associated asymptotic theory naturally render a testing procedure when all the covariates are continuous. Zhu et al. (2018) also developed a model-free variable screening procedure that ranks the estimated interval quantile index for the relationship between the outcome and each continuous covariate. While enjoying desirable theoretical properties (e.g., sure screening property) and appealing empirical performance, Zhu et al.

(2018)'s procedures would encounter difficulty when some covariates are discrete. In addition, as the interval quantile index is oriented to study the relationship between two variables, it is not straightforward to adapt Zhu et al. (2018)'s procedures to simultaneously evaluate multiple covariates in terms of their relevance to the outcome. This task is often needed in practice in order to sensibly account for inherent data hierarchy structure due to biological, spatial, or temporal factors.

1.2.3 Existing Methods to Detect Effects for Survival Outcome with Time Dependent Covariate

To account for time-dependent covariates in time-to-event data, the most widely used approach is the time dependent Cox proportional hazards model, which serves as the foundation for handling time-dependent covariates in survival analysis (Therneau and Grambsch, 2000; Cox and Oakes, 2018). A pivotal contribution was made by Andersen and Gill (1982), which presents a counting process formulation of the Cox proportional hazards model, allowing for the incorporation of time-dependent covariates. Among other early works in this area, Robins and Tsiatis (1992) introduces a semi-parametric estimation approach for accelerated failure time (AFT) model with time-dependent covariates. However, both methods have limitations, such as the reliance on assumption for the proportional hazards and the location-shift effects, which may not always hold in real-world applications. Other approaches, such as the joint models for longitudinal and survival data, have also been developed to address time-dependent covariates (Rizopoulos, 2012). These joint models typically rely on various strong assumptions, such as the distributional form of random effects, the shape of the baseline hazard function, and the functional relationship between the longitudinal and survival outcomes. Violation of these assumptions can lead to biased or inconsistent parameter estimates, affecting the accuracy and reliability of the evaluation. Therefore, non-parametric approaches would be appealing in analyzing survival

data with time-dependent covariates, as they require fewer assumptions about the underlying data distribution compared to these model-based approaches.

The landmark approaches also serve as flexible and practical tools for handling time-dependent covariates. As reviewed in Putter (2013), the basic idea of landmark analysis is to define a set of landmark times and restrict the analysis to subjects who are still at risk at each landmark time. Through examining time-dependent covariate effects on survival outcomes at multiple time points, this approach provide valuable insights into the complex relationships between covariates and survival outcomes. While landmark analysis is a useful method for analyzing time-to-event data with time-dependent covariates, most of the existing works (Andersen and Perme, 2010; Van Houwelingen, 2007, for example) are based on the assumption that the effect of the covariate is constant over the specified time period after the landmark time. This assumption may not hold if there is dynamic association between the covariate and the time-to-event disease outcome.

1.3 The Proposed Methods

In this dissertation, we propose to assess the effects of covariates on a response variable by adopting a global view pertaining to the concept of interval quantile independence. The first topic introduces a flexible and robust method for evaluating the effects of a univariate covariate on a time-to-event outcome. The second topic develops a group testing framework to identify variables that are dynamically associated with the outcome, and introduces a screening framework for excluding irrelevant variables in the ultra-high dimensional settings with high confidence. The third topic suggests an index for measuring and testing the departure from interval quantile independence for time-dependent covariates and time-to-event outcomes across a set of landmark times.

The first part of this dissertation focus on the problem to detect the effect of a univariate covariate across an interval of quantile levels on the time-to-event outcome. We develop a new framework for evaluating a survival prognostic factor following the spirit of globally concerned quantile regression. Quantile regression (Koenker and Bassett Jr, 1978), which directly formulates covariate effects on quantile(s) of a response, confers a seminal venue to characterize a dynamic effect of a prognostic factor. Specifically, given a time-to-event outcome T and a covariate $\widetilde{Z}$ (which represents the prognostic factor of interest), a linear quantile regression model may assume,

$$Q_T(\tau|\widetilde{Z}) = \exp\{\boldsymbol{Z}^\intercal \boldsymbol{\theta}_0(\tau)\}, \quad \tau \in \Delta, \tag{1.1}$$

where $\boldsymbol{Z} = (1, \widetilde{Z})^\intercal$, $Q_T(\tau|\widetilde{Z}) \equiv \inf\{t : \Pr(T \leq t|\widetilde{Z}) \geq \tau\}$ denotes the τ-th conditional quantile of T given $\widetilde{Z}$, $\boldsymbol{\theta}_0(\tau) \equiv (\beta_0^{(0)}(\tau), \beta_0^{(1)}(\tau))^\intercal$ is an unknown coefficient vector, and $\Delta \subseteq (0,1)$ is a pre-specified set including the quantile levels of interest. The coefficient $\beta_0^{(1)}(\tau)$ represents the effect of $\tilde{Z}$ on the τ-th conditional quantile of T, and is allowed to change with τ. This implicates that the prognostic factor is permitted to have different effects across different segments of the distribution of the time-to event outcome.

As a proof of concept, we shall confine the scope of this work to the standard survival setting where the time-to-event outcome T is subject to random censoring. Specifically, our proposal is to simultaneously assess the influence of the prognostic factor on a range of quantiles of T, indexed by a τ-interval, $[\tau_L, \tau_U] \subset (0,1)$. As the key rationale, a significant prognostic factor is allowed to have a dynamic τ-varying effect, which may be non-zero throughout the whole τ-interval (i.e. full effect), or only over a part of the τ-interval (i.e. partial effect). Under this view, when model (2.1) with $\Delta = [\tau_L, \tau_U]$ holds, the task of identifying a prognostic factor reduces to

testing the null hypothesis,

$$H_0 : \ \beta_0^{(1)}(\tau) = 0, \quad \tau \in [\tau_L, \tau_U].$$

Moreover, without assuming any models, we may consider the null hypothesis formulated as,

$$H_0^* : \ Q_T(\tau|\widetilde{Z}) = Q_T(\tau) \ \text{ for } \ \tau \in [\tau_L, \tau_U],$$

where $Q_T(\tau) = \inf\{t : \Pr(T \le t) \ge \tau\}$, denoting the τ-th unconditional (or marginal) quantile of T. The null hypothesis H_0^* corresponds to the setting where $\widetilde{Z}$ has no influence on the conditional quantile of T at any quantile level between τ_L and τ_U.

It is remarkable that under mild regularity conditions, H_0^* implies that model (2.1) holds with $\Delta = [\tau_L, \tau_U]$ and $\beta_0^{(1)}(\tau) = 0$ for $\tau \in [\tau_L, \tau_U]$; on the other hand, model (2.1) with $\Delta = [\tau_L, \tau_U]$ and $\beta_0^{(1)}(\tau) = 0$ for $\tau \in [\tau_L, \tau_U]$ implies $Q_T(\tau|\widetilde{Z}) = Q_T(\tau)$ for $\tau \in [\tau_L, \tau_U]$; see Lemma A2.1 in Section 2.5.1. This finding sheds an important insight that a model-based test developed for H_0 may be used towards testing the model-free null hypothesis H_0^*. From an alternative view, this result suggests that the globally concerned quantile regression model (2.1) with $\Delta = [\tau_L, \tau_U]$ can be used as a working model to test H_0^*, which adopts the view that the effect of a prognostic factor can be assessed through contrasting the conditional versus unconditional quantiles of T.

Regarding H_0^*, we study two "omnibus" test statistics constructed based on the estimator of $\boldsymbol{\theta}_0(\tau)$ obtained under the working model (2.1) with $\Delta = [\tau_L, \tau_U]$. One test is a Kolmogorov-Smirnov (K-S) type test statistic defined upon the maximum "signal" strength (i.e. covariate effect) over τ's in $[\tau_L, \tau_U]$. The other one is a CramérVon-Mises (C-V) type test statistics based on the average "signal" strength over τ's in $[\tau_L, \tau_U]$. These two types of test statistics are known to be very sensitive to detect any departure from the null hypothesis H_0 under model (2.1). However, the analytic form of their limit null distributions are generally complex and sometimes intractable.

This challenge is more intense in the quantile regression setting, where coefficient estimates do not have a closed form, and the corresponding asymptotic variance matrix involves unknown density functions (Koenker, 2005). To overcome these difficulties, we propose to approximate the limit null distributions through a resampling procedure that perturbs the influence function associated with the adopted coefficient estimator under the working model (2.1), following similar strategies of Lin et al. (1993) and Li and Peng (2014). We derive a sample-based procedure to estimate the influence function without requiring the correct specification of model (2.1), thereby circumvents directly evaluating the unknown density function via smoothing. The proposed resampling procedure is easy to implement and is shown to perform well even with realistic sample sizes. Moreover, we provide rigorous theoretical justifications for the proposed resampling procedure.

For the second part of this dissertation, we consider the problem to identify variables that are dynamically associated with the outcome, and to exclude irrelevant variables in the ultra-high dimensional settings with high confidence. Let Y denote a continuous outcome and let $\boldsymbol{X} = \{X^{(1)}, \ldots, X^{(p)}\}^T$ denote the vector of the observed covariates. Define

$$H_{0,j}: \; Q_Y(\tau \mid X^{(j)}) = Q_Y(\tau), \text{ a.s. for } \tau \in \Delta \subseteq (0,1).$$

Here and hereafter, for a general random vector $\boldsymbol{V}$, $Q_Y(\tau \mid \boldsymbol{V}) = \inf\{y : \; \text{pr}(Y \leq y|\boldsymbol{V}) \geq \tau\}$ denotes the conditional quantile function of Y given $\boldsymbol{V}$, and $Q_Y(\tau) = \inf\{y : \; \text{pr}(Y \leq y) \geq \tau\}$ denotes the unconditional quantile function of Y. When $X^{(j)}$ is continuous, $H_{0,j}$ refers to the interval quantile independence between Y and $X^{(j)}$ on quantile level intervals Δ and $[0,1]$ respectively for Y and $X^{(j)}$, as termed by Zhu et al. (2018). The consideration of $H_{0,j}$ confers a flexible view for defining relevant variables. In the multivariate setting, a covariate $X^{(j)}$ is considered as relevant or

active if $Q_Y(\tau \mid \boldsymbol{X})$ functionally depends on $X^{(j)}$ for some $\tau \in \Delta \subseteq (0,1)$, where Δ is a pre-specified set of quantile levels. Under this view, the set of relevant variables is defined as $\mathcal{M}_\Delta = \{1 \leq r \leq p : \text{ there exists } \tau \in \Delta \text{ such that } Q_Y(\tau \mid \boldsymbol{X}) \text{ depends on } X^{(r)}\}$. The formulations of $H_{0,j}$ and $\mathcal{M}_\Delta$ take a global perspective to assess covariate effects throughout the range of the outcome distribution indexed by the quantile level interval Δ. Covariates in $\mathcal{M}_\Delta$ are permitted to have dynamic and non-additive effects across different ranges of the outcome.

We propose a new model-free strategy for tackling a generalized version of $H_{0,j}$ that concerns the outcome relevance of one or multiple covariates, which can be either continuous or discrete. Specifically, for a index set for J covariates, $G = \{r_1, \ldots, r_J\} \subseteq \{1, \ldots, p\}$, define $\boldsymbol{X}_G = \{X^{(r_1)}, \ldots, X^{(r_J)})^\intercal$. A null hypothesis of our interest takes the form

$$H_{0,G} : \ Q_Y(\tau \mid \boldsymbol{X}_G) = Q_Y(\tau), \text{ a.s., } \quad \text{for } \ \tau \in \Delta \subseteq (0,1).$$

To address $H_{0,G}$, we propose to employ a "working" linear quantile regression model, which can help determine whether $H_{0,G}$ holds or not. We construct an omnibus test statistic for $H_{0,G}$ from adapting the spirit of the classic Cramér-Von-Mises (C-V) type test statistics under the "working" linear quantile regression model. We establish the asymptotic behaviors of the proposed test statistic without assuming the working model holds.

We further utilize the proposed test statistic as the utility function to develop a new model-free variable screening procedure for ultra-high dimensional data. Given the flexibility of our test statistic in handling multiple covariates simultaneously, the new screening procedure can be performed with covariates pre-grouped by scientific needs or in a random manner for the benefit of saving computational time. We establish the desirable sure screening property for the new screening procedure. As a useful by-product, we can readily transform the new screening procedure to perform conditional variable screening given some known relevant covariates under mild

additional assumptions. We also prove the corresponding conditional sure screening property.

In the third part of this dissertation, we focus on the problem to evaluate the association between a univariate time dependent covariate (continuous, discrete, or a mixture of both) and the time-to-event outcome. Let T be a continuous random variable representing the survival time. Denote $\bar{Z} = \{Z(t), 0 \leq t < \infty\}$ as a time dependent covariate process, where $Z(t)$ is a univariate covariate at $t \geq 0$. Let C denote the censoring time. Define $X = \min(T, C)$ and $\delta = I(T \leq C)$ as the observed survival time and the event status, respectively.

To deal with the time dependent covariates, we adopt the idea to assessing the overall effects over a pre-determined set of landmark times. Specifically, at each landmark time t_0, we assess the effect on individuals who have not yet experienced the event, and then summarize across all landmark times. Let $M(t_0) = I(X > t_0)$ denote the indicator on whether an individual survive up to time t_0. Let $T^r(t_0)$ be the residual survival time, and $V(t_0)$ be a pre-specified functional form of $\bar{Z}(t_0) = \{Z(t), 0 \leq t \leq t_0\}$, the covariate process up to t_0. Define $C^r(t_0) = C - t_0$ as the residual censoring time. The observed residual survival time is then defined as $X^r(t_0) = \min(T^r(t_0), C^r(t_0))$. We assume $(T^r(t_0), V(t_0)) \perp C \mid M(t_0) = 1$ for any $t_0 \in \mathcal{T}$. The observed data at time t_0 is then noted as $\{X_i^r(t_0), \delta_i, V_i(t_0)\}_{\{i; M_i(t_0)=1, i=1...,n\}}$.

Denote $\mathcal{T}$ as a finite set of the landmark time points of interest, and $\Delta = [\tau_L, \tau_U]$ as the interval of quantile levels of interest for $T^r(t_0)$. Based on the problem of interest, we may formulate it as a hypothesis testing problem to test for

$$H_{0,\mathcal{T}} : Q_{T^r(t_0)|V(t_0),M(t_0)=1}(\tau_1) = Q_{T^r(t_0)|M(t_0)=1}(\tau_1) \text{ for } \tau_1 \in \Delta, t_0 \in \mathcal{T}$$

versus its alternative. Here $Q_{T^r(t_0)|M(t_0)=1}(\tau_1)$ denotes the τ_1th quantile of $T^r(t_0)$ conditional on $M(t_0) = 1$, and $Q_{T^r(t_0)|V(t_0),M(t_0)=1}(\tau_1)$ denotes the τ_1th quantile of

$T^r(t_0)$ conditional on $V(t_0)$ and $M(t_0) = 1$. To address $H_{0,\mathcal{T}}$, we propose two model-free indices, which can help determine whether $H_{0,\mathcal{T}}$ holds or not. We establish the asymptotic behaviors of the proposed test statistic. We also propose to approximate the limit null distributions through a perturbation resampling procedure that perturbs the influence function.

1.4 Outline

The rest of this proposal is organized as follows. In Chapter 2, we first briefly review some existing results about the estimation of model (2.1), which we use as a working model for testing H_0^*. We then present the proposed test statistics along with their theoretical properties. A resampling procedure is developed to carry out inference regarding H_0 or H_0^* based on the proposed test statistics. We also discuss some computational strategies to help simplify or improve the implementation of the proposed method. We report extensive simulation studies conducted to evaluate the finite-sample performance of the proposed testing procedures. Our simulation results show that the proposed tests have accurate empirical sizes and can be much more powerful than benchmark methods when assessing a covariate with a dynamic effect. We further demonstrate the usefulness of the proposed testing procedures with a real data example. In Chapter 3, we present the proposed testing method for the classic multivariate setting, and the unconditional and conditional screening procedures for the ultra-high dimensional setting. We report the results from extensive simulation studies. In settings with dynamic effects, the proposed global testing and screening procedures clearly outperform existing approaches that assume constant effects or locally focus on the covariate effects on the mean or a pre-specified quantile of the outcome. We also present an application of the proposed methods to a microarray dataset. In Chapter 4, we present the proposed test statistics along with their the-

oretical properties. A resampling procedure is developed for conducting inference based on the proposed test statistics. Our simulation results show that the proposed tests have accurate empirical sizes and can be much more powerful than benchmark methods when assessing a covariate with a dynamic effect. We further demonstrate the usefulness of the proposed testing procedures by an application to the data from the FIRST study. In Chapter 5, we provide a summary of the completed work and briefly discuss our plan for future work.

Chapter 2

Assessing Dynamic Covariate Effects with Survival Data

2.1 The Proposed Testing Procedures

2.1.1 Formulation of the Testing Problem

In this part, we focus on the problem to detect the effect of a univariate covariate across an interval of quantile levels on the time-to-event outcome. We develop a new framework for evaluating a survival prognostic factor following the spirit of globally concerned quantile regression. Quantile regression (Koenker and Bassett Jr, 1978), which directly formulates covariate effects on quantile(s) of a response, confers a seminal venue to characterize a dynamic effect of a prognostic factor. Specifically, given a time-to-event outcome T and a covariate $\widetilde{Z}$ (which represents the prognostic factor of interest), a linear quantile regression model may assume,

$$Q_T(\tau|\widetilde{Z}) = \exp\{\boldsymbol{Z}^\intercal \boldsymbol{\theta}_0(\tau)\}, \quad \tau \in \Delta, \tag{2.1}$$

where $\boldsymbol{Z} = (1, \widetilde{Z})^\intercal$, $Q_T(\tau|\widetilde{Z}) \equiv \inf\{t : \Pr(T \leq t|\widetilde{Z}) \geq \tau\}$ denotes the τ-th conditional quantile of T given $\widetilde{Z}$, $\boldsymbol{\theta}_0(\tau) \equiv (\beta_0^{(0)}(\tau), \beta_0^{(1)}(\tau))^\intercal$ is an unknown coefficient vector, and $\Delta \subseteq (0, 1)$ is a pre-specified set including the quantile levels of interest. The coefficient $\beta_0^{(1)}(\tau)$ represents the effect of $\tilde{Z}$ on the τ-th conditional quantile of T, and is allowed to change with τ. This implicates that the prognostic factor is permitted to have different effects across different segments of the distribution of the time-to event outcome.

As a proof of concept, we shall confine the scope of this work to the standard survival setting where the time-to-event outcome T is subject to random censoring. Specifically, our proposal is to simultaneously assess the influence of the prognostic factor on a range of quantiles of T, indexed by a τ-interval, $[\tau_L, \tau_U] \subset (0, 1)$. As the key rationale, a significant prognostic factor is allowed to have a dynamic τ-varying effect, which may be non-zero throughout the whole τ-interval (i.e. full effect), or

only over a part of the τ-interval (i.e. partial effect). Under this view, when model (2.1) with $\Delta = [\tau_L, \tau_U]$ holds, the task of identifying a prognostic factor reduces to testing the null hypothesis,

$$H_0 : \ \beta_0^{(1)}(\tau) = 0, \quad \tau \in [\tau_L, \tau_U].$$

Moreover, without assuming any models, we may consider the null hypothesis formulated as,

$$H_0^* : \ Q_T(\tau|\widetilde{Z}) = Q_T(\tau) \ \text{ for } \ \tau \in [\tau_L, \tau_U],$$

where $Q_T(\tau) = \inf\{t : \Pr(T \leq t) \geq \tau\}$, denoting the τ-th unconditional (or marginal) quantile of T. The null hypothesis H_0^* corresponds to the setting where $\widetilde{Z}$ has no influence on the conditional quantile of T at any quantile level between τ_L and τ_U.

It is remarkable that under mild regularity conditions, H_0^* implies that model (2.1) holds with $\Delta = [\tau_L, \tau_U]$ and $\beta_0^{(1)}(\tau) = 0$ for $\tau \in [\tau_L, \tau_U]$; on the other hand, model (2.1) with $\Delta = [\tau_L, \tau_U]$ and $\beta_0^{(1)}(\tau) = 0$ for $\tau \in [\tau_L, \tau_U]$ implies $Q_T(\tau|\widetilde{Z}) = Q_T(\tau)$ for $\tau \in [\tau_L, \tau_U]$; see Lemma A2.1 in Section 2.5.1. This finding sheds an important insight that a model-based test developed for H_0 may be used towards testing the model-free null hypothesis H_0^*. From an alternative view, this result suggests that the globally concerned quantile regression model (2.1) with $\Delta = [\tau_L, \tau_U]$ can be used as a working model to test H_0^*, which adopts the view that the effect of a prognostic factor can be assessed through contrasting the conditional versus unconditional quantiles of T.

2.1.2 Estimation of the Coefficient Function Under the Censored Quantile Regression Model

As explained in Section 2.1.1, we propose to use globally concerned quantile regression as a vehicle to address the testing problem regarding the general null hypothesis H_0^*. The first step is to obtain an estimator of $\boldsymbol{\theta}_0(\tau)$ (and thus $\beta_0^{(1)}(\tau)$) from fitting the

working model (2.1) to the observed data. Here and hereafter, we shall set the Δ in model (2.1) as $\Delta = [\tau_L, \tau_U]$, which is a pre-specified interval within $(0, 1)$. Let C denote time to censoring, $X = \min(T, C)$, and $\delta = I(T \leq C)$. The observed data include n i.i.d. replicates of $(X, \delta, \boldsymbol{Z})$, denoted by $\{(X_i, \delta_i, \boldsymbol{Z}_i)\}_{i=1}^n$.

To estimate $\boldsymbol{\theta}_0(\tau)$ under model (2.1), we choose to adapt the existing results of Peng and Fine (2009) developed for competing risks data to the setting with randomly censored data. Compared to the other available estimators developed by Portnoy (2003) and Peng and Huang (2008), which require $\tau_L = 0$, the estimator derived from Peng and Fine (2009) is more robust to any realistic violation of the global linearity assumed by model (2.1) (Peng, 2021). The influence function associated with Peng and Fine (2009)'s estimator also has a simpler form that can facilitate the development of the corresponding testing procedures.

The estimator of $\boldsymbol{\theta}_0(\tau)$ adapted from Peng and Fine (2009)'s work, denoted by $\widehat{\boldsymbol{\theta}}(\tau)$, is obtained as the solution to the following estimating equation:

$$S_n(\boldsymbol{b}, \tau) = n^{-1/2} \sum_{i=1}^{n} \boldsymbol{Z}_i \left[\frac{I(X_i \leq \exp\{\boldsymbol{Z}_i^{\intercal} \boldsymbol{b}\}) I(\delta_i = 1)}{\widehat{G}(X_i | \boldsymbol{Z}_i)} - \tau \right] = 0, \tag{2.2}$$

where $\widehat{G}(x|\boldsymbol{Z})$ is a reasonable estimator of $G(x|\boldsymbol{Z}) \equiv \Pr(C \geq x|\boldsymbol{Z})$. For simplicity of illustration, in sequel, we shall assume C is independent of $\tilde{Z}$ and thus take $\widehat{G}(x|\boldsymbol{Z})$ as the Kaplan-Meier estimator of the marginal survival function of C, $\widehat{G}(x)$. As noted by Peng and Fine (2009), solving (2.2) can be formulated as a L_1-type minimization problem of the following convex objective function:

$$\begin{aligned} \boldsymbol{U}_n(\boldsymbol{b}, \tau) \;=\; & \sum_{i=1}^{n} I(\delta_i = 1) \left| \frac{\log(X_i)}{\widehat{G}(X_i)} - \boldsymbol{b}^{\intercal} \frac{\boldsymbol{Z}_i}{\widehat{G}(X_i)} \right| + \left| M - \boldsymbol{b}^{\intercal} \sum_{l=1}^{n} \frac{-\boldsymbol{Z}_l I(\delta_l = 1)}{\widehat{G}(X_i)} \right| \\ & + \left| M - \boldsymbol{b}^{\intercal} \sum_{k=1}^{n} (2\tau \boldsymbol{Z}_k) \right|. \end{aligned}$$

Here M is a sufficiently large number. This L_1-type minimization problem can be

easily solved using the `rq()` function in the R package `quantreg` by Koenker (2022).

By the results of Peng and Fine (2009), the estimator $\widehat{\boldsymbol{\theta}}(\tau)$ enjoys desirable asymptotic properties. Specifically, under certain regularity conditions, we have (i) $\lim_{n\to\infty}\sup_{\tau\in[\tau_L,\tau_U]}||\widehat{\boldsymbol{\theta}}(\tau)-\boldsymbol{\theta}_0(\tau)||\to_p 0$; and (ii) $\sqrt{n}\{\widehat{\boldsymbol{\theta}}(\tau)-\boldsymbol{\theta}_0(\tau)\}$ converge weakly to a mean zero Gaussian process for $\tau\in[\tau_L,\tau_U]$ with covariance function $\boldsymbol{\Phi}(\tau',\tau)=E\{\boldsymbol{\xi}_1(\tau')\boldsymbol{\xi}_1(\tau)^\top\}$. Here $\boldsymbol{\xi}_i(\tau)$ $(i=1,\ldots,n)$ are defined as

$$\begin{aligned}\boldsymbol{\xi}_i(\tau) &\equiv \{\boldsymbol{\xi}_i^{(0)}(\tau),\boldsymbol{\xi}_i^{(1)}(\tau)\}^\top \\ &= \{A(\boldsymbol{\theta}_0(\tau))\}^{-1}\left\{\boldsymbol{Z}_i(\frac{I(\log(X_i)\le\boldsymbol{Z}_i^\top\boldsymbol{\theta}_0(\tau),\delta_i=1)}{G(X_i)}-\tau)\right. \\ &\quad \left.-\int_0^\infty \boldsymbol{w}\{\boldsymbol{\theta}_0(\tau),s\}y(s)^{-1}dM_i^G(s)\right\},\end{aligned}$$

where $G(x)=\Pr(C>x)$, $\boldsymbol{A}(\boldsymbol{b})=E[\boldsymbol{Z}\boldsymbol{Z}^\top f(\boldsymbol{Z}^\top\boldsymbol{b}|\boldsymbol{Z})]$ with $f(t|\boldsymbol{Z})$ denoting the conditional density of X given $\boldsymbol{Z}$, $\boldsymbol{w}(\boldsymbol{b},t)=E[\boldsymbol{Z}Y(t)I(X\le\exp\{\boldsymbol{Z}^\top\boldsymbol{b}\})I(\delta=1)G(X)^{-1}]$, and $M_i^G(t)=N_i^G(t)-\int_0^\infty Y_i(s)d\Lambda^G(t)$ with $N_i^G(t)=I(X_i\le t,\delta_i=0)$, $Y_i(t)=I(X_i\ge t)$, $y(t)=\Pr(X\ge t)$, $\lambda^G(t)=\lim_{\Delta\to0}P(C\in(t,t+\Delta)|C\ge t)/\Delta$, and $\Lambda^G(t)=\int_0^t\lambda^G(s)ds$. In addition, $n^{1/2}\{\widehat{\boldsymbol{\theta}}(\tau)-\boldsymbol{\theta}_0(\tau)\}\approx n^{-1/2}\sum_{i=1}^n\boldsymbol{\xi}_i(\tau)$, where $\approx$ indicate asymptotical equivalence uniformly in $\tau\in[\tau_L,\tau_U]$. Consequently, $\boldsymbol{\xi}_i(\tau)$ is referred to as the influence function of $n^{1/2}\{\widehat{\boldsymbol{\theta}}(\tau)-\boldsymbol{\theta}_0(\tau)\}$.

Note that the variance estimation for $\widehat{\boldsymbol{\theta}}(\tau)$ is complicated by the involvement of the unknown density $f(t|\boldsymbol{Z})$ in the asymptotic covariance matrix $\boldsymbol{\Phi}(\tau',\tau)$. As justified by Peng and Fine (2009), a sample-based procedure that avoids smoothing-based density estimation can be used for variance estimation and is outlined below:

(1.a) Compute an consistent variance estimate for $\boldsymbol{S}_n(\boldsymbol{\theta}_0(\tau), \tau)$ given by

$$\widehat{\boldsymbol{\Sigma}}(\tau,\tau) =$$
$$n^{-1}\sum_{i=1}^{n} \boldsymbol{Z}_i^{\otimes 2}\left(\frac{I[\log(X_i) \le \boldsymbol{Z}_i^{\intercal}\widehat{\boldsymbol{\theta}}(\tau)), \delta_i = 1]}{\widehat{G}(X_i)} - \tau\right)^2$$
$$-n^{-1}\sum_{i=1}^{n} I(\delta_i = 0)\left(\sum_{j=1}^{n} \boldsymbol{Z}_j I(X_j \ge X_i) I[\log(X_j) \le \boldsymbol{Z}_i^{\intercal}\widehat{\boldsymbol{\theta}}(\tau), \delta_j = 1]\{\widehat{G}(X_j)\}^{-1}\right/$$
$$\left.\sum_{j=1}^{n} I(X_j \ge X_i)\right)^{\otimes 2},$$

where for a vector $\boldsymbol{a}$, $\boldsymbol{a}^{\otimes 2} = \boldsymbol{a}\boldsymbol{a}^{\intercal}$.

(1.b) Find a symmetric and nonsingular matrix $\boldsymbol{E}_n(\tau) \equiv \{\boldsymbol{e}_{n,0}(\tau), \boldsymbol{e}_{n,1}(\tau)\}$ such that $\{\boldsymbol{E}_n(\tau)\}^2 = \widehat{\boldsymbol{\Sigma}}(\tau,\tau)$.

(1.c) Calculate $\boldsymbol{D}_n(\tau) = \{\boldsymbol{S}_n^{-1}\{\boldsymbol{e}_{n,0}(\tau), \tau\} - \widehat{\boldsymbol{\theta}}(\tau), \boldsymbol{S}_n^{-1}\{\boldsymbol{e}_{n,1}(\tau), \tau\} - \widehat{\boldsymbol{\theta}}(\tau)\}$, where $\boldsymbol{S}_n^{-1}\{\boldsymbol{e}(\tau), \tau\}$ is the solution to the perturbed estimating equation $\boldsymbol{S}_n(\boldsymbol{b}, \tau) = \boldsymbol{e}(\tau)$.

(1.d) Obtain an estimate for the asymptotic variance of $\sqrt{n}\{\widehat{\boldsymbol{\theta}}(\tau) - \boldsymbol{\theta}_0(\tau)\}$ as $\boldsymbol{V}_n(\tau) \equiv n\boldsymbol{D}_n^{\otimes 2}(\tau)$.

Here $\boldsymbol{E}_n(\tau)$ can be computed with the eigenvalue eigenvector decomposition of $\widehat{\boldsymbol{\Sigma}}(\tau,\tau)$ using the R function `eigen()`. As another important remark, the above procedure ensures that the perturbation terms, $\boldsymbol{e}_{n,j}(\tau), j = 1, 2$, have the desired asymptotic order. As a result, this procedure remains valid when $\boldsymbol{e}_{n,j}(\tau)$ in step (1.c) is replaced by $u \cdot \boldsymbol{e}_{n,j}(\tau)$ for some constant u. Based on our numerical experiences, incorporating some constant u can help stabilize variance estimation when sample size is small or τ is close to 0 or 1. Variance estimation based on the above procedure is found to have satisfactory finite sample performance based on some unreported simulation studies.

2.1.3 The Proposed Test Statistics and Theoretical Properties

Express $\widehat{\boldsymbol{\theta}}(\tau) \equiv (\widehat{\beta}^{(0)}(\tau), \widehat{\beta}^{(1)}(\tau))'$ and let $\hat{\sigma}_n^{(1)}(\tau)$ denote the square root of the second diagonal element of $\boldsymbol{V}_n(\tau)$, which corresponds to the variance estimate for $\sqrt{n}\widehat{\beta}^{(1)}(\tau)$ under H_0^*. We propose to construct two "omnibus" test statistics based on $\widehat{\beta}^{(1)}(\tau)$ and $\hat{\sigma}_n^{(1)}(\tau)$:

$$\widehat{T}_{sup}^{(1)} = \sup_{\tau \in [\tau_L, \tau_U]} \left| \frac{\sqrt{n}\widehat{\beta}^{(1)}(\tau)}{\widehat{\sigma}_n^{(1)}(\tau)} \right|,$$

and

$$\widehat{T}_{inte}^{(1)} = \int_{\tau_L}^{\tau_U} \left| \frac{\sqrt{n}\widehat{\beta}^{(1)}(\tau)}{\widehat{\sigma}_n^{(1)}(\tau)} \right|^2 d\tau.$$

These two test statistics mimic the classic Kolmogorov-Smirnov (K-S) test statistic and Cramér-Von-Mises (C-V) test statistic for two-sample distribution comparisons (Darling, 1957). Under model (2.1), $\widehat{T}_{sup}^{(1)}$ and $\widehat{T}_{inte}^{(1)}$ capture the maximum and average magnitude of the covariate effect over $\tau \in [\tau_L, \tau_U]$ respectively. By this design, both test statistics are sensitive to any type of departures from the null hypothesis H_0 and can be used to construct powerful tests for H_0.

Without assuming model (2.1), we can also show that $\widehat{T}_{sup}^{(1)}$ and $\widehat{T}_{inte}^{(1)}$ provide valid tests for H_0^* and have power approaching one under a general class of alternative hypotheses as specified in Theorem 2.2. The key insight is that even when model (2.1) does not hold, $\widehat{\theta}(\tau)$ may still converge in probability to a deterministic function $\widetilde{\boldsymbol{\theta}}(\tau) \equiv (\widetilde{\beta}^{(0)}(\tau), \widetilde{\beta}^{(1)}(\tau))'$ that is the solution to $\boldsymbol{\mu}(\boldsymbol{b}, \tau) \equiv E[\boldsymbol{Z}\{I(\log T \leq \boldsymbol{Z}^\intercal \boldsymbol{b}) - \tau\}] = 0$. It is easy to see that $\widetilde{\boldsymbol{\theta}}(\tau) = \boldsymbol{\theta}_0(\tau)$ under model (2.1). By Lemma A2.1, it follows that under H_0^*, $\widetilde{\beta}^{(1)}(\tau) = 0$ for $\tau \in [\tau_L, \tau_U]$. As detailed in Theorems A2.1 and A2.2 in Section 2.5.1, under certain regularity conditions, we further have $\lim_{n\to\infty} \sup_{\tau \in [\tau_L, \tau_U]} ||\widehat{\boldsymbol{\theta}}(\tau) - \widetilde{\boldsymbol{\theta}}(\tau)|| \to_p 0$, and $\sqrt{n}\{\widehat{\boldsymbol{\theta}}(\tau) - \widetilde{\boldsymbol{\theta}}(\tau)\}$ converge weakly to a mean zero Gaussian process for $\tau \in [\tau_L, \tau_U]$ with covariance function $\widetilde{\boldsymbol{\Phi}}(\tau', \tau) =$

$E\{\widetilde{\boldsymbol{\xi}}_1(\tau')\widetilde{\boldsymbol{\xi}}_1(\tau)^\intercal\}$, where $\widetilde{\boldsymbol{\xi}}_i(\tau)$ $(i = 1, \ldots, n)$ are defined as

$$\begin{aligned}
&\widetilde{\boldsymbol{\xi}}_i(\tau) \equiv \{\widetilde{\xi}_i^{(0)}(\tau), \widetilde{\xi}_i^{(1)}(\tau)\}^\intercal \\
= \;& \{A(\widetilde{\boldsymbol{\theta}}(\tau))\}^{-1} \left\{ \boldsymbol{Z}_i \left(\frac{I(\log(X_i) \le \boldsymbol{Z}_i^\intercal \widetilde{\boldsymbol{\theta}}(\tau), \delta_i = 1)}{G(X_i)} - \tau \right) - \int_0^\infty \boldsymbol{w}\{\widetilde{\boldsymbol{\theta}}(\tau), s\} y(s)^{-1} dM_i^G(s) \right\}.
\end{aligned}$$

A useful by-product from the proof of Theorem A2.2 is that

$$n^{1/2}\{\widehat{\boldsymbol{\theta}}(\tau) - \widetilde{\boldsymbol{\theta}}(\tau)\} \approx n^{-1/2} \sum_{i=1}^{n} \widetilde{\boldsymbol{\xi}}_i(\tau), \tag{2.3}$$

We can prove these results by adapting the arguments of Peng and Fine (2009) which utilize model assumption (2.1) only through using its implication $\boldsymbol{\mu}(\boldsymbol{\theta}_0, \tau) = 0$ for $\tau \in [\tau_L, \tau_U]$. This provides the critical justification for why $\widehat{\beta}^{(1)}(\tau)$ can be used to test H_0^* even when model (2.1) does not hold. The sample-based procedure reviewed in Section 2.1.2 is still applicable to estimate the asymptotic covariance matrix $\widetilde{\boldsymbol{\Phi}}(\tau', \tau)$.

In Theorems 2.1 and 2.2, we establish useful asymptotic properties of $\widehat{T}_{sup}^{(1)}$ and $\widehat{T}_{inte}^{(1)}$ without assuming model (2.1). Specifically, in Theorem 2.1, we provide the limit distributions of the proposed test statistics under the null hypothesis H_0^*:

Theorem 2.1. *Assuming the regularity conditions 2.1–2.5 in Section 2.5 hold, under the null hypothesis H_0 or H_0^*, we have*

$$\begin{aligned}
\widehat{T}_{sup}^{(1)} &= \sup_{\tau \in [\tau_L, \tau_U]} \left| \frac{n^{1/2}\widehat{\beta}^{(1)}(\tau)}{\widehat{\sigma}_n^{(1)}(\tau)} \right| \to_d \sup\{|\mathcal{X}^{(1)}(\tau)|, \tau \in [\tau_L, \tau_U]\} \\
\widehat{T}_{inte}^{(1)} &= \int_{\tau_L}^{\tau_U} \left| \frac{n^{1/2}\widehat{\beta}^{(1)}(\tau)}{\widehat{\sigma}_n^{(1)}(\tau)} \right|^2 d\tau \to_d \int_{\tau_L}^{\tau_U} \{\mathcal{X}^{(1)}(\tau)\}^2 d\tau,
\end{aligned}$$

where $\mathcal{X}^{(1)}(\tau)$ is a mean zero Gaussian process defined in Section 2.5.3.

We also investigate the asymptotic behavior of the proposed test statistics under a general class of alternative hypotheses. The findings are stated in Theorem 2.2.

Theorem 2.2. *Assuming the regularity conditions 2.1–2.5 in Section 2.5 hold,*
(A) $\widehat{T}^{(1)}_{sup}$ *is consistent against the alternative hypothesis*

$$H_{a,1}: \quad \sup_{\tau\in[\tau_L,\tau_U]} \left|\widetilde{\beta}^{(1)}(\tau)\right| > 0.$$

(B) $\widehat{T}^{(1)}_{inte}$ *is consistent against the alternative hypothesis:*

$$H_{a,2}: \quad \int_{\tau_L}^{\tau_U} \{\widetilde{\beta}^{(1)}(\tau)\}^2 d\tau > 0.$$

The results of Theorem 2.2 indicate that the test statistics have power approaching to 1 (as n goes to ∞) under alternative cases subject to very mild constraints. Given the smoothness of $\widetilde{\beta}^{(1)}(\cdot)$, a general scenario that ensures the consistency of both $\widehat{T}^{(1)}_{sup}$ and $\widehat{T}^{(1)}_{inte}$ can be described as

> $\widetilde{H}_a$: There exists an interval $[\tau_1, \tau_2] \subseteq [\tau_L, \tau_U]$ such that $|\widetilde{\beta}^{(1)}_0(\tau)| > 0$ for $\tau \in [\tau_1, \tau_2]$.

This suggests that the proposed tests are powerful to identify a significant prognostic factor even when it only influences a segment of the outcome distribution, not necessarily the whole outcome distribution. This feature is conceptually appealing for handling a dynamic covariate effect, which may not have similar effect strength across different quantiles. The detailed proofs for Theorems 2.1 and 2.2 can be found in Section 2.5.3.

2.1.4 The Proposed Resampling Procedure to Obtain p values

The results in Theorem 2.1 suggest that $\widehat{T}^{(1)}_{sup}$ and $\widehat{T}^{(1)}_{inte}$, like the classic K-S test statistic and C-V test statistic, have complex, non-standard limit null distributions.

This motivates us to develop a resampling-based procedure to approximate their limit null distributions and obtain the corresponding p values for testing H_0^*.

Our key strategy is to approximate the distribution of $n^{1/2}\{\widehat{\beta}^{(1)}(\tau) - \widetilde{\beta}_0^{(1)}(\tau)\}$, which reduces to $n^{1/2}\widehat{\beta}^{(1)}(\tau)$ under H_0, through perturbing the influence function $\widetilde{\xi}_i^{(1)}(\tau)$, which is the second component of $\widetilde{\boldsymbol{\xi}}_i(\tau)$. Similarly ideas were used by other authors, for example, Lin et al. (1993) and Li and Peng (2014). The core justification of our proposal is provided by equation (2.3), which suggests that $n^{-1/2}\sum_{i=1}^n \widetilde{\xi}_i^{(1)}(\tau)\iota_i/\widehat{\sigma}_n^{(1)}(\tau)$ may be used to approximate $\sqrt{n}\widehat{\beta}^{(1)}(\tau)/\widehat{\sigma}_n^{(1)}(\tau)$, where $\{\iota_i\}_{i=1}^n$ are i.i.d. standard normal variates.

Specifically, we take the following steps:

(2.a) Generate B independent sets of $\{\iota_i^b\}_{i=1}^n$, where $\{\iota_i^b\}_{i=1}^n$ are independent random variables from a standard normal distribution and $b = 1, 2, \ldots, B$.

(2.b) Compute the estimates for the influence function $\widetilde{\xi}_i^{(1)}(\tau)$ as the second component of

$$\widehat{\boldsymbol{\xi}}_i(\tau) = \{\widehat{A}(\widehat{\boldsymbol{\theta}}(\tau))\}^{-1}\left\{\boldsymbol{Z}_i\left(\frac{I[\log(X_i) \le \boldsymbol{Z}_i^{\intercal}\widehat{\boldsymbol{\theta}}(\tau)), \delta_i = 1]}{\widehat{G}(X_i)} - \tau\right)\right.$$
$$\left. - I(\delta_i = 0)\frac{\sum_{j=1}^n \boldsymbol{Z}_j I(X_j \ge X_i) I[\log(X_j) \le \boldsymbol{Z}_j^{\intercal}\widehat{\boldsymbol{\theta}}(\tau), \delta_j = 1]\{\widehat{G}(X_j)\}^{-1}}{\sum_{j=1}^n I(X_j \ge X_i)}\right\},$$

where $\widehat{\boldsymbol{A}}\{\widehat{\boldsymbol{\theta}}(\tau)\}^{-1} = n^{1/2}\boldsymbol{D}_n(\tau)\boldsymbol{E}_n(\tau)^{-1}$.

(2.c) For $b = 1, \ldots, B$, calculate

$$\widehat{T}_{sup,b}^{(1)} = \sup_{\tau\in[\tau_L,\tau_U]}\left|\frac{n^{-1/2}\sum_{i=1}^n \widehat{\xi}_i^{(1)}(\tau)\iota_i^b}{\widehat{\sigma}_n^{(1)}(\tau)}\right| \quad \text{and} \quad \widehat{T}_{inte,b}^{(1)} = \int_{\tau_L}^{\tau_U}\left|\frac{n^{-1/2}\sum_{i=1}^n \widehat{\xi}_i^{(1)}(\tau)\iota_i^b}{\widehat{\sigma}_n^{(1)}(\tau)}\right|^2 d\tau,$$

where $\widehat{\xi}_i^{(1)}(\tau)$ is the second component of $\widehat{\boldsymbol{\xi}}_i(\tau)$.

(2.d) The p values based on $\widehat{T}^{(1)}_{sup}$ and $\widehat{T}^{(1)}_{inte}$ are calculated respectively as

$$p^{(1)}_{sup} = \sum_{b=1}^{B} I(\widehat{T}^{(1)}_{sup,b} > \widehat{T}^{(1)}_{sup})/B \quad \text{and} \quad p^{(1)}_{inte} = \sum_{b=1}^{B} I(\widehat{T}^{(1)}_{inte,b} > \widehat{T}^{(1)}_{inte})/B.$$

The resampling procedure presented above is easy to implement without involving smoothing. The rigorous theoretical justification for the presented resampling procedure is provided in Section 2.5.4.

2.1.5 Some Computational Considerations

Note that $\widehat{\beta}^{(1)}(\tau)$ and $\widehat{\sigma}^{(1)}_n(\tau)$ are piecewise constant; thus an exact calculation of the supremum or integration involved in $\widehat{T}^{(1)}_{sup}$ and $\widehat{T}^{(1)}_{inte}$ is possible. Alternatively, we may follow the recommendation of Zheng et al. (2015) to compute $\widehat{T}^{(1)}_{sup}$ and $\widehat{T}^{(1)}_{inte}$ based on a simpler piecewise-constant approximation of $\widehat{R}(\tau) \equiv \widehat{\beta}^{(1)}(\tau)/\widehat{\sigma}^{(1)}_n(\tau)$ on a pre-determined fine τ-grid, $\mathcal{G} \equiv \tau_L = \tau_1 < \tau_2 < \ldots < \tau_{N_*} = \tau_U$, with the grid size $\max_{1\le l\le N_*-1}(\tau_{l+1} - \tau_l) = o(n^{-1/2})$. In this case, the proposed test statistics can be calculated as

$$\widehat{T}^{(1)}_{sup} = \sqrt{n}\max\{\hat{R}(\tau_l) : 1 \le l \le N_*\}, \quad \widehat{T}^{(1)}_{inte} = \sum_{l=1}^{N_*-1} n\{\widehat{R}(\tau_l)\}^2(\tau_{l+1} - \tau_l). \tag{2.4}$$

When n is not large, the sample-based variance estimation (i.e. the computation of $\widehat{\sigma}^{(1)}_n(\tau)$) sometimes is not stable. Our remedy is to replace the $\boldsymbol{e}_{n,j}(\tau)$ in step (1.c) (see Section 2.1.2) with $u \cdot \boldsymbol{e}_{n,j}(\tau)$, where u is a pre-specified constant. We develop the following algorithm to determine a good choice of the adjusting constant u among a set of candidate values, $\mathcal{U} = \{1, 2, \ldots, U\}$.

(3.a) For each $u \in \mathcal{U}$, calculate $\widehat{R}(\tau; u) \equiv \widehat{\beta}^{(1)}(\tau)/\widehat{\sigma}^{(1)}_n(\tau; u)$ for $\tau \in \mathcal{G}$, where $\widehat{\sigma}^{(1)}_n(\tau; u)$ is the $\widehat{\sigma}^{(1)}_n(\tau)$ computed with the adjusting constant u.

(3.b) For each $u \in \mathcal{U}$, calculate $\widehat{R}^*(u) = \max_{\tau\in\mathcal{G}} \widehat{R}(\tau; u)$ and $\widehat{R}^\dagger(u) = \text{median}_{\tau\in\mathcal{G}}\widehat{R}(\tau; u)$.

(3.c) For each $u \in \mathcal{U}$, calculate $\widetilde{R}(u) = \max_{\tau\in\mathcal{G}} \max\{\boldsymbol{V}_n(\tau; u)\} - \min_{\tau\in\mathcal{G}} \min\{\boldsymbol{V}_n(\tau; u)\}$, where $\boldsymbol{V}_n(\tau; u)$ is $\boldsymbol{V}_n(\tau)$ computed with the adjusting constant u. Here, for a matrix $\boldsymbol{A}$, $\max(\boldsymbol{A})$ (or $\min(\boldsymbol{A})$) denotes the largest (or the smallest) component of the matrix $\boldsymbol{A}$.

(3.d) Assign a large positive value to $A^{[0]}$ and $B^{[0]}$, say 10^5. Set $k = 1$ and $u^{[0]} = U+1$.

(i) If $\widehat{R}^*(k) - \widehat{R}^\dagger(k) < A^{[k-1]}$ and $\widetilde{R}(k) < B^{[k-1]}$, then let $A^{[k]} = \widehat{R}^*(k) - \widehat{R}^\dagger(k)$, $B^{[k]} = \widetilde{R}(k)$, and $u^{[k]} = k$. Otherwise, let $A^{[k]} = A^{[k-1]}$, $B^{[k]} = B^{[k-1]}$ and $u^{[k]} = u^{[k-1]}$.

(ii) Increase k by 1 and go back to (i) until $k > U$.

(3.e) If $u^{[U]} < U + 1$, then choose u as $u^{[U]}$. Otherwise, no appropriate u can be selected from $\mathcal{U}$.

By this algorithm, we provide an empirical strategy to select u based on two estimation instability measures: (A) $\widehat{R}^*(k) - \widehat{R}^\dagger(k)$, which reflects the spread of $\widehat{R}(\tau) \equiv \widehat{\beta}^{(1)}(\tau)/\widehat{\sigma}_n^{(1)}(\tau)$ over τ given $u = k$; (B) $\widetilde{R}(k)$, which measures the maximum fluctuation of the estimated variance matrices across τ given $u = k$. It is clear that both measures would be large when unstable variance estimation occurs. Our algorithm first compares them with pre-specified initial values, $A^{[0]}$ and $B^{[0]}$, to rule out the occurrence of obviously outlying estimates of $\widehat{R}(\tau)$ or $\widehat{\sigma}_n^{(1)}(\tau)$. Once these two measures are found to meet the stability criteria set by the initial values with some $u \in \mathcal{U}$, the algorithm will proceed to check if other u's can yield smaller values of the instability measures. The output from this algorithm is either the value of u that produces the smallest instability measures, or an error message indicating that none of the constants in $\mathcal{U}$ can lead to stable estimation required by the proposed testing procedure. Based on our numerical experiences, setting $\mathcal{U} = \{1, 2, \ldots, 6\}$, which corresponds to $U = 6$, works well for small sample sizes such as 200 or 400. In a rare case where this algorithm fails to identify an appropriate u, we recommend adaptively increasing

the value of U until an appropriate u can be identified. Our extensive numerical experiences suggest that incorporating the adjusting constant u selected by this algorithm results in good and stable numerical performance of the proposed tests. The algorithm can be easily generalized to allow $\mathcal{U}$ to include non-integer values.

2.2 Numerical Studies

We conduct extensive simulation studies to investigate the finite-sample performance of the proposed resampling-based testing procedures. To simulate randomly censored data, we consider six setups where T and $\widetilde{Z}$ follow different relationships. In all setups, we generate $\widetilde{Z}$ from $Uniform(0,1)$ and generate censoring time C from $Uniform(U_L, U_U)$, where U_L and U_U are properly specified to produce 15% or 30% censoring. Let $\Phi(\cdot)$ denote the cumulative distribution function of the standard normal distribution. The six simulation set-ups are described as follows.

(I) Setup I: Generate T such that $Q_\tau\{\log(T)\} = \Phi^{-1}(\tau)$. Set $(U_L, U_U) = (2, 3.8)$ to produce 15% censoring, and set $(U_L, U_U) = (1, 2.5)$ to produce 30% censoring.

(II) Setup II: Generate T such that $Q_\tau\{\log(T)\} = 0.2X + \Phi^{-1}(\tau)$. Set $(U_L, U_U) = (2.5, 3.9)$ to produce 15% censoring and set $(U_L, U_U) = (1.2, 2.8)$ to produce 30% censoring.

(III) Setup III: Generate T such that $Q_\tau\{\log(T)\} = 0.5X + \Phi^{-1}(\tau)$. Set $(U_L, U_U) = (2.7, 4.9)$ to produce 15% censoring, and set $(U_L, U_U) = (1.5, 3)$ to produce 30% censoring.

(IV) Setup IV: Generate T such that $Q_\tau\{\log(T)\} = l_4(\tau)X + \Phi^{-1}(\tau)$, where $l_4(\tau)$ is as plotted in Figure 2.1. Set $(U_L, U_U) = (2, 3.9)$ to produce 15% censoring, and set $(U_L, U_U) = (1, 2.5)$ to produce 30% censoring.

(V) Setup V: Generate T such that $Q_\tau\{\log(T)\} = l_5(\tau)X + \Phi^{-1}(\tau)$, where $l_5(\tau)$ is as plotted in Figure 2.1. Set $(U_L, U_U) = (5.2, 6.5)$ to produce 15% censoring, and set $(U_L, U_U) = (1.5, 3.5)$ to produce 30% censoring.

(VI) Setup VI: Generate T such that $Q_\tau\{\log(T)\} = l_6(\tau)X + \Phi^{-1}(\tau)$, where $l_6(\tau)$ is as plotted in Figure 2.1. Set $(U_L, U_U) = (3.5, 5.5)$ to produce 15% censoring, and set $(U_L, U_U) = (1.1, 3.5)$ to produce 30% censoring.

Under all setups, model (1) holds for $\tau \in (0, 1)$ and thus for $\tau \in [0.1, 0.6]$, a pre-specified τ-interval of interest $[\tau_L, \tau_U]$. In Figure 2.1, we plot the true coefficient function $\beta_0^{(1)}(\tau)$ for each setup. It is easy to see that setup (I) represents a null case, where $\widetilde{Z}$ has no effect on any quantile of T. Setup (II) and (III) are two setups where $\widetilde{Z}$ has nonzero constant effects over all $\tau \in [0.1, 0.6]$. The constant effect in setup (II) has a magnitude of 0.2, which is smaller than that in setup (III), which is 0.5. In setups (IV), (V), and (VII), $\widetilde{Z}$ has a dynamic effect varying across different τ's. More specifically, $\widetilde{Z}$ has a partial effect over the τ-interval $[0.1, 0.49]$ in setup (IV). In setup (V), the magnitude of $\widetilde{Z}$'s effect is symmetric around 0.5, while the sign of the effect is opposite for $\tau < 0.5$ and for $\tau > 0.5$, and the effect equals 0 at $\tau = 0.5$. In setup (VI), the τ-varying effect pattern of $\widetilde{Z}$ is similar to that in setup (V) except that there is a small interval around 0.5 where $\widetilde{Z}$ has no effect in setup(VI).

We compare the proposed method with the Wald test based on the Cox PH model, denoted by "CPH (Wald)", as well as the Wald test based on the locally concerned quantile regression that focuses on $\tau = 0.4, 0.5$, or 0.6, denoted by "CQR (Wald)". To implement CQR (Wald), we adopt Peng and Huang (2008)'s estimates with variance estimated by bootstrapping. The resampling size used for both CQR (Wald) and the proposed testing procedures is set as 2500. In the sequel, we shall refer the testing procedures based on $\widehat{T}_{sup}^{(1)}$ and $\widehat{T}_{inte}^{(1)}$ respectively to as GST and GIT. For all the methods, we consider sample sizes 200, 400, and 800. We set $\mathcal{U} = \{1, \ldots, 6\}$ when implementing the algorithm for selecting the constant u.

In Table 2.1, we report the empirical rejection rates based on 1000 simulations. The results in setup I show that the proposed GIT, and the existing tests, CQR (Wald) and CPH (Wald), have empirical sizes quite close to the nominal level 0.05. The proposed GST yields relatively larger empirical type I errors as compared to the other tests. The empirical size of GST equals 0.1 when the sample size is 200 but decreases to 0.077 when the sample size increases to 800. Such an anti-conservative behavior of GST is not surprising because the K-S type test statistic is defined based on the largest value of $\widehat{\beta}^{(1)}(\tau)/\widehat{\sigma}_n^{(1)}(\tau)$ over $\tau \in [0.1, 0.6]$, which is more sensitive to a possible outlying value of $\widehat{\sigma}_n^{(1)}(\tau)$ at some τ.

When the quantile effect of $\widetilde{Z}$ is constant over τ (i.e. setups (II) and (III)), we note that in setup (II) where the effect size (i.e. magnitude of the constant effect) is relatively small, CPH (Wald) has lower empirical power as compared to the proposed GIT and GST, and the power improvement associated with the proposed GIT and GST is more evident with the smaller sample size 200. In setup (III), where the effect size is larger, CPH (Wald) still generally has lower empirical power compared to the proposed tests but its empirical power becomes comparable to that of GIT when the sample size is large (i.e. $n = 800$). These observations suggest that even in the trivial constant effect cases, the proposed tests can outperform the traditional Cox regression based tests in data scenarios with small effect sizes or sample sizes. In both setups (II) and (III), the locally concerned CQR (Wald) consistently yields lower empirical power than the proposed globally concerned GIT and GST. This reflects the power benefit resulted from integrating information on covariate effects on different quantiles as in GST and GIT, rather than focusing on the covariate effect on a single quantile as in CQR (Wald).

In setups (IV), (V), and (VI), the effect of $\widetilde{Z}$ is τ-varying, reflecting its dynamic association with T. In these cases, CPH (Wald), which assumes a constant covariate effect, can have poor power to detect the dynamic effect of $\widetilde{Z}$ (e.g. 8.3% empirical

power in setup (VI) with $n = 800$ in the presence of 30% censoring), while the proposed GST and GIT may yield much higher power (e.g. >99% power in setup (VI) with $n = 800$ in the presence of 30% censoring). The locally concerned CQR (Wald) can have higher power than CPH (Wald) when the targeted quantile level is within the τ-region where $\beta_0^{(1)}(\tau)$ is non-zero. When the targeted quantile level is outside the τ-region with non-zero effect, such as $\tau = 0.6$ in setup (IV) or $\tau = 0.5$ in setups (V) and (VI), the CQR (Wald) has even poorer power compared to CPH (Wald). This is well expected because these cases may serve as the null cases for the locally concerned CQR (Wald). This confirms that CQR (Wald) is inadequate to capture the meaningful effect of $\widetilde{Z}$ that is manifested at non-targeted quantiles.

We compare the simulation results across settings that are only differed by the censoring distribution. For each relationship between $\widetilde{Z}$ and T specified by setups (I)-(VI), we consider three different censoring distributions to yield 0%, 15%, and 30% censoring. The results for settings with 15% and 30% censoring are presented in Table 2.1 and the results based on uncensored data are presented in Table A2.1 in Section 2.5.5. From our comparisons, we find that quantile regression based tests, including GST, GIT and CQR (Wald), demonstrate small variations in empirical powers as the censoring rate (or distribution) changes. In cases with a constant covariate effect, the Cox regression based test, CPH (Wald), also has similar performance among settings with different censoring rates. However, in setup (V), where the covariate effect is not constant over τ, CPH (Wald) has reasonably good power when there is no censoring or only 15% censoring, but its performance deteriorates considerably when the censoring rate is increased to 30%. We have a similar observation for CPH (Wald) in setup (VI). A reasonable interpretation of these observations is that the capacity to detect a dynamic effect can be weakened by incorrectly assuming a constant proportional hazard effect and can be further attenuated by the missing data from censoring.

We also investigate whether the proposed tests are sensitive to the choice of $\mathcal{U}$.

We conduct additional simulation studies with $\mathcal{U}$ set as $\{1, \ldots, 3\}$, $\{1, \ldots, 6\}$, and $\{1, \ldots, 12\}$ for the six set-ups with 15% censoring. The results are summarized in Table A2.2 in Section 2.5. From this table, we note that GIT is quite robust to the change in $\mathcal{U}$, while GST demonstrates more variations across different choices of $\mathcal{U}$. Another observation is that GIT becomes less sensitive to the change in $\mathcal{U}$ when the sample size becomes larger. A possible explanation for these results is similar to that for the observed anti-conservative behavior of GST. That is, GST, by its construction, is sensitive to any outlying value of $\widehat{\sigma}_n^{(1)}(\tau)$ with $\tau \in [\tau_L, \tau_U]$, which is more likely to occur when the sample size is not large.

Aligning with the definitions of the proposed tests, the simulation results suggest that GST, as compared to GIT, is more sensitive to detect a departure from the null hypothesis, yielding higher power. This observation is also consistent with the anti-conservative behavior of GST observed in the null cases, which is reflected by empirical sizes notably greater than 0.05. With a smaller sample size, such as $n = 200$, GST can produce quite elevated type I errors, while GIT yields more reasonable empirical sizes. Therefore, in practice, one may need to exercise caution for applying GST to a small dataset, for which we recommend using GIT instead.

In summary, our simulation results demonstrate the proposed testing procedures have robust satisfactory performance for detecting a covariate of either a constant or dynamic effect. The new tests tend to exhibit greater advantages over benchmark approaches when the covariate presents a dynamic effect, or the covariate has a constant effect but of a small magnitude.

2.3 Real Example with Dialysis Data

To illustrate the utility of the proposed testing framework, we apply our method to investigate the prognostic factors for dialysis survival based on a dataset collected

Table 2.1: Empirical rejection rate based on 1000 simulations.

Set-up	n	Proposed Test		CQR (Wald)			CPH (Wald)
		GST	GIT	$\tau = 0.4$	$\tau = 0.5$	$\tau = 0.6$	
15% censoring							
I	200	0.100	0.073	0.066	0.062	0.057	0.049
	400	0.091	0.078	0.072	0.072	0.066	0.051
	800	0.077	0.055	0.064	0.063	0.059	0.061
II	200	0.234	0.167	0.117	0.131	0.117	0.115
	400	0.275	0.214	0.155	0.153	0.150	0.178
	800	0.410	0.362	0.277	0.265	0.247	0.322
III	200	0.566	0.485	0.359	0.401	0.360	0.450
	400	0.786	0.772	0.585	0.592	0.576	0.722
	800	0.957	0.957	0.873	0.887	0.865	0.960
IV	200	0.377	0.254	0.097	0.060	0.053	0.063
	400	0.652	0.478	0.116	0.065	0.063	0.067
	800	0.939	0.816	0.148	0.047	0.058	0.090
V	200	0.653	0.464	0.143	0.070	0.118	0.260
	400	0.937	0.827	0.208	0.071	0.153	0.458
	800	0.999	0.993	0.291	0.053	0.279	0.757
VI	200	0.731	0.552	0.149	0.062	0.086	0.125
	400	0.971	0.896	0.198	0.055	0.095	0.201
	800	1.000	0.995	0.260	0.033	0.142	0.364
30% censoring							
I	200	0.171	0.095	0.062	0.060	0.048	0.047
	400	0.110	0.085	0.069	0.074	0.065	0.056
	800	0.066	0.052	0.063	0.059	0.050	0.038
II	200	0.302	0.186	0.115	0.122	0.105	0.122
	400	0.305	0.221	0.152	0.156	0.138	0.188
	800	0.411	0.359	0.277	0.259	0.245	0.298
III	200	0.681	0.539	0.360	0.393	0.322	0.432
	400	0.828	0.791	0.585	0.590	0.534	0.703
	800	0.959	0.957	0.874	0.877	0.855	0.952
IV	200	0.440	0.271	0.101	0.061	0.044	0.056
	400	0.668	0.480	0.115	0.065	0.062	0.085
	800	0.947	0.804	0.150	0.048	0.046	0.089
V	200	0.799	0.573	0.135	0.069	0.103	0.092
	400	0.960	0.846	0.206	0.068	0.135	0.140
	800	1.000	0.993	0.292	0.054	0.282	0.211
VI	200	0.803	0.587	0.148	0.063	0.077	0.053
	400	0.978	0.903	0.199	0.052	0.082	0.064
	800	1.000	0.995	0.263	0.033	0.141	0.083

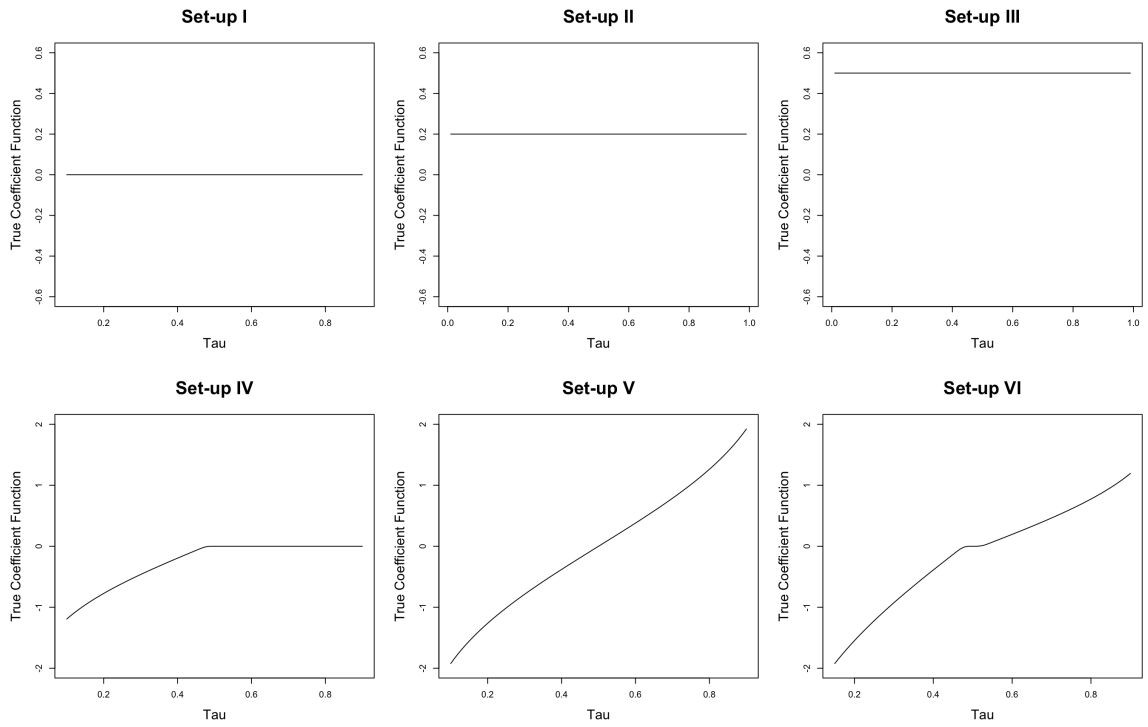


Figure 2.1: The true coefficient function for all simulation set-ups.

from a cohort of 191 incident dialysis patients (Kutner et al., 2002). In this dataset, time to death is censored in about 35% of dialysis patients due to either renal transplantation or end of the study as of December 31, 2005. In our analysis, we consider six potential prognostic factors (or covariates), which include age in years (AGE), indicator of reporting fish consumption over the first year of dialysis (FISHH), the indicator for baseline HD dialysis modality (BHDPD); whether the patient has severe symptoms of restless leg syndrome or not (BLEGS); whether or not education level is equal or higher than college (HIEDU); and the indicator of being in the black race group (BLACK). In our analyses, we standardize AGE by subtracting the sample mean and then dividing the resulting quantity by the sample standard deviation.

As a part of exploratory analyses, we check the proportional hazard assumption for each covariate based on Grambsch and Therneau (1994)'s method, using the R function `cox.zph()` in the R package `survival`. The p-values corresponding to AGE, FISHH, BHDPD, BLEGS, HIEDU and BLACK are 0.43, 0.63, 0.55, 0.0006, 0.047 and

0.0004, respectively. These results suggest that the proportional hazard assumption may be violated for BLEGS, HIEDU and BLACK.

We fit model (2.1) for time to death (i.e. T) with each covariate separately. We set $[\tau_L, \tau_U]$ as $[0.1, 0.6]$ for FISHH, BLGES, HIEDU, and BLACK, but set $[\tau_L, \tau_U]$ as $[0.1, 0.54]$ and $[0.1, 0.49]$ respectively for AGE and BHDPD. This is because the estimation of $\beta_0^{(1)}(\tau)$ based on Peng and Fine (2009) does not converge for some τ's larger than 0.54 and 0.49 when $\tilde{Z}$ is AGE or BHDPD. Figure 2.2 presents the estimated coefficients with the pointwise 95% confidence interval across $\tau \in [\tau_L, \tau_U]$. It is suggested by Figure 2.2 that AGE and BLACK have strong and persistent effects across all or most quantiles of time to death, implying an apparent survival advantage for younger or black patients. For each of the rest covariates, FISHH, BHDPD, BLEGS, or HIEDU, we note a partial effect pattern. For example, FISHH and BLEGS may only impact some lower quantiles of the survival time. BHDPD and HIEDU may only have quantile effects in the τ-intervals, $[0.15, 0.3]$ and $[0.3, 0.4]$, respectively. These observations suggest the presence of dynamic covariate effects as well as the need to appropriately accommodate such dynamic covariate effects.

To evaluate each potential prognostic factor considered, we apply the proposed testing procedures, GST and GIT, along with the benchmark methods, CPH (Wald) and CQR (Wald), as described in Section 2.2. Table 2.2 summarizes the p values obtained from different methods for evaluating each covariate. We note that all tests consistently suggest a strong effect of AGE or BLACK on the survival time. The locally concerned quantile regression tests, CQR (Wald), reveal τ-varying effects of FISHH, BHDPD, BLEGS, and HIEDU. For example, BLEGS may significantly influence the 10th and 20th quantiles of the survival time but not the 30th, 40th, 50th, 60th of quantiles. HIEDU may also have a partial effect, influencing some quantiles, such as the 30th and 40th quantiles, but not the other quantiles. The classic Cox regression based test, CPH (Wald), however, fails to capture the partial effects of

Table 2.2: A summary of p-values for each covariates with different methods.

Covariate	Proposed Test		CQR (Wald)						CPH (Wald)
	GST	GIT	$\tau = 0.1$	$\tau = 0.2$	$\tau = 0.3$	$\tau = 0.4$	$\tau = 0.5$	$\tau = 0.6$	
AGE	<0.001	<0.001	0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
FISHH	<0.001	0.018	0.037	0.055	0.036	0.214	0.473	0.316	0.026
BHDPD	<0.001	0.005	0.090	0.021	0.152	0.228	0.229	0.030	0.008
BLEGS	<0.001	0.001	<0.001	0.001	0.062	0.082	0.091	0.507	0.349
HIEDU	0.013	0.093	0.596	0.137	0.003	0.032	0.068	0.241	0.245
BLACK	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001

BLEGS and HIEDU. The p values for testing the effect of BLEGS and HIEDU based on CPH (Wald) are 0.35 and 0.25 respectively. This is possibly caused by imposing a restrictive static view on how a covariate can influence the survival time. In contrast, the proposed GIT and GST, through simultaneously examining covariate effects at quantile levels $[\tau_L, \tau_U]$, are able to detect the partial effect of BLEGS, with small p values ≤ 0.001 and to suggest a trend toward the association between HIEDU and the survival time, with marginal p values 0.01 and 0.09. The proposed GIT and GST also provide some evidence for the dynamic prognostic value of FISHH and BHDPD for dialysis survival. For example, as suggested by CQR (Wald), fish consumption in the first year may benefit dialysis patients with shorter survival time but may manifest little effect on the long-term survival. In general, our analysis results are consistent with the analyses of Peng and Huang (2008) based on multivariate censored quantile regression model. This example demonstrates the good practical utility of the proposed methods when varying covariate effects are present.

2.4 Remarks

In this chapter we develop a new testing framework for evaluating a survival prognostic factor. The main thrust of the new framework lies in its flexibility of accommodating a dynamic covariate effect, which is achieved through adapting the spirit of globally concerned quantile regression. Our testing procedures are conveniently developed based on existing results on fitting a working quantile regression model

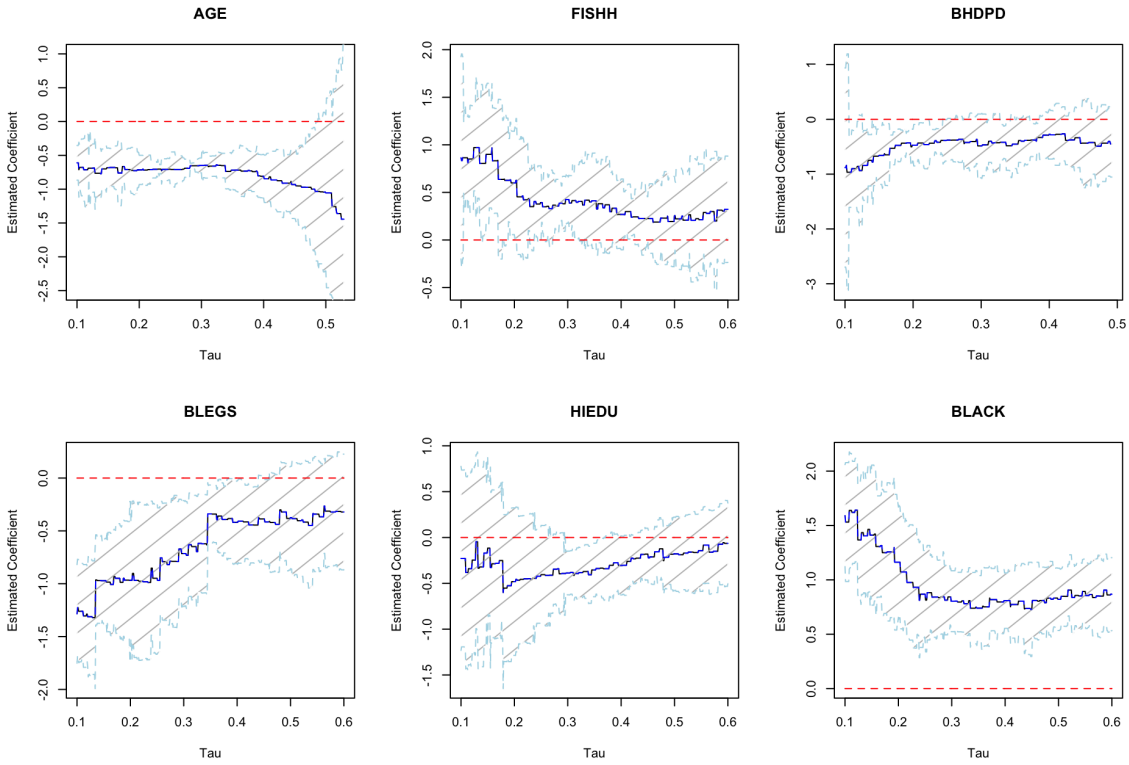


Figure 2.2: The estimated coefficient with the 95% confidence interval for the covariates based on the censored quantile regression model on the dialysis data.

with randomly censored data. It is important to note that the validity of the testing procedures does not require that the working model is the true model. Moreover, the proposed methods can be readily extended to handle more complex survival outcomes, such as time to event subject to competing risks.

Also, we would like to point out that $Q_T(\tau|\widetilde{Z}) = Q_T(\tau)$ for $\tau \in (0,1)$ implies the statistical independence between T and $\widetilde{Z}$. Nevertheless, in this chapter, we confine our attention to H_0^* with τ_U less than 1. This is because right censoring typically precludes the information on the upper tail of the distribution of T, and thus $Q_T(\tau)$ or $Q_T(\tau|\tilde{Z})$ can become non-identifiable as τ approaches 1. The null hypothesis H_0^* entails a weaker version of the independence between T and $\tilde{Z}$ that can be better assessed with right censored data. Rejecting H_0^* can provide evidence for the dependence between T and $\tilde{Z}$, while accepting H_0^* may not sufficiently indicate the independence between T and $\tilde{Z}$.

Another commendable extension is to generalize the current null hypothesis and testing procedures to permit evaluating multiple prognostic factors simultaneously. This chapter also lays a key foundation for developing a nonparametric screening method for helping identify useful prognostic factors among a large number of candidates. These extensions will be reported in the next chapter.

2.5 Appendix

2.5.1 Lemma A2.1 and the Proof

Lemma A2.1. *Suppose the conditional distribution function of T given $\widetilde{Z} = \widetilde{z}$ is continuous and strictly monotone for all possible values of $\widetilde{z}$. Then $Q_T(\tau|\widetilde{Z}) = Q_T(\tau)$ for $\tau \in [\tau_L, \tau_U]$ is equivalent to model* (2.1) *holds with $\Delta = [\tau_L, \tau_U]$ and $\beta_0^{(1)}(\tau) = 0$ for $\tau \in [\tau_L, \tau_U]$.*

Proof of Lemma A2.1. Suppose we have $Q_T(\tau|\widetilde{Z}) = Q_T(\tau)$ for $\tau \in [\tau_L, \tau_U]$. It

is clear that for $\tau \in [\tau_L, \tau_U]$, we can write $Q_T(\tau|\widetilde{Z}) = \exp\{\boldsymbol{Z}^\intercal \boldsymbol{\theta}_0(\tau)\}$ with $\boldsymbol{\theta}_0(\tau) = (\log Q_T(\tau), 0)^\intercal$. This means that model (2.1) holds with $\Delta = [\tau_L, \tau_U]$ and $\beta_0^{(1)}(\tau) = 0$ for $\tau \in [\tau_L, \tau_U]$.

Suppose model (2.1) holds with $\Delta = [\tau_L, \tau_U]$ and $\beta_0^{(1)}(\tau) = 0$ for $\tau \in [\tau_L, \tau_U]$. This means, $Q_T(\tau|\widetilde{Z}) = \exp\{\beta_0^{(0)}(\tau)\}$ for $\tau \in [\tau_L, \tau_U]$. Given that the conditional distribution function of T given $\widetilde{Z}$ is continuous and strictly monotone, it follows from the definition of $Q_T(\tau|\widetilde{Z})$ that $\Pr(T \leq \exp\{\beta_0^{(0)}(\tau)\}|\widetilde{Z}) = \tau$ for $\tau \in [\tau_L, \tau_U]$. Taking expectation on both sides of this equality with respect to $\widetilde{Z}$, we then get $\Pr(T \leq \exp\{\beta_0^{(0)}(\tau)\} = \tau$ for $\tau \in [\tau_L, \tau_U]$. Given the continuity and strict monotonicity of the distribution function of T, which is implied by the continuity and strict monotonicity of the conditional distribution function of T given $\widetilde{Z}$, this implies that $\exp\{\beta_0^{(0)}(\tau)\} = Q_T(\tau)$. Thus, $Q_T(\tau|\widetilde{Z}) = Q_T(\tau)$ for $\tau \in [\tau_L, \tau_U]$. This completes the proof of Lemma A2.1.

2.5.2 Asymptotic Properties Without Assuming the Censored Quantile Regression Model

We assume the following regularity conditions:

Condition 2.1. *There exist a constant v such that $P(C = v) > 0$ and $P(C > v) = 0$.*

Condition 2.2. *$\tilde{Z}$ is uniformly bounded, i.e. $\sup_i |\widetilde{Z}_i| < \infty$.*

Condition 2.3. *(i) $\widetilde{\boldsymbol{\theta}}(\tau)$ is Lipschitz continuous for $\tau \in [\tau_L, \tau_U]$; (ii) $f(y|\boldsymbol{z})$ is bounded above uniformly in y and $\boldsymbol{z}$, where $f(y|\boldsymbol{z})$ denotes the conditional density of X given $\boldsymbol{Z} = \boldsymbol{z}$.*

Condition 2.4. *For some $\rho_0 > 0$ and $c_0 > 0$,$\inf_{\boldsymbol{b} \in \mathcal{B}(\rho_0)} eigmin\boldsymbol{A}(\boldsymbol{b}) \geq c_0$, where $\mathcal{B}(\rho) = \{\boldsymbol{b} \in R^2 : \inf_{\tau \in [\tau_L, \tau_U]} ||\boldsymbol{b} - \widetilde{\boldsymbol{\theta}}(\tau)|| \leq \rho\}$ and $\boldsymbol{A}(\boldsymbol{b}) = E[\boldsymbol{Z}\boldsymbol{Z}^\intercal f(\boldsymbol{Z}^\intercal \boldsymbol{b}|\boldsymbol{Z})]$. Here $||\cdot||$ is the Euclidean norm and $eigmin\boldsymbol{A}(\boldsymbol{b})$ represents the minimal eigenvalue of $\boldsymbol{A}(\boldsymbol{b})$.*

Condition 2.1 is adopted to simplify the theoretical arguments to ensure that $\widehat{G}(\cdot)$ is consistent for $G(\cdot)$. This condition is usually satisfied in studies subject to administrative censoring. Condition 2.2 imposes covariate boundedness. Condition 2.3 assumes that the limit coefficient process is smooth and the conditional density distribution is bounded and smooth. Condition 2.4 requires that the asymptotic limit of $\boldsymbol{U}_n(\boldsymbol{b}, \tau)$ is strictly convex in a neighborhood of $\widetilde{\boldsymbol{\theta}}(\tau)$ for $\tau \in [\tau_L, \tau_U]$, implying the uniqueness of the solution to $\mu(\boldsymbol{b}, \tau) \equiv E\{\boldsymbol{Z} I(\log T \leq \boldsymbol{Z}^\intercal \boldsymbol{b}) - \tau)\} = 0$. This plays a critical role in establishing the uniform convergence of $\widehat{\boldsymbol{\theta}}(\tau)$ to $\widetilde{\boldsymbol{\theta}}(\tau)$.

Theorem A2.1. *Under regularity conditions 2.1–2.4, we have*

$$\lim_{n\to\infty} \sup_{\tau\in[\tau_L,\tau_U]} ||\widehat{\boldsymbol{\theta}}(\tau) - \widetilde{\boldsymbol{\theta}}(\tau)|| \to_p 0.$$

Theorem A2.2. *Under regularity conditions 2.1–2.4, we have* $\sqrt{n}(\widehat{\boldsymbol{\theta}}(\tau) - \widetilde{\boldsymbol{\theta}}(\tau))$ *converge weakly to a mean zero Gaussian process for* $\tau \in [\tau_L, \tau_U]$ *with covariance*

$$\widetilde{\boldsymbol{\Phi}}(\tau', \tau) = E\{\widetilde{\boldsymbol{\xi}}_1(\tau')\widetilde{\boldsymbol{\xi}}_1(\tau)^\intercal\}.$$

The proofs of Theorems A2.1 and A2.2 closely resemble the proofs in Peng and Fine (2009) and thus are omitted.

2.5.3 Proofs of Theorem 2.1 and 2.2

We assume one additional regularity condition:

Condition 2.5. $\inf_{\tau\in[\tau_L,\tau_U]} \sigma^{(1)}(\tau) > 0$, *where* $\{\sigma^{(1)}(\tau)\}^2$ *is the second diagonal element of* $\widetilde{\boldsymbol{\Phi}}(\tau, \tau)$.

Proof of Theorem 2.1

Following the lines of Peng and Fine (2009), we can show that the sample-based variance estimation procedure presented in Section 2.1.2 provides consistent variance

estimation, which implies $\sup_{\tau\in(\tau_L,\tau_U]}|\widehat{\sigma}_n^{(1)}(\tau)-\sigma^{(1)}(\tau)|\to_p 0$.

Note that under the null hypothesis H_0^*, we have $\widetilde{\beta}^{(1)}(\tau)=0$ and consequently,

$$\begin{aligned} n^{1/2}\widehat{R}(\tau) &= \frac{n^{1/2}\{\widehat{\beta}^{(1)}(\tau)-\widetilde{\beta}^{(1)}(\tau)\}}{\widehat{\sigma}_n^{(1)}(\tau)} \\ &= \frac{n^{1/2}\{\widehat{\beta}(\tau)-\widetilde{\beta}^{(1)}(\tau)\}}{\sigma^{(1)}(\tau)}\left(\frac{\sigma^{(1)}(\tau)}{\widehat{\sigma}_n^{(1)}(\tau)}-1\right)+\frac{n^{1/2}\{\widehat{\beta}^{(1)}(\tau)-\widetilde{\beta}^{(1)}(\tau)\}}{\sigma^{(1)}(\tau)} \quad (2.5)\end{aligned}$$

By Theorem A2.2, $n^{1/2}\{\widehat{\beta}^{(1)}(\tau)-\widetilde{\beta}^{(1)}(\tau)\}/\sigma^{(1)}(\tau)$ converges weakly to a mean zero Gaussian process $\mathcal{X}^{(1)}(\tau)$ with covariance process

$$\widetilde{\Phi}^{(1)}(\tau,\tau')=\frac{\widetilde{\mathbf{\Phi}}^{(2,2)}(\tau,\tau')}{\sigma^{(1)}(\tau)\sigma^{(1)}(\tau')},$$

where $\widetilde{\mathbf{\Phi}}^{(2,2)}(\tau,\tau')$ denotes the element in the second row and the second column of $\widetilde{\mathbf{\Phi}}(\tau,\tau')$. In addition, condition 2.5 and $\sup_{\tau\in(\tau_L,\tau_U]}|\widehat{\sigma}_n^{(1)}(\tau)-\sigma^{(1)}(\tau)|\to_p 0$ imply $\sup_{\tau\in(\tau_L,\tau_U]}\left|\frac{\sigma^{(1)}(\tau)}{\widehat{\sigma}_n^{(1)}(\tau)}-1\right|\to_p 0$. Applying the result of Theorem A2.2 and the Slutsky's Theorem (line 11 of Example 1.4.7 in Boucheron et al. (2013)) to (2.5), we then get $n^{1/2}\widehat{R}(\tau)\to_d\mathcal{X}^{(1)}(\tau)$ in $l^\infty(\mathcal{F}_T)$, where $l^\infty(S)$ is the collection of all bounded functions $f:S\mapsto R$ for any index set S and $\mathcal{F}_T=\{\frac{\widetilde{\boldsymbol{\xi}}_1^{(1)}(\boldsymbol{c},\tau)}{\sigma^{(1)}(\tau)},\boldsymbol{c}\in R^2,\tau\in[\tau_L,\tau_U]\}$. Then, by the extended continuous mapping theorem (Theorem 1.11.1 in van der Vaart et al. (1996)), we can establish the limiting null distribution for $\widehat{T}_{sup}^{(1)}$ and $\widehat{T}_{inte}^{(1)}$ as

$$\begin{aligned} \widehat{T}_{sup}^{(1)} &= \sup_{\tau\in[\tau_L,\tau_U]}\left|\frac{n^{1/2}\widehat{\beta}^{(1)}(\tau)}{\widehat{\sigma}_n^{(1)}(\tau)}\right|=\sup_{\tau\in[\tau_L,\tau_U]}\left|n^{1/2}\widehat{R}(\tau)\right|\to_d\sup\{|\mathcal{X}^{(1)}(\tau)|,\tau\in[\tau_L,\tau_U]\}, \\ \widehat{T}_{inte}^{(1)} &= \int_{\tau_L}^{\tau_U}\left|\frac{n^{1/2}\widehat{\beta}^{(1)}(\tau)}{\widehat{\sigma}_n^{(1)}(\tau)}\right|^2 d\tau=\int_{\tau_L}^{\tau_U}\left|n^{1/2}\widehat{R}(\tau)\right|^2 d\tau\to_d\int_{\tau_L}^{\tau_U}\{\mathcal{X}^{(1)}(\tau)\}^2 d\tau. \end{aligned}$$

This completes the proof of Theorem 2.1.

Proof of Theorem 2.2

We first investigate the asymptotic limit of $\widehat{T}_{sup}^{(1)}$ under the alternative hypothesis

$H_{a,1}$. Simple algebra shows that

$$\begin{aligned}\widehat{T}^{(1)}_{sup} = & \sup_{\tau\in[\tau_L,\tau_U]}\left|\frac{n^{1/2}\widehat{\beta}^{(1)}(\tau)}{\widehat{\sigma}^{(1)}_n(\tau)}\right| = \sup_{\tau\in[\tau_L,\tau_U]}\left|\frac{n^{1/2}\widetilde{\beta}^{(1)}(\tau)}{\widehat{\sigma}^{(1)}_n(\tau)} + \frac{n^{1/2}(\widehat{\beta}^{(1)}(\tau)-\widetilde{\beta}^{(1)}(\tau))}{\widehat{\sigma}^{(1)}_n(\tau)}\right| \\ \geq & \sup_{\tau\in[\tau_L,\tau_U]}\left|\frac{n^{1/2}\widetilde{\beta}^{(1)}(\tau)}{\widehat{\sigma}^{(1)}_n(\tau)}\right| - \sup_{\tau\in[\tau_L,\tau_U]}\left|\frac{n^{1/2}(\widehat{\beta}^{(1)}(\tau)-\widetilde{\beta}^{(1)}(\tau))}{\widehat{\sigma}^{(1)}_n(\tau)}\right| \equiv \widehat{T}^{(1)}_{sup,1} - \widehat{T}^{(1)}_{sup,2}.\end{aligned}$$

By the extended continuous mapping theorem, we can show that the $\widehat{T}^{(1)}_{sup,2}$ converges in distribution to $\sup_{\tau\in[\tau_L,\tau_U]}|\mathcal{X}^{(1)}(\tau)|$ and thus is $O_p(1)$. At the same time, given $\sup_{\tau\in(\tau_L,\tau_U]}|\widehat{\sigma}^{(1)}_n(\tau)-\sigma^{(1)}(\tau)| \to_p 0$, under condition 2.5, we get $n^{-1/2}\widehat{T}^{(1)}_{sup,1} \to_p \nu_0$, where $\nu_0 = \sup_{\tau\in[\tau_L,\tau_U]}\left|\frac{\widetilde{\beta}^{(1)}(\tau)}{\sigma^{(1)}(\tau)}\right|$.

Under the alternative hypothesis $H_{a,1}$ and condition 2.5, we have $\nu_0 > 0$, and hence $P(n^{-1/2}\widehat{T}^{(1)}_{sup,1} > \nu_0/2) \to P(\nu_0 > \nu_0/2) = 1$ as $n \to \infty$. Furthermore, for any $a > 0$, we have $n^{-1/2}\widehat{T}^{(1)}_{sup,2} + a\cdot n^{-1/2} = o_p(1)$, which implies $P(n^{-1/2}\widehat{T}^{(1)}_{sup,2} + a\cdot n^{-1/2} > \nu_0/2) \to 0$ as $n \to \infty$. Note that

$$\begin{aligned}P(\widehat{T}^{(1)}_{sup} > a) & \geq P(n^{-1/2}\widehat{T}^{(1)}_{sup,1} > n^{-1/2}\widehat{T}^{(1)}_{sup,2} + a\cdot n^{-1/2}) \\ & \geq P(n^{-1/2}\widehat{T}^{(1)}_{sup,1} > \nu_0/2) - P(n^{-1/2}\widehat{T}^{(1)}_{sup,2} + a\cdot n^{-1/2} > \nu_0/2).\end{aligned}$$

It then follows that $P(\widehat{T}^{(1)}_{sup} > a) \to 1$ as $n \to \infty$ under the alternative hypothesis $H_{a,1}$. This immediately implies that $\widehat{T}^{(1)}_{sup}$ is a consistent test against $H_{a,1}$ because $P(\widehat{T}^{(1)}_{sup} > C_{sup,\alpha}) \to 1$ as $n \to \infty$ given $H_{a,1}$ holds, where $C_{sup,\alpha}$ denotes the α-level critical value determined upon the limit null distribution of $\widehat{T}^{(1)}_{sup}$, which is greater than 0.

Next, we consider $\widehat{T}^{(1)}_{inte}$ under the alternative hypothesis $H_{a,2}$. Write $\widehat{T}^{(1)}_{inte}$ as

$$\begin{aligned}
\widehat{T}^{(1)}_{inte} = & \int_{\tau_L}^{\tau_U} \left| \frac{n^{1/2}\widehat{\beta}^{(1)}(\tau)}{\widehat{\sigma}^{(1)}_n(\tau)} \right|^2 d\tau = \int_{\tau_L}^{\tau_U} \left| \frac{n^{1/2}\widetilde{\beta}^{(1)}(\tau)}{\widehat{\sigma}^{(1)}_n(\tau)} - \frac{n^{1/2}(\widetilde{\beta}^{(1)}(\tau) - \widehat{\beta}^{(1)}(\tau))}{\widehat{\sigma}^{(1)}_n(\tau)} \right|^2 d\tau \\
\geq & \int_{\tau_L}^{\tau_U} \left| \frac{n^{1/2}\widetilde{\beta}^{(1)}(\tau)}{\widehat{\sigma}^{(1)}_n(\tau)} \right|^2 d\tau - \int_{\tau_L}^{\tau_U} 2 \left| \frac{n^{1/2}\widetilde{\beta}^{(1)}(\tau)}{\widehat{\sigma}^{(1)}_n(\tau)} \right| \cdot \left| \frac{n^{1/2}(\widetilde{\beta}^{(1)}(\tau) - \widehat{\beta}^{(1)}(\tau))}{\widehat{\sigma}^{(1)}_n(\tau)} \right| d\tau \\
\equiv & \widehat{T}^{(1)}_{inte,1} - \widehat{T}^{(1)}_{inte,2}.
\end{aligned}$$

By the continuous mapping theorem, combined with $\sup_{\tau\in(\tau_L,\tau_U]} |\widehat{\sigma}^{(1)}_n(\tau) - \sigma^{(1)}(\tau)| \to_p 0$ and condition 2.5, we get $n^{-1}\widehat{T}^{(1)}_{inte,1} \to_p \nu^*_0$, where $\nu^*_0 = \int_{\tau_L}^{\tau_U} \left| \frac{\widetilde{\beta}^{(1)}(\tau)}{\sigma^{(1)}(\tau)} \right|^2 d\tau$, and

$$n^{-1/2}\widehat{T}^{(1)}_{inte,2} \to_d \int_{\tau_L}^{\tau_U} 2 \left| \frac{\widetilde{\beta}^{(1)}(\tau)}{\sigma^{(1)}(\tau)} \right| \cdot \{\mathcal{X}^{(1)}(\tau)\} d\tau$$

and thus $O_p(1)$. By condition 2.5, the alternative hypothesis $H_{a,2}$ implies $\nu^*_0 > 0$. Then following the same arguments for showing $P(\widehat{T}^{(1)}_{sup} > a) \to 1$ for any $a > 0$ based on the results that $n^{-1/2}\widehat{T}^{(1)}_{sup,1} \to_p \nu_0 > 0$ and $\widehat{T}^{(1)}_{sup,2} = O_p(1)$, we can prove that $P(n^{-1/2}\widehat{T}^{(1)}_{inte} > a) \to 1$ as $n \to \infty$ for any $a > 0$ under $H_{a,2}$. This implies that $P(\widehat{T}^{(1)}_{inte} > a) \to 1$ as $n \to \infty$ for any $a > 0$ under $H_{a,2}$. Therefore, $\widehat{T}^{(1)}_{inte}$ is a consistent test against the alternative hypothesis $H_{a,2}$.

2.5.4 Justification for the Proposed Resampling Procedure

Given the observed data denoted by $\{\boldsymbol{O}_i\}^n_{i=1} \equiv \{(X_i, \delta_i, \tilde{Z}_i)\}^n_{i=1}$, since $\{\iota^b_i\}^n_{i=1}$ are i.i.d. standard normal random variables, we have

$$\begin{aligned}
& E\left\{ \frac{n^{-1/2}\sum_{i=1}^n \widehat{\xi}^{(1)}_i(\tau)\iota^b_i}{\widehat{\sigma}^{(1)}_n(\tau)} \cdot \frac{n^{-1/2}\sum_{i=1}^n \widehat{\xi}^{(1)}_i(\tau')\iota^b_i}{\widehat{\sigma}^{(1)}_n(\tau')} \middle| \{\boldsymbol{O}_i\}^n_{i=1} \right\} \\
= & \; n^{-1}\sum_{i=1}^n \frac{\widehat{\xi}^{(1)}_i(\tau)\widehat{\xi}^{(1)}_i(\tau')}{\widehat{\sigma}^{(1)}_n(\tau)\widehat{\sigma}^{(1)}_n(\tau')} \to_p \widetilde{\Phi}^{(1)}(\tau, \tau').
\end{aligned}$$

By the arguments of Lin et al. (1993), the distribution of $n^{-1/2}\sum_{i=1}^{n}\widehat{\xi}_i^{(1)}(\tau)\iota_i^b/\widehat{\sigma}_n^{(1)}(\tau)$ converges weakly to $\mathcal{X}^{(1)}(\tau)$, the same limit as that of $n^{1/2}\{\widehat{\beta}^{(1)}(\tau)-\widetilde{\beta}^{(1)}(\tau)\}/\widehat{\sigma}_n^{(1)}(\tau)$, for almost all realizations of $\{\boldsymbol{O}_i\}_{i=1}^n$. Applying the extended continuous mapping theorem as in the proof of Theorem A2.2, we have that under H_0^*, the conditional distribution of $\widehat{T}_{sup,b}^{(1)}$ (or $\widehat{T}_{inte,b}^{(1)}$) given the observed data is asymptotically equivalent to the unconditional distributions of $T_{sup}^{(1)}$ (or $T_{inte}^{(1)}$). This justifies using the resampling procedure in Section 2.1.4 to obtain the p values of the proposed tests.

2.5.5 Additional simulation results

Table A2.1: Empirical rejection rate for the uncensored case based on 1000 simulations.

Set-up	n	Proposed Test		CQR (Wald)			CPH (Wald)
		GST	GIT	$\tau = 0.4$	$\tau = 0.5$	$\tau = 0.6$	
I	200	0.098	0.070	0.055	0.052	0.056	0.048
	400	0.093	0.075	0.069	0.064	0.060	0.047
	800	0.076	0.058	0.053	0.053	0.048	0.061
II	200	0.215	0.156	0.104	0.108	0.108	0.121
	400	0.275	0.216	0.162	0.156	0.139	0.183
	800	0.420	0.372	0.276	0.265	0.238	0.328
III	200	0.541	0.478	0.344	0.374	0.337	0.456
	400	0.790	0.771	0.589	0.595	0.590	0.745
	800	0.958	0.961	0.883	0.886	0.873	0.963
IV	200	0.378	0.250	0.074	0.045	0.049	0.060
	400	0.656	0.476	0.101	0.056	0.055	0.049
	800	0.935	0.808	0.118	0.034	0.045	0.085
V	200	0.618	0.452	0.106	0.057	0.121	0.428
	400	0.939	0.828	0.169	0.071	0.165	0.737
	800	1.000	0.994	0.255	0.041	0.313	0.968
VI	200	0.729	0.543	0.095	0.047	0.088	0.228
	400	0.971	0.898	0.154	0.048	0.097	0.446
	800	1.000	0.995	0.243	0.020	0.154	0.756

Table A2.2: Empirical rejection rate for the proposed test with different choices of $\mathcal{U}$ on the six set-ups subject to 15% censoring based on 1000 simulations.

Set-up	n	$\mathcal{U} = \{1, \ldots, 3\}$		$\mathcal{U} = \{1, \ldots, 6\}$		$\mathcal{U} = \{1, \ldots, 12\}$	
		GST	GIT	GST	GIT	GST	GIT
I	200	0.128	0.074	0.091	0.067	0.092	0.065
	400	0.126	0.079	0.086	0.067	0.081	0.063
	800	0.112	0.060	0.080	0.059	0.072	0.058
II	200	0.287	0.178	0.228	0.161	0.225	0.158
	400	0.359	0.231	0.283	0.218	0.256	0.206
	800	0.472	0.379	0.415	0.369	0.376	0.361
III	200	0.666	0.549	0.593	0.510	0.585	0.513
	400	0.841	0.789	0.779	0.773	0.761	0.757
	800	0.975	0.964	0.956	0.956	0.940	0.957
IV	200	0.427	0.257	0.364	0.242	0.362	0.243
	400	0.702	0.490	0.666	0.470	0.649	0.471
	800	0.952	0.808	0.942	0.811	0.936	0.808
V	200	0.695	0.478	0.649	0.452	0.649	0.446
	400	0.962	0.850	0.948	0.831	0.944	0.827
	800	1.000	0.994	1.000	0.991	1.000	0.993
VI	200	0.768	0.558	0.723	0.534	0.726	0.535
	400	0.981	0.902	0.971	0.894	0.970	0.892
	800	1.000	0.997	1.000	0.997	1.000	0.998

Chapter 3

Global Group Testing and Screening With Dynamic Effects

3.1 Problem and Motivation

In this part, we consider the problem to identify variables that are dynamically associated with the outcome, and to exclude irrelevant variables in the ultra-high dimensional settings with high confidence. Let Y denote a continuous outcome and let $\boldsymbol{X} = \{X^{(1)}, \ldots, X^{(p)}\}^T$ denote the vector of the observed covariates. Define

$$H_{0,j}: \; Q_Y(\tau \mid X^{(j)}) = Q_Y(\tau), \text{ a.s. for } \tau \in \Delta \subseteq (0,1).$$

Here and hereafter, for a general random vector $\boldsymbol{V}$, $Q_Y(\tau \mid \boldsymbol{V}) = \inf\{y : \; \text{pr}(Y \leq y|\boldsymbol{V}) \geq \tau\}$ denotes the conditional quantile function of Y given $\boldsymbol{V}$, and $Q_Y(\tau) = \inf\{y : \; \text{pr}(Y \leq y) \geq \tau\}$ denotes the unconditional quantile function of Y. When $X^{(j)}$ is continuous, $H_{0,j}$ refers to the interval quantile independence between Y and $X^{(j)}$ on quantile level intervals Δ and $[0,1]$ respectively for Y and $X^{(j)}$, as termed by Zhu et al. (2018). The consideration of $H_{0,j}$ confers a flexible view for defining relevant variables. In the multivariate setting, a covariate $X^{(j)}$ is considered as relevant or active if $Q_Y(\tau \mid \boldsymbol{X})$ functionally depends on $X^{(j)}$ for some $\tau \in \Delta \subseteq (0,1)$, where Δ is a pre-specified set of quantile levels. Under this view, the set of relevant variables is defined as $\mathcal{M}_\Delta = \{1 \leq r \leq p : \text{ there exists } \tau \in \Delta \text{ such that } Q_Y(\tau \mid \boldsymbol{X}) \text{ depends on } X^{(r)}\}$. The formulations of $H_{0,j}$ and $\mathcal{M}_\Delta$ take a global perspective to assess covariate effects throughout the range of the outcome distribution indexed by the quantile level interval Δ. Covariates in $\mathcal{M}_\Delta$ are permitted to have dynamic and non-additive effects across different ranges of the outcome.

Motivated from the first topic, we propose a new model-free strategy for tackling a generalized version of $H_{0,j}$ that concerns the outcome relevance of one or multiple covariates, which can be either continuous or discrete. Specifically, for a index set for J covariates, $G = \{r_1, \ldots, r_J\} \subseteq \{1, \ldots, p\}$, define $\boldsymbol{X}_G = \{X^{(r_1)}, \ldots, X^{(r_J)})^\intercal$. A null

hypothesis of our interest takes the form

$$H_{0,G}: \ Q_Y(\tau \mid \boldsymbol{X}_G) = Q_Y(\tau), \ \text{a.s.}, \quad \text{for} \ \ \tau \in \Delta \subseteq (0,1).$$

To address $H_{0,G}$, we propose to employ a "working" linear quantile regression model, which can help determine whether $H_{0,G}$ holds or not. We further utilize the proposed test statistic as the utility function to develop a new model-free variable screening procedure for ultra-high dimensional data. Given the flexibility of our test statistic in handling multiple covariates simultaneously, the new screening procedure can be performed with covariates pre-grouped by scientific needs or in a random manner for the benefit of saving computational time. As a useful by-product, we can readily transform the new screening procedure to perform conditional variable screening given some known relevant covariates under mild additional assumptions.

3.2 The Proposed Global Testing Framework

3.2.1 Formulation of the Proposed Test Statistic

Without loss of generality, let $G = \{1, \ldots, J\}$ and express the quantile interval Δ as $[\tau_L, \tau_U]$ with $0 \leq \tau_L < \tau_U \leq 1$. Define $\boldsymbol{Z} = (1, \boldsymbol{X}_G^\intercal)^\intercal$. As introduced in Section 3.1, the null hypothesis of interest is

$$H_{0,G}: \ Q_Y(\tau \mid \boldsymbol{X}_G) = Q_Y(\tau), \ \text{a.s.}, \ \ \text{for} \ \ \tau \in [\tau_L, \tau_U]. \tag{3.1}$$

The observed data consist of n independently identically distributed (i.i.d.) replicates of $(Y, \boldsymbol{Z})$, denoted as $\{(Y_i, \boldsymbol{Z}_i), i = 1, \ldots, n\}$. We assume that the conditional distribution of Y given $\boldsymbol{X}_G$ is continuous and strictly monotone and $E(\boldsymbol{Z}\boldsymbol{Z}^\intercal)$ is positive definite.

To address $H_{0,G}$, we uncover a useful connection between $H_{0,G}$ and a "working"

linear quantile regression model:

$$Q_Y(\tau \mid \boldsymbol{Z}) = \boldsymbol{Z}^\intercal \boldsymbol{\theta}_0(\tau), \quad \tau \in [\tau_L, \tau_U], \tag{3.2}$$

where $\boldsymbol{\theta}_0(\tau) = \{\alpha_0(\tau), \beta_0^{(1)}(\tau), \ldots, \beta_0^{(J)}(\tau)\}^\intercal$ is a vector of unknown coefficients. A key fact is that $H_{0,G}$ holds if and only if model (3.2) holds with $\beta_0^{(j)}(\tau) = 0$ for $\tau \in [\tau_L, \tau_U]$ for $j = 1, \ldots, J$; see Lemma A3.1 and its proof in Section 3.7.

To utilize this connection, we consider an estimator of $\boldsymbol{\theta}_0(\tau)$ defined as the solution to the standard score estimating equation for linear quantile regression,

$$\boldsymbol{S}_n(\boldsymbol{b}, \tau) = n^{-1/2} \sum_{i=1}^{n} \boldsymbol{Z}_i [I(Y_i \leq \boldsymbol{Z}_i^\intercal \boldsymbol{b}) - \tau] = 0, \tag{3.3}$$

with respect to $\boldsymbol{b}$ (Koenker and Bassett Jr, 1978), denoted by $\widehat{\boldsymbol{\theta}}(\tau) = \{\widehat{\alpha}_0(\tau), \widehat{\beta}^{(1)}, \ldots, \widehat{\beta}^{(J)}(\tau)\}^\intercal$. It is important to note that, without assuming the working model (3.2), $\widehat{\boldsymbol{\theta}}(\tau)$ may uniformly converge to $\widetilde{\boldsymbol{\theta}}(\tau) = \{\widetilde{\alpha}_0(\tau), \widetilde{\beta}^{(1)}, \ldots, \widetilde{\beta}^{(J)}(\tau)\}^\intercal$ over $\tau \in [\tau_L, \tau_U]$, where $\widetilde{\boldsymbol{\theta}}(\tau)$ is the solution to the equation, $\boldsymbol{\mu}(\boldsymbol{b}, \tau) = E[\boldsymbol{Z}\{I(Y \leq \boldsymbol{Z}^\intercal \boldsymbol{b}) - \tau\}] = 0$, with respect to $\boldsymbol{b} \in R^{J+1}$; see Theorem A3.1 in Section 3.7. By Lemma A3.2 in Section 3.7, the solution to $\boldsymbol{\mu}(\boldsymbol{b}, \tau) = 0$ uniquely exists and $H_{0,G}$ implies $\{\widetilde{\beta}^{(1)}(\tau), \ldots, \widetilde{\beta}^{(J)}(\tau)\}^\intercal = \boldsymbol{0}$. Motivated by these results, we propose to test the departure of $H_{0,G}$ by using the deviation of $\{\widetilde{\beta}^{(1)}(\tau), \ldots, \widetilde{\beta}^{(J)}(\tau)\}^\intercal$ from $\boldsymbol{0} \in R^J$ for $\tau \in [\tau_L, \tau_U]$.

Employing the connection between $H_{0,G}$ and the working model (3.2) permits leveraging existing inferential tools and software for quantile regression to facilitate the task of testing $H_{0,G}$ based on $\widehat{\boldsymbol{\theta}}(\cdot)$. It also provides an intuitive way to interpret $\{\widehat{\beta}^{(1)}(\tau), \ldots, \widehat{\beta}^{(J)}(\tau)\}^\intercal$, which would capture the covariate effects on the τ-th quantile of the outcome, possibly varying over τ, when the working model holds.

Specifically, we propose to construct the test statistic for $H_{0,G}$ as

$$\widehat{T}_{UC} = \max_{j \in G = \{1, \ldots, J\}} \widehat{T}_{inte}^{(j)},$$

where $\widehat{T}^{(j)}_{inte} = \int_{\tau_L}^{\tau_U} \left| n^{1/2}\widehat{\beta}^{(j)}(\tau)/\widehat{\sigma}^{(j)}_n(\tau) \right|^2 d\tau$ and $\widehat{\sigma}^{(j)2}_n(\tau)$ is the variance estimate for $n^{1/2}\{\widehat{\beta}^{(j)}(\tau) - \widetilde{\beta}^{(j)}(\tau)\}$ elaborated later. The construction of $\widehat{T}_{UC}$ reflects the idea of first utilizing the squared $\widetilde{\beta}^{(j)}(\tau)$ to capture the local influence of $X^{(j)}$ at the single τ, integrating the local effect over $\tau \in [\tau_L, \tau_U]$ to assess the global effect of $X^{(j)}$, and then taking the maximum global effect across all covariates. Such a test statistic shares a similar spirit of the Cramér-Von-Mises (C-V) test statistic and is expected to be sensitive to any departure of $(\widetilde{\beta}^{(1)}(\tau), \ldots, \widetilde{\beta}^{(J)}(\tau))^\intercal$ from the constant zero function. In Theorem 3.1, we establish the limit null distribution of $\widehat{T}_{UC}$. The proof is provided in Section 3.7.

Theorem 3.1. *Suppose the regularity conditions 3.3 and 3.4 in Section 3.7 hold. Under the null hypothesis $H_{0,G}$, we have*

$$\widehat{T}_{UC} \quad \rightarrow_d \max_{j=1,\ldots,J} \left\{ \int_{\tau_L}^{\tau_U} [\mathcal{X}^{(j)}(\tau)]^2 d\tau \right\},$$

where $\mathcal{X}^{(j)}(\tau)$ is a mean zero Gaussian process defined in Section 3.7, $j = 1, \ldots, J$.

3.2.2 The Proposed Global Testing Procedure

Given the connection between $\widetilde{\boldsymbol{\theta}}(\tau)$ and the working model (3.2), we can readily obtain $\widehat{\beta}^{(j)}(\tau)$ by using the `rq()` function in the R package `quantreg`. As detailed in Theorem A3.2 in Section 3.7, under certain regularity conditions, $n^{1/2}(\widehat{\boldsymbol{\theta}}(\tau) - \widetilde{\boldsymbol{\theta}}(\tau))$ converges weakly to a mean zero Gaussian process for $\tau \in [\tau_L, \tau_U]$ with covariance $\boldsymbol{\Phi}(\tau', \tau) = E\{\boldsymbol{\xi}_i(\tau')\boldsymbol{\xi}_i(\tau)^\intercal\}$, where the influence function $\boldsymbol{\xi}_i(\tau)$ is defined in Theorem A3.2. The asymptotic result allows us to obtain the variance estimate $\widehat{\sigma}^{(j)}_n(\tau)$ from adapting Peng and Fine (2009)'s sample-based inference procedure as outlined below:

(1.a) Compute $\widehat{\boldsymbol{\Sigma}}(\tau, \tau) = n^{-1} \sum_{i=1}^n \boldsymbol{Z}_i \boldsymbol{Z}_i^\intercal \{I[Y_i \leq \boldsymbol{Z}_i^\intercal \widehat{\boldsymbol{\theta}}(\tau)] - \tau\}^2$.

(1.b) Conduct eigenvalue eigenvector decomposition for $\widehat{\boldsymbol{\Sigma}}(\tau, \tau)$ using `eigen()` func-

tion in R to find the matrix $\boldsymbol{E}_n(\tau) = \{\boldsymbol{e}_{n,0}(\tau), \ldots, \boldsymbol{e}_{n,J}(\tau)\}$ such that $\{\boldsymbol{E}_n(\tau)\}^2 = \widehat{\boldsymbol{\Sigma}}(\tau, \tau)$.

(1.c) Solve the perturbed estimating equation $\boldsymbol{S}_n(\boldsymbol{c}, \tau) = \boldsymbol{e}_{n,j}(\tau)$ for $j = 1, \ldots, J$ and denote the solution as $\boldsymbol{S}_n^{-1}\{\boldsymbol{e}_{n,j}(\tau), \tau\}$.

(1.d) Calculate $\boldsymbol{D}_n(\tau) = \{\boldsymbol{S}_n^{-1}\{\boldsymbol{e}_{n,0}(\tau), \tau\} - \widehat{\boldsymbol{\theta}}(\tau), \ldots, \boldsymbol{S}_n^{-1}\{\boldsymbol{e}_{n,J}(\tau), \tau\} - \widehat{\boldsymbol{\theta}}(\tau)\}$. Compute an estimate for the asymptotic variance of $n^{1/2}\{\widehat{\boldsymbol{\theta}}(\tau) - \widetilde{\boldsymbol{\theta}}(\tau)\}$ as $\boldsymbol{V}_n(\tau) \equiv n\boldsymbol{D}_n^{\otimes 2}(\tau)$. Obtain $\widehat{\sigma}_n^{(j)2}(\tau)$ as the $j + 1^{\text{th}}$ diagonal component of $\boldsymbol{V}_n(\tau)$.

Remark 1: The above procedure remains valid if we replace $\boldsymbol{e}_{n,j}(\tau)$ with $u \cdot \boldsymbol{e}_{n,j}(\tau)$ for some constant u in step (1.c). Based on our numerical experiences, employing a properly selected u may help stabilize the variance estimation especially when the sample size is not large. An empirical strategy to choose u is detailed in Section 3.7.

The result in Theorem 3.1 indicates that the asymptotic null distribution of the proposed test statistic is non-standard. We develop a perturbation resampling procedure to obtain the p value from testing $H_{0,G}$ based on the proposed test statistic. The resampling procedure is outline as follow.

(2.a) Generate B independent sets of $\{\iota_i^b\}_{i=1}^n$, where $\{\iota_i^b\}_{i=1}^n$ are independent random variables from a standard normal distribution for $b = 1, \ldots, B$.

(2.b) Calculate $\widehat{\boldsymbol{\xi}}_i(\tau) = \{\widehat{A}(\widehat{\boldsymbol{\theta}}(\tau))\}^{-1}\boldsymbol{Z}_i[I(Y_i \leq \boldsymbol{Z}_i^{\intercal}\widetilde{\boldsymbol{\theta}}(\tau)) - \tau]$, where $\{\widehat{A}(\widehat{\boldsymbol{\theta}}(\tau))\}^{-1}$ is obtained from $\{\widehat{A}(\widehat{\boldsymbol{\theta}}(\tau))\}^{-1} = n^{1/2}\boldsymbol{D}_n(\tau)\boldsymbol{E}_n(\tau)^{-1}$.

(2.c) For $b = 1, \ldots, B$, calculate

$$\widehat{T}_{UC,b} = \max_{j=1,\ldots,J} \left\{ \int_{\tau_L}^{\tau_U} \left| n^{-1/2} \sum_{i=1}^{n} \widehat{\xi}_i^{(j)}(\tau)\iota_i^b / \widehat{\sigma}_n^{(j)}(\tau) \right|^2 d\tau \right\},$$

where $\widehat{\xi}_i^{(j)}(\tau)$ is the $j + 1^{\text{th}}$ component of $\widehat{\boldsymbol{\xi}}_i(\tau)$.

(2.d) The p value is calculated by $p_{UC} = \sum_{b=1}^{B} I(\widehat{T}_{UC,b} > \widehat{T}_{UC})/B$.

Similar resampling procedures were used in other settings, such as Lin et al. (1993), Li and Peng (2014), and Cui and Peng (2022). The key idea is to approximate the limit null distribution through perturbing the influence function $\boldsymbol{\xi}_i(\tau)$. The above resampling procedure is easy to implement without involving smoothing. Justification for this procedure is provided in Section 3.7.

In Theorem A3.3 in Section 3.7, we further show that the proposed test statistic $\widehat{T}_{UC}$ is consistent against the alternative hypothesis,

$$H_{a,G}: \text{For some } j_1 \in \{1,\ldots,J\}, \text{there exists } \tau \in [\tau_L, \tau_U] \text{ such that } |\widetilde{\beta}^{(j_1)}(\tau)| > 0.$$

This result suggests promising power of the proposed procedure for detecting departures from $H_{0,G}$.

3.3 Variable Screening in Ultra-high Dimensional Setting

3.3.1 The Proposed Unconditional Screening Framework

Consider the ultra-high dimensional setting, where $p = O(\exp(n^c))$ for a positive $c < 1$. Suppose the observed covariates are grouped as $(\boldsymbol{X}_{G_1}^\intercal, \ldots, \boldsymbol{X}_{G_L}^\intercal)^\intercal$, where $G_1, \ldots, G_L$ are non-overlapping covariate index sets and $\cup_{l=1}^{L} G_l = \{1,\ldots,p\}$. While both p and L may depend on the sample size n, we omit n from their notation for presentation simplicity. Assume that the sizes of G_l's ($l = 1,\ldots,L$) are finite and uniformly bounded; thus p and L are of the same asymptotic order.

In practice, the grouping of covariates may be motivated by scientific needs, for example, grouping genes according to biological pathways. The special case with $p = L$ corresponds to the regular scenario where no grouping is imposed to covariates. Thus, a unified definition of the set of relevant covariates, with or without grouping, is

$M_{[\tau_L,\tau_U]} = \{G_l : 1 \le l \le L \text{ and there exists } \tau \in [\tau_L, \tau_U] \text{ such that } Q_\tau(Y|\boldsymbol{X}) \text{ depends on } X_{G_l}\}$. Assume that the cardinality of $M_{[\tau_L,\tau_U]}$ is smaller than the sample size n. Let $\lfloor \cdot \rfloor$ and $\lceil \cdot \rceil$ denote the floor and ceiling operators respectively.

We propose the following variable screening procedure:

(3.a) Normalize $\boldsymbol{X}$ and Y.

(3.b) For each index set G_l, compute $\widehat{T}_{UC}$ for H_{0,G_l} and denote it by $w_{1,l},\ l = 1, \ldots, L$.

(3.c) Sort $\{1, \ldots, L\}$ according to $w_{1,l}$ in a decreasing order.

(3.d) Keep $\boldsymbol{X}_{G_l}$'s with $w_{1,l}$ greater than a pre-defined threshold ν_n or keep a pre-specified number (e.g., $\lfloor n/\log n \rfloor$) of covariates on the top of the list obtained from (3.c).

By the above procedure with some pre-determined threshold value ν_n, the set of remaining variables is defined as $\widehat{M}_{[\tau_L,\tau_U]} = \{G_l : 1 \le l \le L,\ w_{1,l} \ge \nu_n\}$.

When there is no particular scientific reasons to group covariates, we have the variable screening problem with $L = p$. In this case, we may consider an alternative two-step screening procedure that first filters covariates by randomly formed groups and then conducts a second-step single covariate screening. Our numerical investigation shows that such a two-step procedure may considerably reduce computational time while preserving similar screening performance. Specifically, the two-step screening procedure includes the following steps:

(4.a) Normalize $\boldsymbol{X}$ and Y.

(4.b) Perform the first-step group-level screening:

(i) Shuffle the index set of the covariates $\{1, \ldots, p\}$ to $\{r_1, \ldots, r_p\}$.

(ii) With a pre-determined group size S_G, compute $L = \lceil p/S_G \rceil$. Divide the first $(L-1) \cdot S_G$ covariates into $L-1$ groups of equal size S_G. The Lth

group includes the last $p - S_G \cdot (L-1)$ elements. Denote the resulting grouped covariates as $\{\boldsymbol{X}_{G_1}, \ldots, \boldsymbol{X}_{G_L}\}$.

(iii) Apply steps (3.b)–(3.d) to the grouped covariates $\{\boldsymbol{X}_{G_1}, \ldots, \boldsymbol{X}_{G_L}\}$ with a pre-sepecified threshold $\nu_{n,1}$ or a pre-specified number (e.g. $[n/\log n]$) of groups on the top of the list.

(4.c) Express the set of remaining variables from (4.b) in terms of individual covariates, $\{X^{(r_1)}, \ldots, X^{(r_M)}\}$, and then perform the second-step individual-level screening:

(i) Obtain $\widehat{T}_{UC}$ for $H_{0,\{r_m\}}$, denoted by $w_{2,m}$, for $m = 1, \ldots, M$.

(ii) Sort $\{X^{(r_1)}, \ldots, X^{(r_M)}\}$ according to $w_{2,m}$ in a decreasing order.

(iii) Keep the covariates with $w_{2,m}$ greater than a pre-sepecified threshold $\nu_{n,2}$ or keep a pre-specified number (e.g. $[n/\log n]$) of covariates on the top of the list obtained from (4.c) (ii).

With this two-step screening procedure, the set of remaining variables is defined as $\widehat{M}^G_{[\tau_L,\tau_U]} = \{r_m : 1 \le m \le M,\ w_{2,m} \ge \nu_{n,2}\}$.

We establish the sure screening property for the proposed unconditional screening procedures. Let $\widehat{T}^{(G_l)}_{UC}$ denote the proposed test statistic $\widehat{T}_{UC}$ for H_{0,G_l} and define

$$T^{(G_l)}_{UC} = \max_{j \in G_l} \int_{\tau_L}^{\tau_U} \left| n^{1/2} \widetilde{\beta}^{(j)}(\tau) / \sigma^{(j)}(\tau) \right|^2 d\tau,$$

where $\{\sigma^{(j)}(\tau)\}^2$ is the $j+1$th diagonal element of $\boldsymbol{\Phi}(\tau, \tau)$ defined in Theorem A3.2. In Theorem 3.2, we establish the exponential probability bounds for $|n^{-1}\widehat{T}^{(G_l)}_{UC} - n^{-1}T^{(G_l)}_{UC}|$, which serve as the key step to derive the sure screening property.

Theorem 3.2. *Given that the regularity conditions 3.1-3.4 in Section 3.7 hold. For*

any $c > 0$ *and* $1/4 < \zeta \leq 1/2$*, there exists positive constant* υ *and* η *such that*

$$pr(\max_{1\leq l\leq L} |n^{-1}\widehat{T}_{UC}^{(G_l)} - n^{-1}T_{UC}^{(G_l)}| \geq cn^{\zeta-1/2}) \leq p\upsilon \exp[-\eta n^{4\zeta-1} - \log(n^{\zeta-1/2})]$$

for sufficiently large n*.*

Next, we introduce Corollary 3.1 and Corollary 3.2 to establish the sure screening property of both the proposed one-step screening procedure outlined in (3.a)–(3.d) and the proposed two-step screening procedure outlined in (4.a)–(4.c), respectively.

Corollary 3.1 (Sure screening property for the one-step screening procedures)**.** *Suppose that the regularity conditions 3.1-3.5 in Section 3.7 hold. If we take the threshold value* $\nu_n = \delta^* n^{\zeta-1/2}$ *with* $\delta^* \leq \alpha_0/2$*, then there exists positive constant* a_1 *and* b_1*, such that*

$$pr(M_{[\tau_L,\tau_U]} \subseteq \widehat{M}_{[\tau_L,\tau_U]}) \geq 1 - S_{[\tau_L,\tau_U]} \cdot a_1 \exp[-b_1 n^{4\zeta-1} - \log(n^{\zeta-1/2})]$$

for sufficient large n*, where* $S_{\tau_L,\tau_U} = |M_{[\tau_L,\tau_U]}|$ *is the cardinality of* $M_{[\tau_L,\tau_U]}$*. In particular,* $pr(M_{[\tau_L,\tau_U]} \subseteq \widehat{M}_{[\tau_L,\tau_U]}) \to 1$ *as* $n \to \infty$*.*

Corollary 3.2 (Sure screening property for the two-step screening procedure)**.** *Suppose that the regularity conditions 3.1-3.4 and 3.6 in Section 3.7 hold. If we take the threshold value* $\nu_{n,1} = \delta^* n^{\zeta-1/2}$ *and* $\nu_{n,2} = \delta^{**} n^{\zeta-1/2}$ *with* $0 < \delta^* \leq \alpha_0/2$ *and* $0 < \delta^{**} \leq \alpha_0/2$*, respectively, then there exists positive constant* a_2 *and* b_2*, such that*

$$pr(M_{[\tau_L,\tau_U]} \subseteq \widehat{M}^G_{[\tau_L,\tau_U]}) \;\; \geq \;\; 1 - S_{[\tau_L,\tau_U]} \cdot a_2 \exp[-b_2 n^{4\zeta-1} - \log(n^{\zeta-1/2})]$$

for sufficient large n*, where* $S_{[\tau_L,\tau_U]} = |M_{[\tau_L,\tau_U]}|$ *is the cardinality of* $M_{[\tau_L,\tau_U]}$*. In particular,* $pr(M_{[\tau_L,\tau_U]} \subseteq \widehat{M}^G_{[\tau_L,\tau_U]}) \to 1$ *as* $n \to \infty$*.*

The proofs of Theorem 3.2, Corollary 3.1 and Corollary 3.2 are provided in Section 3.7.

3.3.2 A Generalization to Conditional Screening

In practice, a set of covariates may be known to relate to the outcome by existing knowledge. In many studies, assessing the relative importance of the other covariates in the presence of the known relevant covariates is of interest. This confers a conditional screening problem (Barut et al., 2016). By the proposed testing strategy, we can readily generalize the screening procedures presented in Section 3.3.1 to conduct conditional variable screening.

Let $\boldsymbol{X}_{\mathcal{C}}$ denote the set of relevant covariates known from prior knowledge, and denote the rest of covariates as $\boldsymbol{X}_{-\mathcal{C}}$. Suppose $\boldsymbol{X}_{-\mathcal{C}}$ is grouped as $\{\boldsymbol{X}_{G_{c,1}}, \ldots, \boldsymbol{X}_{G_{c,L_c}}\}$. When $L_c = q$, no grouping is imposed to $\boldsymbol{X}_{-\mathcal{C}}$. Adapting the global perspective taken in the proposed unconditional screening framework, we consider $\boldsymbol{X}_{G_{c,l}}$ as conditionally irrelevant to the outcome if $Q_Y(\tau \mid \boldsymbol{X}_{\mathcal{C}}, \boldsymbol{X})$ does not depend on $\boldsymbol{X}_{G_{c,l}}$ for $\tau \in [\tau_L, \tau_U]$. Under this view, screening out conditionally irrelevant covariates is naturally linked to the problem of testing $H_{c,G_{c,l}}: \ Q_Y(\tau \mid \boldsymbol{X}_{\mathcal{C}}, \boldsymbol{X}_{G_{c,l}}) = Q_Y(\tau \mid \boldsymbol{X}_{\mathcal{C}})$ for $\tau \in [\tau_L, \tau_U]$. We assume that $\boldsymbol{X}_{\mathcal{C}}$ has a known type of relationship with the outcome. For simplicity, we assume that $Q_Y(\tau \mid \boldsymbol{X}_{\mathcal{C}})$ is linearly related to $\boldsymbol{X}_{\mathcal{C}}$ for $\tau \in [\tau_L, \tau_U]$. Similar to the finding in the unconditional setting, $H_{c,G_{c,l}}$ holds if and only if the working linear quantile regression model

$$Q_Y(\tau \mid \boldsymbol{X}_{\mathcal{C}}, \boldsymbol{X}_{G_{c,l}}) = \alpha_c(\tau) + \boldsymbol{X}_{\mathcal{C}}{}^{\mathsf{T}}\boldsymbol{\beta}_{c,1} + \boldsymbol{X}_{G_{c,l}}^{\mathsf{T}}\boldsymbol{\beta}_{c,2}, \quad \tau \in [\tau_L, \tau_U], \tag{3.4}$$

holds with $\boldsymbol{\beta}_{c,2} = \mathbf{0}$. This fact naturally motivates the following conditional variable screening procedure:

(5.a) Normalize $\boldsymbol{X}$ and Y.

(5.b) For each index set $G_{c,l}$, compute a conditional test statistic $\widehat{T}_C$ for $H_{c,G_{c,l}}$, which is obtained in the same manner as that for $\widehat{T}_{UC}$ except that the working linear

quantile regression model includes $\boldsymbol{X}_{\mathcal{C}}$ in addition to $\boldsymbol{X}_{G_{c,l}}$. Denote the resulting $\widehat{T}_C$ by $w_{c,l},\ l = 1, \ldots, L_c$.

(5.c) Sort $\{\boldsymbol{X}_{G_{c,1}}, \ldots, \boldsymbol{X}_{G_{c,L_c}}\}$ according to $w_{c,l}$ in a decreasing order.

(5.d) Keep $\boldsymbol{X}_{c,G_l}$'s with $w_{c,l}$ greater than a pre-defined threshold $\nu_{c,n}$ or keep a pre-specified number (e.g., $\lfloor n/\log n \rfloor$) of covariates on the top of the list obtained from (5.c).

By the above procedure with some pre-determined threshold value $\nu_{c,n}$, the set of remaining variables is defined as $\widehat{M}^{\mathcal{C}}_{[\tau_L,\tau_U]} = \{G_{c,l} : 1 \leq l \leq L_c,\ w_{c,l} \geq \nu_{c,n}\}$.

We establish the sure screening property for the proposed conditional screening procedure. Denote the conditional test statistic $\widehat{T}_C$ for $H_{c,G_{c,l}}$ by $\widehat{T}_C^{(G_{c,l})}$. Let $T_C^{(G_{c,l})}$ be $\widehat{T}_C^{(G_{c,l})}$ with the coefficient estimate and variance estimate replaced by their population analogues. Define the set of conditionally relevant covariates as $M^{(\mathcal{C})}_{\tau_L,\tau_U} = \{G_{c,l} :\ 1 \leq l \leq L_c,$ there exists $\tau \in [\tau_L, \tau_U]$ such that $Q_Y(\tau \mid \boldsymbol{X})$ depends on $\boldsymbol{X}_{G_{c,l}}\}$. The results of the exponential tail probability bound for $|n^{-1}\widehat{T}_C^{(G_{c,l})} - n^{-1}T_C^{(G_{c,l})}|$ and the sure screening property are summarized in Theorem 3.3 and Corollary 3.3, respectively.

Theorem 3.3. *Given that the regularity conditions 3.1-3.4 in Section 3.7 hold. Assume that the number of variables in $\boldsymbol{X}_{(\mathcal{C})}$, noted as C, is finite. For any $c > 0$ and $1/4 < \zeta \leq 1/2$, there exists positive constant υ^* and η^* such that*

$$pr(\max_{1\leq l\leq q} |n^{-1}\widehat{T}_C^{(G_{c,l})} - n^{-1}T_C^{(G_{c,l})}| \geq cn^{\zeta-1/2}) \leq q\upsilon^* \exp[-\eta^* n^{4\zeta-1} - \log(n^{\zeta-1/2})]$$

for sufficiently large n.

Corollary 3.3 (Sure screening property for the conditional screening procedure)**.** *Given that the regularity conditions 3.1-3.4 and 3.7 in Section 3.7 hold. If we take the threshold value $\nu_{c,n} = \delta^{\mathcal{C}} n^{\zeta-1/2}$ with $\delta^{\mathcal{C}} \leq \alpha_0/2$, then there exists positive constant*

Table 3.1: Summary of simulation set-ups for evaluating the proposed testing procedure. $\Phi^{-1}(\cdot)$ is the inverse cumulative distribution for standard Normal distribution; and $F^{-1}_{Cauchy}(\cdot)$ is the inverse cumulative distribution for standard Cauchy distribution.

Univariate Case			
Set-up	Model	X	
U1	$Q_Y(\tau \mid X) = 10 + \Phi^{-1}(\tau)$	U(0,10)	
U2	$Q_Y(\tau \mid X) = 10 + 0.05 \cdot X + \Phi^{-1}(\tau)$	U(0,10)	
U3	$Q_Y(\tau \mid X) = 10 + q_u(\tau) \cdot X + \Phi^{-1}(\tau)$	U(0,10)	
U4	$Q_Y(\tau \mid X) = 10 + 0.5 \cdot \Phi^{-1}(\tau) \cdot X + \Phi^{-1}(\tau)$	U(0,10)	
U5	$Q_Y(\tau \mid X) = 10 + 0.5 \cdot X^{1/3} + F^{-1}_{Cauchy}(\tau)$	U(0,10)	
Multivariate Case			
Set-up	Model	X_1	X_2
M1	$Q_Y(\tau \mid X) = 10 + \Phi^{-1}(\tau)$	U(0,10)	10B(0.5)
M2	$Q_Y(\tau \mid X) = 10 + 0.1 \cdot X_1 + \Phi^{-1}(\tau)$	U(0,10)	10B(0.5)
M3	$Q_Y(\tau \mid X) = 10 + 0.05 \cdot X_1 + 0.05 \cdot X_2 + \Phi^{-1}(\tau)$	U(0,10)	10B(0.5)
M4	$Q_Y(\tau \mid X) = 10 + q_{m_1}(\tau) \cdot X_1 + q_{m_2}(\tau) \cdot X_2 + \Phi^{-1}(\tau)$	U(0,10)	10B(0.5)
M5	$Q_{Y\mid \boldsymbol{X}}(\tau) = 10 + 0.5 \cdot X_1^{1/3} I(X_1 > 3) + 0.5 \cdot X_2^{1/3} I(X_1 \leq 3) + F^{-1}_{Cauchy}(\tau)$	U(0,10)	10B(0.5)

a_3 and b_3, such that

$$pr(M^{(\mathcal{C})}_{[\tau_L,\tau_U]} \subseteq \widehat{M}^{(\mathcal{C})}_{[\tau_L,\tau_U]}) \geq 1 - S^{(\mathcal{C})}_{[\tau_L,\tau_U]} \cdot a_3 \exp[-b_3 n^{4\zeta-1} - \log(n^{\zeta-1/2})]$$

for sufficient large n, where $S^{(\mathcal{C})}_{[\tau_L,\tau_U]} = |M^{(\mathcal{C})}_{[\tau_L,\tau_U]}|$ is the cardinality of $M^{(\mathcal{C})}_{[\tau_L,\tau_U]}$. In particular, $pr(M^{(\mathcal{C})}_{[\tau_L,\tau_U]} \subseteq \widehat{M}^{(\mathcal{C})}_{[\tau_L,\tau_U]}) \to 1$ as $n \to \infty$.

The proofs for Theorem 3.3 and Corollary 3.3 are provided in Sections S2.2 and S2.3 of the Appendix.

3.4 Numerical Studies

3.4.1 Simulation Studies for Evaluating the Proposed Testing Procedure

We first evaluate the proposed testing procedure for $H_{0,G}$ in univariate settings where $\boldsymbol{X}_G$ contains one covariate X. The specific models for generating (X, Y) are presented in Table 3.1. In set-ups U1-U4, the working model (3.2) holds for $\tau \in [0.2, 0.8]$, which

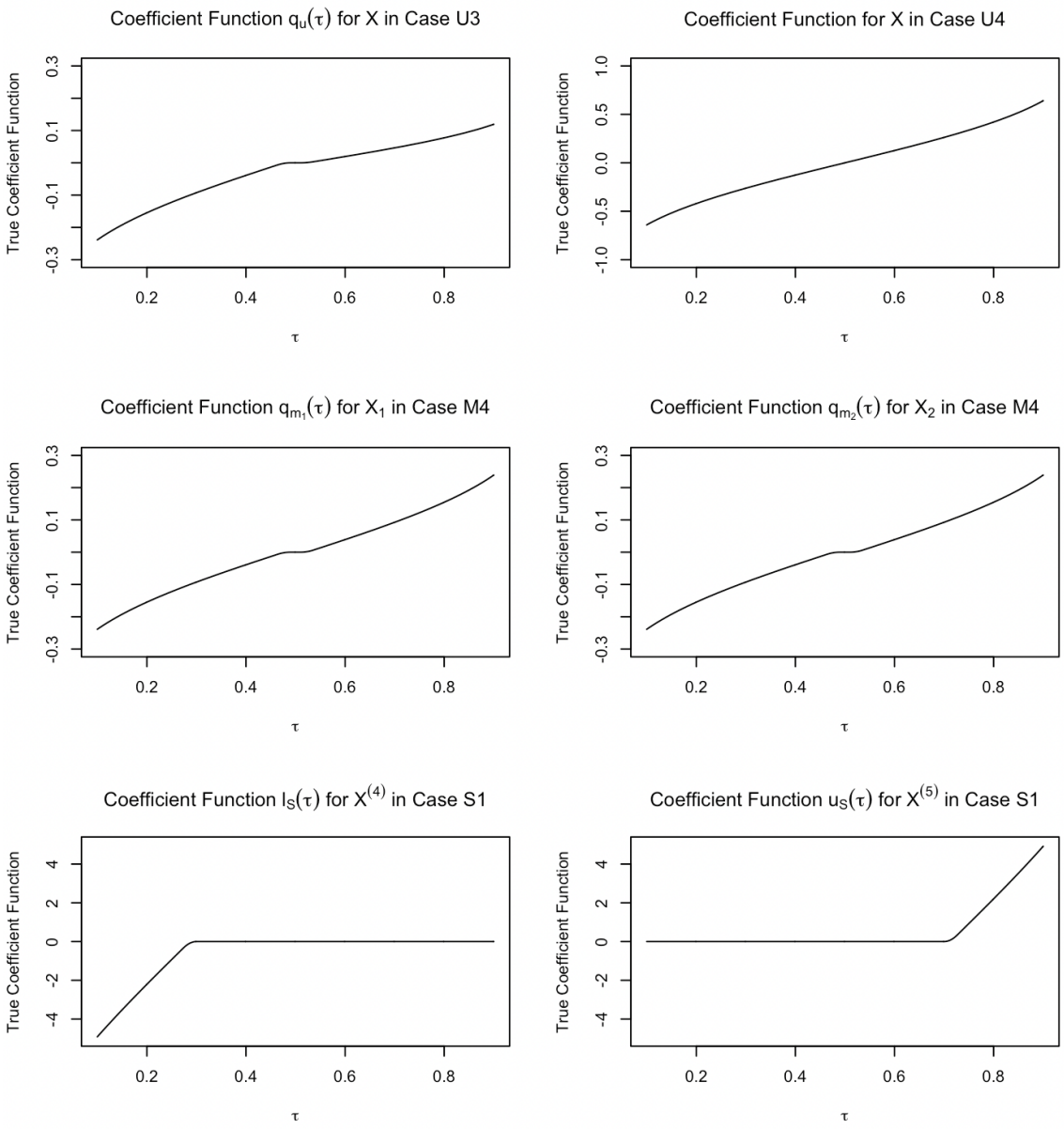


Figure 3.1: True coefficient functions for the simulation set-ups.

Table 3.2: Empirical rejection rates with 1000 replicates for $\boldsymbol{X}_G = X$. GIT: the proposed test based on $\widehat{T}_{UC}$; AQI: Zhu et al. (2018)'s method; Q_S and Q_W: rank score test (Gutenbrunner et al., 1993a) and Wald test (Koenker and Bassett Jr, 1982) based on quantile regression models; L_W: the Wald test based on linear regression.

		$\tau \in [0.2, 0.8]$		τ=0.4		τ=0.5		τ=0.6		
Set-up	n	GIT	AQI	Q_S	Q_W	Q_S	Q_W	Q_S	Q_W	L_W
U1	200	0.053	0.040	0.035	0.055	0.052	0.042	0.044	0.048	0.049
	400	0.050	0.057	0.045	0.048	0.040	0.045	0.039	0.033	0.045
U2	200	0.492	0.443	0.357	0.338	0.366	0.343	0.354	0.336	0.524
	400	0.788	0.755	0.637	0.623	0.629	0.638	0.636	0.622	0.824
U3	200	0.272	0.231	0.103	0.109	0.050	0.028	0.069	0.066	0.121
	400	0.578	0.515	0.171	0.156	0.042	0.029	0.083	0.079	0.207
U4	200	0.948	0.951	0.285	0.221	0.089	0.049	0.277	0.225	0.052
	400	1.000	1.000	0.480	0.407	0.064	0.051	0.495	0.389	0.065
U5	200	0.469	0.415	0.379	0.308	0.423	0.344	0.343	0.284	0.041
	400	0.742	0.683	0.668	0.596	0.703	0.687	0.635	0.593	0.038

Table 3.3: Empirical rejection rates with 1000 replicates for $\boldsymbol{X}_G = (X_1, X_2)^\intercal$. GIT: the proposed test based on $\widehat{T}_{UC}$, Q_S and Q_W: rank score test (Gutenbrunner et al., 1993a) and Wald test (Koenker and Bassett Jr, 1982) based on quantile regression models; ANOVA: the analysis of variance test for overall significance based on linear regression.

		$\tau \in [0.2, 0.8]$	τ=0.4		τ=0.5		τ=0.6		
Set-up	n	GIT	Q_S	Q_W	Q_S	Q_W	Q_S	Q_W	ANOVA
M1	200	0.058	0.050	0.055	0.053	0.048	0.050	0.047	0.044
	400	0.045	0.042	0.041	0.041	0.038	0.042	0.047	0.040
M2	200	0.933	0.805	0.791	0.827	0.799	0.823	0.801	0.952
	400	1.000	0.987	0.982	0.988	0.986	0.989	0.985	0.998
M3	200	0.900	0.809	0.797	0.831	0.812	0.811	0.802	0.959
	400	0.999	0.991	0.987	0.992	0.991	0.988	0.980	0.999
M4	200	0.409	0.169	0.131	0.052	0.036	0.116	0.111	0.051
	400	0.821	0.228	0.210	0.040	0.028	0.231	0.203	0.045
M5	200	0.415	0.452	0.325	0.384	0.345	0.268	0.264	0.025
	400	0.710	0.758	0.716	0.690	0.698	0.543	0.572	0.039

is the τ-interval of interest. Set-up U5 gives a scenario where the working linear quantile regression model does not hold. It is easy to see that U1 is a null case, where X has no effect on Y. In U2, a standard linear model holds and X has a constant effect on Y over $\tau \in [0.2, 0.8]$. U3 and U4 are two set-ups with dynamic effects varying across different τ's. The true coefficient functions in set-ups U3 and U4 are presented in Figure 3.1. In set-up U5, X takes a non-linear functional form to influence Y and thus the working model is not satisfied.

We compare the following testing procedures:

GIT: the proposed test based on $\widehat{T}_{UC}$ with $[\tau_L, \tau_U] = [0.2, 0.8]$;

AQI: the test proposed in Zhu et al. (2018), with the quantile interval set as $[0.2, 0.8]$ for Y and $[0, 1]$ for X;

Q_S: rank score test (Gutenbrunner et al., 1993a)

Q_W: Wald test (Koenker and Bassett Jr, 1982)

L_W: Wald test based on linear regression.

In each setting, the significance level is set as 0.05. We consider sample sizes, 200 and 400.

Table 3.2 presents the empirical rejection rates based on 1000 simulations. In the null case U1, all methods yield empirical sizes close to the nominal level of 0.05. In set-ups U3–U5, where dynamic covariate effects are present, we observe that the proposed method and Zhu et al. (2018)'s method, which are designed to capture global effects throughout $\tau \in [0.2, 0.8]$, yield much higher power than tests which target the local effect on a single τ or the mean when dynamic effects are present, for example set-ups U3 and U4. These demonstrate substantial power gains resulted from integrating information across quantiles in the presence of dynamic covariate

effects. In addition, we observe that the proposed method and Zhu et al. (2018)'s method have comparable performance in the univariate settings.

We also evaluate the proposed testing procedure in multivariate settings, where $\boldsymbol{X}_G$ includes two covariates X_1 and X_2. To illustrate the utility of our method for handling both continuous and discrete variables, we generate X_1 as a continuous variable and X_2 as a discrete variable. We consider five settings M1–M5 with configuration details shown in Table 3.1. M1 is the null case, where both X_1 and X_2 have no effects on Y. M2 corresponds to the case where only X_1 influences Y and its effect is constant. In M3, both X_1 and X_2 have constant covariate effects on Y. In M4, X_1 and X_2 have partial effects on Y. The true coefficient functions, $q_{m_1}(\tau)$ and $q_{m_2}(\tau)$, are shown in Figure 3.1. M5 is a set-up where X_1 and X_2 influence Y in a non-standard way and the working model (3.2) does not hold. In all multivariate settings, Zhu et al. (2018) is no longer applicable. We compare the proposed GIT to Q_S and Q_W with $\tau = 0.4$, 0.5, or 0.6 and the analysis of variance test for overall significance based on linear regression (ANOVA).

Table 3.3 reports the empirical rejection rates of these tests based on 1000 simulations. All methods have empirical sizes close to the nominal level 0.05 in the null case M1. The empirical power of all tests grows as the sample size increases. When there are varying covariate effects, such as in set-ups M4–M5, the proposed method can yield much higher power than tests, Q_S, Q_W, and ANOVA, which target local covariate effects on a single τ or the mean. These results suggest that good utility of the proposed tests to detect the existence of either constant or dynamic covariate effects, no matter the covariates are continuous or discrete.

Table 3.4: Simulation results for unconditional procedures based on 500 replicates. MMMS: median minimum model size; RSD: the robust standard deviation; pr($X^{(j)}$): probability of selecting $X^{(j)}$; pr($\mathcal{A}$): probability of selecting all covariates in $\mathcal{A}$; GIT: the proposed one-step unconditional procedure; GOT: the proposed two-step unconditional procedure with $S_G = 2$; SIS: Fan and Lv (2008)'s method; QaSIS: He et al. (2013)'s method; AQI: Zhu et al. (2018)'s method.

		MMS		pr($X^{(j)}$)					
Set-up	Method	MMMS	RSD	$X^{(1)}$	$X^{(2)}$	$X^{(3)}$	$X^{(4)}$	$X^{(5)}$	pr($\mathcal{A}$)
S1	GIT	6	2	1.000	1.000	1.000	1.000	0.996	0.996
	SIS	5	0	1.000	1.000	1.000	1.000	0.998	0.998
	QaSIS(0.25)	8	10	0.976	0.990	0.988	0.982	0.910	0.884
	QaSIS(0.5)	6	3	0.990	0.998	0.998	0.992	0.972	0.958
	QaSIS(0.75)	8	9	0.994	0.992	0.988	0.968	0.912	0.896
	AQI	5	0	1.000	1.000	1.000	1.000	1.000	1.000
	GOT($S_G = 2$)	7	2	0.998	0.996	1.000	0.998	0.992	0.990
S2	GIT	5	1	1.000	1.000	1.000	1.000	0.994	0.994
	SIS	50	264	0.694	0.694	0.688	0.628	0.554	0.466
	QaSIS(0.25)	97	130	0.482	0.544	0.518	0.392	0.268	0.142
	QaSIS(0.5)	20	20	0.978	0.988	0.972	0.956	0.854	0.822
	QaSIS(0.75)	98	128	0.448	0.536	0.490	0.410	0.286	0.178
	AQI	5	1	1.000	1.000	1.000	0.998	0.998	0.996
	GOT($S_G = 2$)	6	1	0.998	0.996	1.000	1.000	0.992	0.990
S3	GIT	12	7	1.000	1.000	1.000	1.000	1.000	1.000
	SIS	214	798	0.358	0.402	0.446	1.000	1.000	0.300
	QaSIS(0.25)	25	10	0.992	1.000	0.982	1.000	1.000	0.976
	QaSIS(0.5)	27	9	0.996	1.000	0.998	1.000	1.000	0.994
	QaSIS(0.75)	90	77	0.406	0.562	0.640	1.000	1.000	0.340
	AQI	20	6	1.000	1.000	1.000	1.000	1.000	1.000
	GOT($S_G = 2$)	13	7	1.000	1.000	1.000	1.000	1.000	1.000
S4	GIT	5	1	1.000	1.000	1.000	1.000	1.000	1.000
	SIS	2194	3860	0.260	0.266	0.256	0.214	0.196	0.140
	QaSIS(0.25)	33	37	0.912	0.950	0.992	0.954	0.888	0.794
	QaSIS(0.5)	6	2	1.000	1.000	1.000	1.000	1.000	1.000
	QaSIS(0.75)	10	9	1.000	1.000	1.000	1.000	1.000	1.000
	AQI	5	0	1.000	1.000	1.000	1.000	1.000	1.000
	GOT($S_G = 2$)	6	1	1.000	1.000	1.000	1.000	1.000	1.000
S4*	GIT	5	1	1.000	1.000	1.000	1.000	1.000	1.000
	SIS	2180	4139	0.332	0.264	0.376	0.306	0.346	0.186
	QaSIS(0.25)	5	0	1.000	1.000	1.000	1.000	1.000	1.000
	QaSIS(0.5)	5	0	1.000	1.000	1.000	1.000	1.000	1.000
	QaSIS(0.75)	8	4	1.000	0.912	1.000	1.000	1.000	0.912
	AQI	–	–	–	–	–	–	–	–
	GOT($S_G = 2$)	6	1	1.000	1.000	1.000	1.000	1.000	1.000

3.4.2 Simulation Studies for Evaluating the Proposed Screening Procedures

We conduct simulation studies to evaluate the performance of the proposed one-step screening procedure in (3.a)–(3.d), denoted by GIT, and the proposed two-step procedure in (4.a)–(4.c) with $S_G = 2$, denoted by GOT. For comparisons, we consider existing approaches, including Fan and Lv (2008)'s method, denoted by SIS, He et al. (2013)'s method at $\tau = 0.25, 0.5$ or 0.75, denoted by QsSIS(τ), as well as Zhu et al. (2018)'s method with quantile interval sets, $[0.2, 0.8]$ for Y and $[0, 1]$ for X, denoted by AQI. When implementing He et al. (2013)'s method, we set the number of basis as 3. To assess the performance of these screening methods, we use the median minimum model size of the selected models required for sure screening, and the robust standard deviation, defined as the interquartile range of minimum model size, and the probability of selecting each $X^{(j)}$, and the probability of selecting all covariates in $\mathcal{A}$ when top $[n/\log(n)]$ covariates are maintained.

The simulation set-ups are described as follows:

S1 ($n = 200$, $p = 2000$): $Y = 0.2(X^{(1)} + 0.8X^{(2)} + 0.6X^{(3)} + 0.4X^{(4)} + 0.2X^{(5)}) + \varepsilon$, where ε follow the standard normal distribution.

S2 ($n = 200$, $p = 2000$): $Y = 0.2(X^{(1)} + 0.8X^{(2)} + 0.6X^{(3)} + 0.4X^{(4)} + 0.2X^{(5)}) + \exp(Z) \cdot \varepsilon$, where Z and ε follow the standard normal distribution.

S3 ($n = 400$, $p = 5000$): $Y = X^{(1)}I(X^{(1)} > 0) + X^{(2)}I(X^{(1)} \leq 0) + \exp(X^{(19)} + X^{(20)}) + \exp(X^{(3)}) \cdot \varepsilon$, where ε follows the standard normal distribution.

S4 ($n = 400$, $p = 5000$): $Q_Y(\tau|\boldsymbol{X}) = 3X^{(1)}I(X^{(1)} > 0) + 3X^{(3)}I(X^{(1)} \leq 0) + l_S(\tau) \cdot X^{(4)} + u_S(\tau) \cdot X^{(5)} + (s(X^{(2)}) + 1)^2 \cdot Q_\varepsilon(\tau)$, where $s(a) = (a - E(a))/sd(a)$, $l_S(\tau)$ and $u_S(\tau)$ are plotted in Figure 3.1, and ε follows standard Cauchy distribution.

In the above set-ups, the covariates $\boldsymbol{X} = \{X^{(1)}, \ldots, X^{(p)}\}^T$ are generated from multivariate normal distribution with mean zero and covariance matrix $\boldsymbol{\Sigma} = (0.9^{|k-k'|})_{p\times p}$. The error terms Z and ε are independent of $\boldsymbol{X}$. It is easy to see that the relevant covariate set is $\mathcal{A} = \{X^{(1)}, X^{(2)}, X^{(3)}, X^{(4)}, X^{(5)}\}$ for set-ups S1, S2 and S4. In set-up S3, the relevant/active covariates are more separated from each other with $\mathcal{A} = \{X^{(1)}, X^{(2)}, X^{(3)}, X^{(19)}, X^{(20)}\}$. We further consider an additional set-up S4*, which is the same as S4 except that we transform half of the covariates to discrete covariates. Specifically, in S4*, we first use the same way to generate $\{X^{(1)}, X^{(2)}, \ldots, X^{(p)}\}^T$, and then dichotomize $\{X^{(2)}, X^{(4)}, \ldots, X^{(2\lfloor p/2 \rfloor)}\}^T$ at 0 to generate binary covariates defined as $I(X^{(j)} < 0)$ $(j = 2, 4, \ldots, 2\lfloor p/2 \rfloor)$. In this case, Zhu et al. (2018)'s method can not be applied. To implement He et al. (2013)'s method, we use the linear option due to singular issues.

In Table 3.4, we summarize the screening results based on 500 simulations. In set-up S1, where the error term follows the normal distribution and the relevant covariates are highly correlated with each other, we observe that all the methods perform quite well. In set-up S2, which differs from S1 only by the error distribution, we notice that there is substantial deterioration with the performance of SIS. The number of covariates needed for sure screening along with its variability inflates substantially from MMMS(RSD)= 5(0) to 50(264), and the probability of retaining all relevant covariates drops significantly from 1.00 to 0.47. In the other three set-ups, S3, S4 and S4*, we have similar observations regarding the under-performance of Fan and Lv (2008). Such observations are not surprising and are likely caused by the fact that the normal error assumption is no longer valid in these settings. Also, we notice that He et al. (2013) has varying performance for different τ's. For example, in set-up S2, He et al. (2013)'s method with $\tau = 0.5$ may select the relevant covariates over 80% of times; while by He et al. (2013)'s method with $\tau = 0.25$ or 0.75 the probability of keeping all relevant variables reduces to be below 20%. Compared to He et al. (2013)'s

method, which focuses on local effects, the screening procedures that examine global effects, such as Zhu et al. (2018)'s method and the proposed methods, GIT and GOT, demonstrate better performance, as reflected by larger selection probabilities, pr($\mathcal{A}$), and smaller model sizes measured by MMMS and RSD. A reasonable interpretation is that the global testing procedures leverage information across different τ's, thereby producing higher detection power.

In set-ups S1, S2, and S4, where relevant covariates are strongly correlated, the proposed methods, GIT and GOT, and Zhu et al. (2018)'s method, AQI, have similar performance. In set-up S3, the relevant covariates are separated into two clusters with one cluster including $X^{(1)}$, $X^{(2)}$, and $X^{(3)}$ and the other including $X^{(19)}$ and $X^{(20)}$. In addition, $X^{(19)}$ and $X^{(20)}$ have stronger covariate effects than $X^{(1)}$, $X^{(2)}$, and $X^{(3)}$. In this case, though the proposed methods and Zhu et al. (2018)'s method all have high probabilities of selecting relevant covariates, Zhu et al. (2018)'s method yields relatively larger model size as compared to the proposed methods. This is caused by the tendency of Zhu et al. (2018)'s method to select "neighboring" covariates around $X^{(19)}$ and $X^{(20)}$, such as $X^{(18)}$ or $X^{(21)}$. Since these covariates are highly corrected with $X^{(19)}$ and $X^{(20)}$ (which have strong effects on the outcome), Zhu et al. (2018)'s method may catch the trails of these neighboring covariates by producing interval quantile independence indices comparable to or even higher than those for $X^{(1)}, X^{(2)}$, and $X^{(3)}$. Despite this discrepancy, we think that the proposed methods and Zhu et al. (2018)'s method have quite comparable performance in variable screening when all covariates are continuous, while the proposed methods offer flexibility to naturally accommodate discrete covariates.

Table 3.5: The Median Minimum Model Size (MMMS), the sampling variability of minimum model size (RSD) and the selection probability across 500 simulations. MMMS: median minimum model size; RSD: the robust standard deviation; pr($\mathcal{A}$): probability of selecting all covariates in $\mathcal{A}$; IA: irrelevant variables; SIS: Fan and Lv (2008)'s method; GIT: the proposed one-step unconditional procedure; CSIS: Barut et al. (2016)'s method; CGIT, the proposed conditional screening procedure.

			MMS		
Set-up	$\mathcal{A}$	Method	MMMS	RSD	pr($\mathcal{A}$)
CS1(i)	$X^{(6)}$	SIS	1995	0	0.000
		GIT	1995	0	0.000
		CSIS($X^{(1)}$-$X^{(5)}$)	1	0	1.000
		CSIS($X^{(1)}$-$X^{(5)}$, 5 IAs)	1	0	1.000
		CSIS(5 IAs)	1	0	1.000
		CGIT($X^{(1)}$-$X^{(5)}$)	1	0	1.000
		CGIT($X^{(1)}$-$X^{(5)}$, 5 IAs)	1	0	1.000
		CGIT(5 IAs)	1	0	1.000
CS1(ii)	$X^{(6)}$	SIS	1991	240	0.006
		GIT	1995	0	0.000
		CSIS($X^{(1)}$-$X^{(5)}$)	1	44	0.744
		CSIS($X^{(1)}$-$X^{(5)}$, 5 IAs)	1	58	0.730
		CSIS(5 IAs)	8	303	0.608
		CGIT($X^{(1)}$-$X^{(5)}$)	1	0	1.000
		CGIT($X^{(1)}$-$X^{(5)}$, 5 IAs)	1	0	1.000
		CGIT(5 IAs)	1	0	0.998
CS2(i)	$X^{(2000)}$	SIS	1999	0	0.000
		GIT	1999	0	0.000
		CSIS($X^{(1)}$-$X^{(5)}$)	1	0	1.000
		CSIS($X^{(1)}$-$X^{(5)}$, 5 IAs)	1	0	1.000
		CSIS(5 IAs)	1490	768	0.004
		CGIT($X^{(1)}$-$X^{(5)}$)	1	0	1.000
		CGIT($X^{(1)}$-$X^{(5)}$, 5 IAs)	1	0	1.000
		CGIT(5 IAs)	1044	1137	0.048
CS2(ii)	$X^{(2000)}$	SIS	1999	124	0.012
		GIT	1999	0	0.000
		CSIS($X^{(1)}$-$X^{(5)}$)	751	1208	0.110
		CSIS($X^{(1)}$-$X^{(5)}$, 5 IAs)	898	1090	0.050
		CSIS(5 IAs)	1320	939	0.000
		CGIT($X^{(1)}$-$X^{(5)}$)	1	0	1.000
		CGIT($X^{(1)}$-$X^{(5)}$, 5 IAs)	1	0	1.000
		CGIT(5 IAs)	1046	1075	0.044
CS3	$X^{(2000)}$	SIS	1675	1184	0.058
		GIT	1824	698	0.004
		CSIS($X^{(1)}$-$X^{(5)}$)	627	1174	0.130
		CSIS($X^{(1)}$-$X^{(5)}$, 5 IAs)	789	1048	0.046
		CSIS(5 IAs)	857	1055	0.042
		CGIT($X^{(1)}$-$X^{(5)}$)	1	1	0.988
		CGIT($X^{(1)}$-$X^{(5)}$, 5 IAs)	1	0	0.998
		CGIT(5 IAs)	21	62	0.626

3.4.3 Simulation Studies for Evaluating the Proposed Conditional Screening Procedure

We conduct simulation studies to investigate the performance of the proposed conditional screening procedure outlined in (5.a)–(5.d) in the main manuscript. In the simulations, we compare it with the proposed unconditional screening procedures, Fan and Lv (2008), as well as the conditional screening procedure proposed by Barut et al. (2016). We consider scenarios, where $\boldsymbol{X}_{\mathcal{C}}$ include relevant variables only, some relevant variables along with some randomly selected irrelevant variables, or irrelevant variables only. The same criteria described in Section 3.4.2 are used to evaluate the performance for these methods.

We consider the following simulation set-ups :

CS1 $(n = 200,\, p = 2000)$: $Y = 3X^{(1)} + 3X^{(2)} + 3X^{(3)} + 3X^{(4)} + 3X^{(5)} - 7.5X^{(6)} + \varepsilon$, where each covariate follows the standard normal distribution with equal correlation 0.5, and the error term ε follows (i) standard normal or (ii) standard Cauchy distribution.

CS2 $(n = 200,\, p = 2000)$: $Y = 10X^{(1)} + X^{(2000)} + \varepsilon$, where each covariate follows the standard normal distribution with equal correlation 0.9 except for $X^{(2000)}$, which is independent of all the other covariates, and the error term ε follows (i) standard normal or (ii) standard Cauchy distribution.

CS3 $(n = 200,\, p = 2000)$: $Y = 10X^{(1)} + (3X^{(2000)} + 4)^2 \cdot \varepsilon$, where each covariate follows the standard normal distribution with equal correlation 0.9 except for $X^{(2000)}$, which is independent of all the other covariates, and the error term ε follows the standard normal distribution.

CS1(i) and CS2(i) are the same set-ups considered by Barut et al. (2016). In CS1(i), $X^{(6)}$ is a hidden variable based on marginal screening procedures. This is due to

the fact that in this case, the covariance between $X^{(j)}$ and Y is 0 given $\beta^{*(j)} = -\sum_{k \neq j} \beta^{*(k)} \Sigma_{kj} / \Sigma_{jj}$, where Σ_{kj} denotes the (k, j)th element of $\boldsymbol{\Sigma} = \text{var}(\boldsymbol{X})$. In CS2(i), $X^{(2000)}$ has the least priority to be included by marginal screening procedures. This is because in this case, $\text{cov}(X^{(1)}, Y) = 10$, $\text{cov}(X^{(2000)}, Y) = 1$, and $\text{cov}(X^{(j)}, Y) = 9$ for $j = 2. \ldots, 1999$. This means, the marginal covariance between each inactive variable and Y is larger than the covariance between the active variable $X^{(2000)}$ and Y. Set-up CS1(ii) and CS2(ii) differ from CS1(i) and CS2(i) only by the error distribution. In CS3, which is a set-up adapted from CS2(i), $X^{(2000)}$ has a non-linear varying effect on the quantiles of Y.

Table 3.5 presents the results based on 500 simulations. Within the parenthese following the notation "CSIS" or "CGIT", we indicate the conditioning covariates $\boldsymbol{X}_{\mathcal{C}}$ specified in these conditional screening procedures. We see that the unconditional screening procedures, SIS and GIT, barely select hidden active variables, for example, $X^{(6)}$ in set-up CS1(i) and $X^{(2000)}$ in set-up CS2(i). The conditional procedures, CSIS and CGIT, work well in set-up CS1(i) regardless whether or not the conditioning covariates are active or not. In set-up CS2(i), the conditional screening procedures yield lower probabilities of selecting relevant covariates but still perform well when some of conditioning variables are active. Based on the results for CS1(i) and CS2(i), we may conclude that the proposed method and Barut et al. (2016) have comparable performance under standard linear models with normal random errors.

However, when the error term is not normally distributed or there exists a dynamic covariate effect, such as in settings CS1(ii), CS2(ii) and CS3, we notice that the performance of Barut et al. (2016) deteriorates significantly. For example, in set-up CS(ii), the selection probability of $X^{(6)}$ by Barut et al. (2016) decreases to 0.74, 0.73 and 0.61 regardless the choices of the conditioning variables. The deterioration of Barut et al. (2016)'s performance is much more evident in set-ups CS2(ii) and CS3, as reflected by low probabilities of selecting the conditionally relevant covariate, which

are around or below 0.10, and are much smaller than those produced by the proposed conditional screening procedure CGIT. These results strongly support the advantage of taking a global view for assessing covariate effects, particularly in the presence of dynamic covariate effects.

3.5 Real Example with Microarray Data

We apply the proposed methods to a microarray dataset (Scheetz et al., 2006), which contains the gene expression levels of 31,098 probe sets on 120 12-week old male offsprings of rats. With this dataset, one interest is to identify the set of genes related to gene *TRIM32*, which is a known predictor for genetically heterogeneous diseases including Muscular Dystrophy, Limb-Girdle, Autosomal Recessive 8 and Bardet-Biedl Syndrome 11. The probe id corresponding to gene *TRIM32* is "1389163_at".

We first illustrate the utility of the proposed global testing procedure through evaluating the marginal relevance of six example genes to the expression level of *TRIM32*. To test the effect of each of these genes, we apply the proposed global test, Wald tests for linear quantile regression with $\tau = 0.25, 0.5$ and 0.75, Wald tests for linear regression with outliers and after the removal of outliers based on Cook's Distance, and Wald test based on the robust linear regression (Hampel et al., 1986). Table 3.6 presents the p values obtained from these different tests. For the genes with probe id "1367462_at" and "1372996_at", their effects are captured by linear regression after removing outliers, robust linear regression, quantile regression based tests with most choices of τ, as well as the proposed test. The test based on standard linear regression does not detect the effect of either of these two genes, likely due to the "diluting" influence from the outliers. We have opposite findings regarding the effects of the genes with probe id "1367479_at" and "1367525_at". As hinted by the scatter plots in Figure 3.2, these discrepant results are likely caused by the presence

Table 3.6: Summary for the p values of six example genes. GIT: the proposed test based on $\widehat{T}_{UC}$; Q_S and Q_W: rank score test (Gutenbrunner et al., 1993a) and Wald test (Koenker and Bassett Jr, 1982) for quantile regression models; L_W and $L_W(\text{rm})$: Wald tests for linear regression with outliers and after the removal of outliers based on Cook's Distance; RL_W: the Wald test based on the robust linear regression.

Probe id	GIT	$Q_W(0.25)$	$Q_W(0.5)$	$Q_W(0.75)$	L_W	$L_W(\text{rm})$	RL_W
1367462_at	0.0004	0.0530	0.0018	0.0007	0.4673	0.0005	0.0005
1372996_at	0.0024	0.0012	0.0257	0.0111	0.9714	0.0049	0.0078
1367479_at	0.7616	0.2396	0.5638	0.6088	0.0176	0.1663	0.4939
1367525_at	0.9640	0.5331	0.9326	0.8139	0.0082	0.9819	0.7016
1379467_at	0.0040	0.1097	0.3106	0.0293	0.4370	0.5645	0.5439
1381314_at	0.0184	0.0951	0.0352	0.7773	0.1069	0.4604	0.1371

of a few outliers, which are not appropriately handled by standard linear regression and thus leads to spurious effect estimates. These demonstrate the robustness of the proposed testing procedure against outliers.

As suggested by exploratory marginal linear quantile regression analyses (see the third column of Figure 3.2), constant location-shift effects may not be adequate for the genes with probe id, "1379467_at" and "1381314_at", but are presumed by linear regression based tests. In this case, the local quantile regression based tests separately examine the effects of these genes at different quantile levels; thus it is not surprising that the resulting p values suggest significant effects at some τ's but not at the other τ's. All linear regression based tests fail to capture the effects of these two genes. This may reflect effect attenuation resulted from assuming a varying effect as constant. The proposed test, by taking a global perspective for assessing effects, sensibly support the relevance of these two genes to the outcome.

We apply the proposed screening procedures to help identify outcome-relevant genes out of 31,097 genes. In our analyses, we first perform the proposed unconditional screening procedures to filter out most irrelevant genes. Specifically, we keep the genes ranked top $[2n/\log n] = 50$. With the remaining genes, we perform Zheng et al. (2015)'s globally adaptive quantile regression method with $\tau \in [0.2, 0.8]$ for further variable selection. We also analyze the same data by alternative combinations

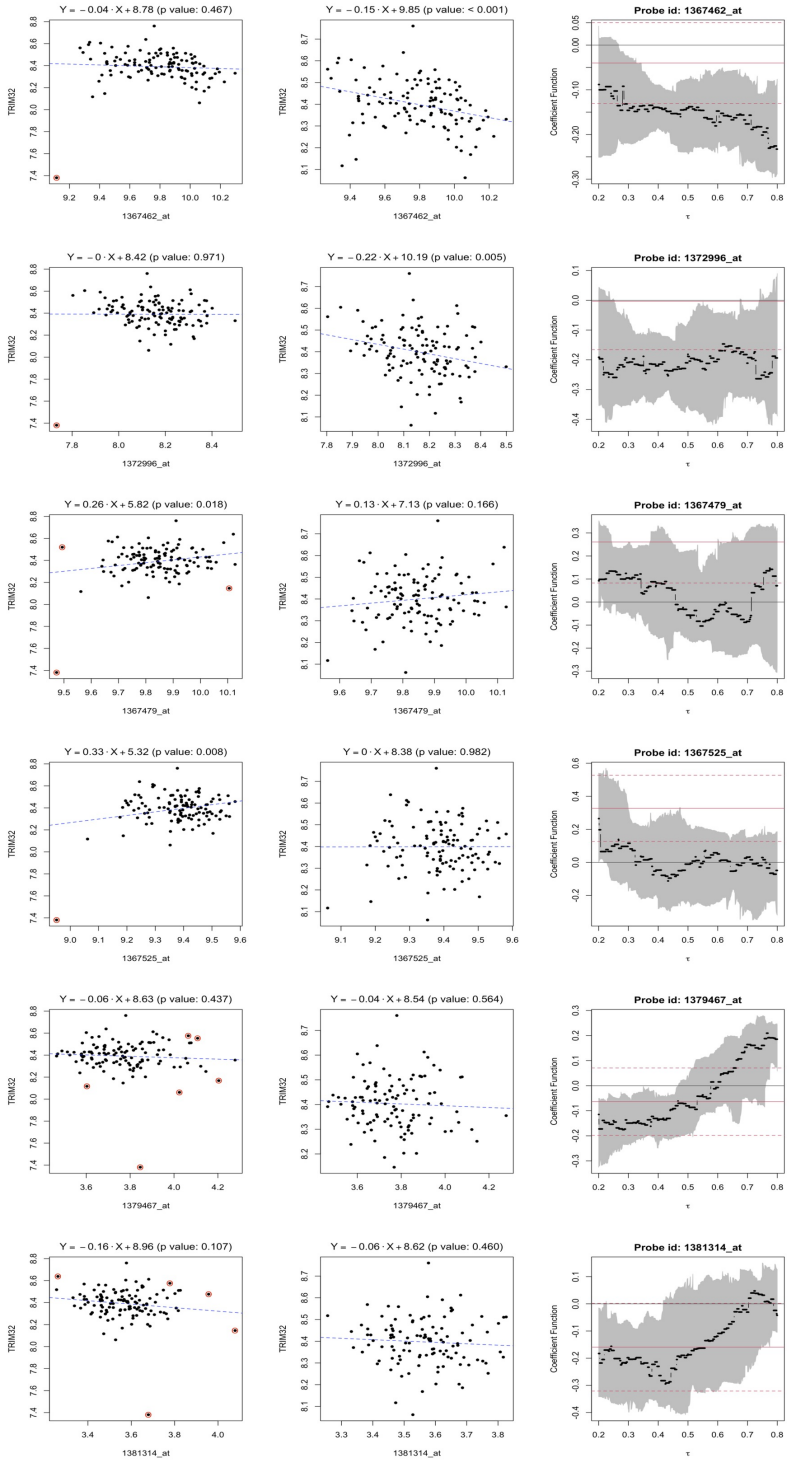


Figure 3.2: Scatter plots and the coefficient function for six example probes: the second column show show the scatter plots with the outliers removed based on the Cook's distance and the regression lines from linear regression (blue dashed lines); the first column show the scatter plots with outliers circled red and the regression lines from linear regression (blue dashed lines). The third column show the linear quantile regression's fitted coefficients (black lines) with the region within the 95% confidence intervals (shaded regions), along with the linear regression's fitted coefficients (red solid lines) with the 95% confidence interval (red dashed lines).

of screening and variable selection approaches, including Fan and Lv (2008) coupled with adaptive Lasso for linear regression (Zou, 2006), He et al. (2013) coupled with locally concerned quantile regression with adaptive Lasso penalty (Belloni and Chernozhukov, 2011) for $\tau = 0.25$, 0.5, or 0.75, and Zhu et al. (2018) coupled with Zheng et al. (2015)'s globally adaptive quantile regression method with $\tau \in [0.2, 0.8]$. When applying each approach, we determine the tuning parameter in the variable selection step by cross validation.

The heatmap presented in Figure 3.3 informs the sets of genes selected by different approaches and also displays the Pearson's correlation in expression level between the genes selected by the proposed one-step approach and the genes selected by the other approaches. With the same variable selection procedure, using the proposed global tests for variable screening leads to more parsimonious selection of genes as compared to adopting Zhu et al. (2018) which also takes a global view for variable screening. We observe that the gene with probe id "1393510_at" selected from using the proposed methods is also selected from using Fan and Lv (2008), He et al. (2013) and Zhu et al. (2018).All genes selected based on the proposed one-step approach have moderate or high correlations with at least one gene selected by the other approaches. This observation may help endorse the sensible gene selection by the proposed approach based on the results from several benchmark approaches.

For each approach, we further assess the quantile prediction performance. To compare across different approaches, we adjust the tuning parameter in the variable selection step so that all approaches select the same number of genes. For a given number of selected genes, denoted by g, following the approach developed by Li and Peng (2017), we measure the quantile prediction error as

$$\widehat{PE}^{(g)} = n^{-1} \sum_{i=1}^{n} \int_{\tau_L}^{\tau_U} \rho_\tau [Y_i - \boldsymbol{X}_{S,i}^{\mathrm{T}} \widehat{\boldsymbol{\theta}}_S(\tau)] d\tau,$$

where $\rho_\tau(u) = u\{\tau - I(u < 0)\}$, $\boldsymbol{X}_S$ represents the express levels of the selected genes, and $\widehat{\boldsymbol{\theta}}_S(\tau)$ represents the estimated regression quantiles derived from the final model fitting at the variable selection step. Under a linear regression model or a local quantile regression model, the $\widehat{\boldsymbol{\theta}}_S(\tau)$ is extrapolated as a constant function over τ equal to the regression coefficient estimate.

In Table 3.7, we report $\widehat{PE}^{(g)}$ with $g = 1, \ldots, 10$ resulted from all the approaches considered. We see that the estimated prediction errors associated with the proposed methodsare always comparable or smaller than those associated with the other methods. For example, with $g = 5$, the estimated prediction errors associated with the proposed methods are both around 0.11 and smaller than the other approaches.

We also apply conditional screening procedures with the conditioning covariates representing the two important genes suggested in Scheetz et al. (2006), *Abca4* and *Opn1sw* with probe ID "1384603_at" and "1388025_at", respectively. We pair the proposed method and Barut et al. (2016) respectively with globally adaptive quantile regression method (Zheng et al., 2015) and linear regression with adaptive LASSO for variable selection. The heatmap presented in Figure 3.4 indicates that CGIT yields much more sparse gene selection results as compared to CSIS, and the expression level of genes selected by CGIT are well correlated with those of the genes selected CSIS. The results in Table 3.7 show that the estimated prediction error is 0.19 based on linear regression with only *Abca4* and *Opn1sw* as covariates. The prediction errors decrease when the conditionally relevant covariate set includes additional covariates identified from variable screening and variable selection. The prediction errors associated with CGIT are smaller than those associated with CSIS most times. The prediction error reduction from using CGIT instead of CSIS, is more apparent when there are fewer selected genes. This may indirectly imply that CGIT-based approach, as compared to CSIS-based approach, may give higher priority to genes with more predictive power and thus leads to large gains in prediction when the "model size" is

Table 3.7: The estimated prediction error $\widehat{PE}^{(g)}(\Delta)$ for different model size. GIT: the proposed one-step unconditional procedure; GOT: the proposed two-step unconditional procedure with $S_G = 2$; AQI: Zhu et al. (2018)'s method; SIS: Fan and Lv (2008)'s method; QaSIS: He et al. (2013)'s method; CGIT: the proposed conditional screening procedure; CSIS: Barut et al. (2016)'s method.

Method	Size:1	Size:2	Size:3	Size:4	Size:5	Size:6	Size:7	Size:8	Size:9	Size:10
A. Unconditional screening										
GIT	0.178	0.171	0.118	0.117	0.116	0.115	0.114	0.113	0.109	0.108
GOT	0.177	0.172	0.169	0.127	0.113	0.112	0.111	0.111	0.111	0.111
AQI	0.171	0.157	0.142	0.141	0.140	0.134	0.133	0.124	0.118	0.108
SIS	0.185	0.172	0.159	0.144	0.137	0.123	0.113	0.109	0.109	0.108
QaSIS ($\tau = 0.25$)	0.195	0.179	0.179	0.167	0.166	0.162	0.156	0.155	0.144	0.143
QaSIS ($\tau = 0.5$)	0.150	0.148	0.138	0.129	0.122	0.120	0.119	0.119	0.117	0.116
QaSIS ($\tau = 0.75$)	0.180	0.173	0.158	0.154	0.155	0.135	0.131	0.130	0.125	0.124
B. Conditional screening (*Abca4*, *Opn1sw*)										
CGIT	–	–	0.161	0.163	0.163	0.156	0.113	0.104	0.103	0.103
CSIS	–	–	0.178	0.169	0.142	0.130	0.120	0.117	0.112	0.105

small.

3.6 Remarks

In this work, we develop a new testing and screening framework that can help determine outcome-relevant covariates in classic univariate and multivariate settings and ultra-high dimensional settings. The proposed methods sensibly adopt a global perspective that examines covariate effects over a continuum of outcome quantiles. Such a global perspective shares a similar spirit with the concept of globally concerned quantile regression proposed by Zheng et al. (2015). Nevertheless, Zheng et al. (2015)'s work is hinged upon the assumption of a global linear quantile regression model, while our testing procedures tackle a non-model-based null hypothesis and the corresponding screening procedure is model-free. Our numerical studies strongly support the advantages of the proposed methods over existing locally concerned methods, particularly in data settings with dynamic covariate effects.

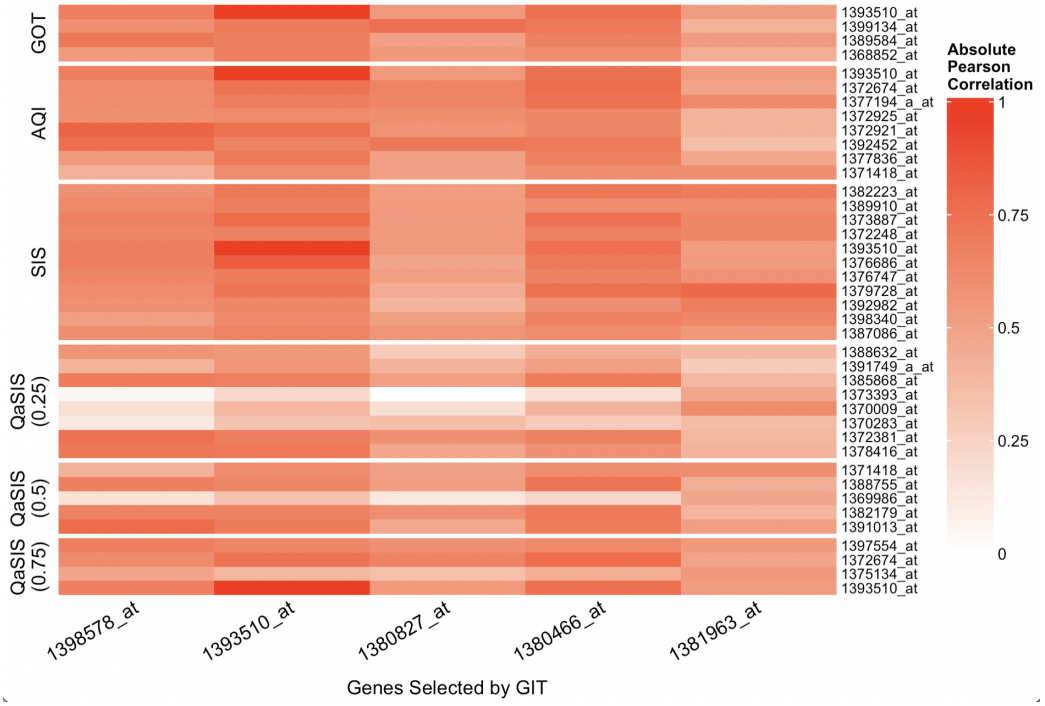


Figure 3.3: Heat map for the correlation between the proposed unconditional screening procedure and the other methods. GIT: the proposed one-step unconditional procedure; GOT: the proposed two-step unconditional procedure with $S_G = 2$; AQI: Zhu et al. (2018)'s method; SIS: Fan and Lv (2008)'s method; QaSIS: He et al. (2013)'s method.

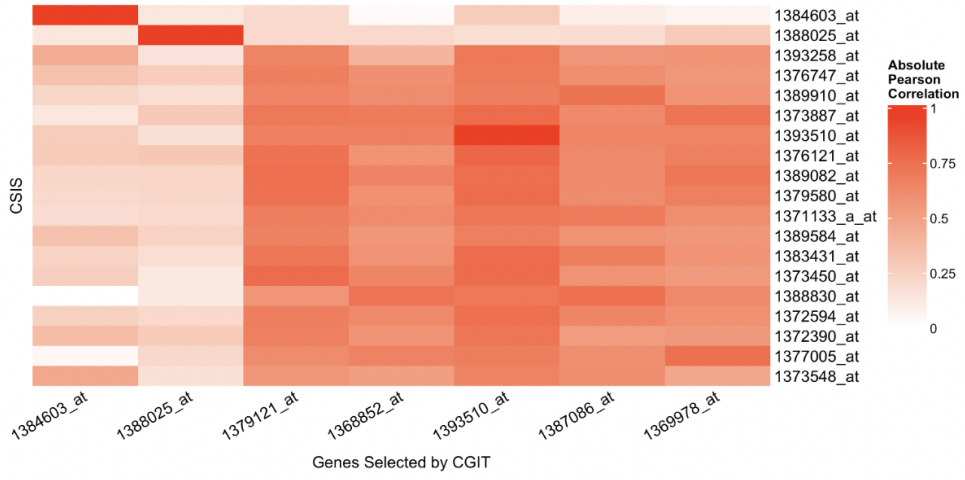


Figure 3.4: Heat map for the correlation between the proposed conditional screening method (CSIS) and Barut et al. (2016)'s method (CSIS).

3.7 Appendix

3.7.1 Theoretical Proofs for The Proposed Global Testing Framework

3.7.1.1 Some Necessary Notations and Regularity Conditions

Denote $\mu(\boldsymbol{b},\tau) = E[\boldsymbol{Z}\{I(Y \leq \boldsymbol{Z}^\intercal \boldsymbol{b}) - \tau\}|\boldsymbol{Z}]$ and let $\widetilde{\boldsymbol{\theta}}(\tau)$ be the solution to $\mu(\boldsymbol{b},\tau) = 0$. We state the following necessary regularity conditions.

Condition 3.1. *$\boldsymbol{Z}$ is uniformly bounded, i.e., $sup_i\|\boldsymbol{Z}_i\| < \infty$.*

Condition 3.2. *(i) $\widetilde{\boldsymbol{\theta}}(\tau)$ is Lipschitz continuous for $\tau \in [\tau_L, \tau_U]$; (ii) $f(y|\boldsymbol{z})$ is bounded above uniformly in y and $\boldsymbol{z}$, where $f(y|\boldsymbol{z}) = dF(y|\boldsymbol{z})/dt$ is Lipschitz continuous for all $y \in R$.*

Condition 3.3. *For some $\rho_0 > 0$ and $c_0 > 0$,$\inf_{\boldsymbol{b}\in\mathcal{B}(\rho_0)} eigmin\boldsymbol{A}(\boldsymbol{b}) \geq c_0$, where $\mathcal{B}(\rho) = \{\boldsymbol{b} \in R^{J+1} : \inf_{\tau\in[\tau_L,\tau_U]} ||\boldsymbol{b} - \widetilde{\boldsymbol{\theta}}(\tau)|| \leq \rho\}$ and $\boldsymbol{A}(\boldsymbol{b}) = E[\boldsymbol{Z}\boldsymbol{Z}^\intercal f(\boldsymbol{Z}^\intercal\boldsymbol{b}|\boldsymbol{Z})]$. Here $||\cdot||$ is the Euclidean norm and $eigmin\boldsymbol{A}(\boldsymbol{b})$ represents the minimal eigenvalue of $\boldsymbol{A}(\boldsymbol{b})$.*

Condition 3.4. *For $j = 1, \ldots, J$, $\inf_{\tau\in[\tau_L,\tau_U]} \sigma^{(j)}(\tau) > 0$, where $\sigma^{(j)}(\tau)$ is the $j+1^{th}$ diagonal element of the variance matrix $\boldsymbol{\Phi}(\tau,\tau)$*

Condition 3.1 is related to the boundedness of the covariates. Condition 3.2 assumes that the coefficient process is smooth and the conditional probability density distribution is bounded and smooth. Both Condition 3.1 and Condition 3.2 are standard assumptions for quantile regression methods and can be satisfied in most practical cases. Condition 3.3 requires that the asymptotic limit of $\boldsymbol{U}_n(\boldsymbol{b},\tau)$ is strictly convex in a neighborhood of $\widetilde{\boldsymbol{\theta}}(\tau)$ for $\tau \in [\tau_L, \tau_U]$, from which we can infer that $\widetilde{\boldsymbol{\theta}}(\tau)$ would be the unique solution to $\mu(\boldsymbol{b},\tau) = E[\boldsymbol{Z}\{I(Y \leq \boldsymbol{Z}^T\boldsymbol{b}) - \tau\}|\boldsymbol{Z}] = 0$, as well as establishing the asymptotic properties. Condition 3.4 states the lower boundedness of the standard deviation of the coefficient.

3.7.1.2 Lemma A3.1, Lemma A3.2 and the Proofs

Lemma A3.1. *Suppose the conditional distribution function of Y given $\boldsymbol{X}_G = \boldsymbol{x}$ is continuous and strictly monotone for all possible values of $\boldsymbol{x}$, and $E(\boldsymbol{Z}\boldsymbol{Z}^\intercal)$ is positive definite. Then $Q_Y(\tau|\boldsymbol{X}_G) = Q_Y(\tau)$ almost surely for $\tau \in [\tau_L, \tau_U]$ is equivalent to model (2) holds with $\beta_0^{(j)}(\tau) = 0$ for $\tau \in [\tau_L, \tau_U]$ for $j = 1, \ldots, J$*

Proof. One the one hand, suppose we have $Q_Y(\tau|\boldsymbol{X}_G) = Q_Y(\tau)$ almost surely for $\tau \in [\tau_L, \tau_U]$. It is clear that we can write $Q_Y(\tau|\boldsymbol{X}_G) = \boldsymbol{Z}^\intercal\boldsymbol{\theta}_0(\tau)$ with $\boldsymbol{\theta}_0(\tau) = (Q_Y(\tau), 0, \ldots, 0)$. Then, we show that given that $E(\boldsymbol{Z}\boldsymbol{Z}^\intercal)$ is positive definite, $\boldsymbol{\theta}_0(\tau)$ is the only solution to $Q_Y(\tau|\boldsymbol{X}_G) = \boldsymbol{Z}^\intercal\boldsymbol{b}$ by contradiction. Let $\boldsymbol{\gamma}_0(\tau)$ denote another solution to $Q_Y(\tau|\boldsymbol{X}_G) = \boldsymbol{Z}^\intercal\boldsymbol{b}$ with $\boldsymbol{\gamma}_0(\tau) \neq \boldsymbol{\theta}_0(\tau)$. This implies

$$\boldsymbol{Z}^\intercal(\boldsymbol{\theta}_0(\tau) - \boldsymbol{\gamma}_0(\tau)) = 0,$$

which leads to

$$(\boldsymbol{\theta}_0(\tau) - \boldsymbol{\gamma}_0(\tau))^\intercal\boldsymbol{Z}\boldsymbol{Z}^\intercal(\boldsymbol{\theta}_0(\tau) - \boldsymbol{\gamma}_0(\tau)) = 0.$$

Given that $E(\boldsymbol{Z}\boldsymbol{Z}^\intercal)$ is positive definite and $\boldsymbol{\gamma}_0(\tau) - \boldsymbol{\theta}_0(\tau) \neq \boldsymbol{0}$, however, we have

$$(\boldsymbol{\theta}_0(\tau) - \boldsymbol{\gamma}_0(\tau))^\intercal\boldsymbol{Z}\boldsymbol{Z}^\intercal(\boldsymbol{\theta}_0(\tau) - \boldsymbol{\gamma}_0(\tau)) > 0.$$

This leads to a contradict. Thus we have shown that $\boldsymbol{\theta}_0(\tau)$ is the unique solution to $Q_Y(\tau) = \boldsymbol{Z}^\intercal\boldsymbol{b}$, i.e. $\beta_0^{(j)}(\tau) = 0$ for $\tau \in [\tau_L, \tau_U]$ for $j = 1, \ldots, J$.

On the other hand, suppose model (2) holds with $\beta_0^{(j)}(\tau) = 0$ for $\tau \in [\tau_L, \tau_U]$ and $j = 1, \ldots, J$, we can write

$$Q_Y(\tau|\boldsymbol{X}_G) = \alpha_0(\tau).$$

Given that the conditional distribution function of Y given $\boldsymbol{X}_G$ is continuous and

strictly monotone, it follows from the definition of $Q_Y(\tau|\boldsymbol{X}_G)$ that for any $\tau \in [\tau_L, \tau_U]$,

$$\text{pr}(Y \leq \alpha_0(\tau)|\boldsymbol{X}_G) = \tau.$$

Taking expectation on both side with respect to $\boldsymbol{X}_G$, we have

$$\text{pr}(Y \leq \alpha_0(\tau)) = \tau.$$

As the conditional distribution function of Y given $\boldsymbol{X}_G = \boldsymbol{x}$ is continuous and strictly monotone, we have the continuity and strictly monotonicity of conditional distribution function of Y. This implies

$$\alpha_0(\tau) = Q_Y(\tau).$$

Thus we have shown that $Q_Y(\tau|\boldsymbol{X}_G) = Q_Y(\tau)$ for $\tau \in [\tau_L, \tau_U]$. This completes the Proof of Lemma A3.1. □

Lemma A3.2. *Suppose the conditional distribution function of Y given $\boldsymbol{X}_G = \boldsymbol{x}$ is continuous and strictly monotone for all possible values of $\boldsymbol{x}$, and $E(\boldsymbol{Z}\boldsymbol{Z}^\intercal)$ is positive definite. Then the solution to $\boldsymbol{\mu}(\boldsymbol{b}, \tau) \doteq E[\boldsymbol{Z}\{I(Y \leq \boldsymbol{Z}^\intercal\boldsymbol{b}) - \tau\}] = 0$, with respect to $\boldsymbol{b} \in R^{J+1}$, uniquely exists and $Q_Y(\tau|\boldsymbol{X}_G) = Q_Y(\tau)$ almost surely for $\tau \in [\tau_L, \tau_U]$ implies $(\widetilde{\beta}^{(1)}(\tau), \ldots, \widetilde{\beta}^{(J)}(\tau))^\intercal = \boldsymbol{0}$.*

Proof. For $\boldsymbol{\mu}(\boldsymbol{b}, \tau) = E[\boldsymbol{Z}\{I(Y \leq \boldsymbol{Z}^\intercal\boldsymbol{b}) - \tau\}]$, we can write

$$\boldsymbol{\mu}(\boldsymbol{b}, \tau) = E[E_Y\{\boldsymbol{Z}[I(Y \leq \boldsymbol{Z}^\intercal\boldsymbol{b}) - \tau]\}|\boldsymbol{Z}] = E[\boldsymbol{Z}\{F_{Y|\boldsymbol{Z}}(\boldsymbol{Z}^\intercal\boldsymbol{b}) - \tau\}].$$

As the conditional distribution function of Y given $\boldsymbol{X}_G = \boldsymbol{x}$ is continuous and strictly monotone, we learn that $F_{Y|\boldsymbol{Z}}(\boldsymbol{Z}^\intercal\boldsymbol{b})$ is continuous and strictly monotone. This also suggests that the conditional density distribution function $f_{Y|\boldsymbol{Z}}(\boldsymbol{Z}^\intercal\boldsymbol{b}) > 0$. Taking

derivative on $\boldsymbol{\mu}(\boldsymbol{b},\tau)$ with respect to $\boldsymbol{b}$, we have

$$\frac{\partial \boldsymbol{\mu}(\boldsymbol{b},\tau)}{\partial \boldsymbol{b}} = E[\boldsymbol{Z}\boldsymbol{Z}^{\intercal} f_{Y|\boldsymbol{Z}}(\boldsymbol{Z}^{\intercal}\boldsymbol{b})].$$

In the next few steps, we want to show that $E[\boldsymbol{Z}\boldsymbol{Z}^{\intercal} f_{Y|\boldsymbol{Z}}(\boldsymbol{Z}^{\intercal}\boldsymbol{b})]$ is positive definite. First, given that $E(\boldsymbol{Z}\boldsymbol{Z}^{\intercal})$ is positive definite. For any non-zero $\boldsymbol{v} \in R^{J+1}$, we have

$$\boldsymbol{v}^{\intercal} E(\boldsymbol{Z}\boldsymbol{Z}^{\intercal})\boldsymbol{v} = E(\boldsymbol{v}^{\intercal}\boldsymbol{Z}\boldsymbol{Z}^{\intercal}\boldsymbol{v}) > 0,$$

which means that $\boldsymbol{Z}^{\intercal}\boldsymbol{v}$ is non-zero. Then, since $f_{Y|\boldsymbol{Z}}(\boldsymbol{Z}^{\intercal}\boldsymbol{b}) > 0$, we have

$$\boldsymbol{v}^{\intercal}\{\boldsymbol{Z}\boldsymbol{Z}^{\intercal} f_{Y|\boldsymbol{Z}}(\boldsymbol{Z}^{\intercal}\boldsymbol{b})\}\boldsymbol{v} = (\boldsymbol{Z}^{\intercal}\boldsymbol{v})^{\intercal} f_{Y|\boldsymbol{Z}}(\boldsymbol{Z}^{\intercal}\boldsymbol{b})\boldsymbol{Z}^{\intercal}\boldsymbol{v} > 0,$$

which leads to

$$\boldsymbol{v}^{\intercal} E[\boldsymbol{Z}\boldsymbol{Z}^{\intercal} f_{Y|\boldsymbol{Z}}(\boldsymbol{Z}^{\intercal}\boldsymbol{b})]\boldsymbol{v} = E[\boldsymbol{v}^{\intercal}\{\boldsymbol{Z}\boldsymbol{Z}^{\intercal} f_{Y|\boldsymbol{Z}}(\boldsymbol{Z}^{\intercal}\boldsymbol{b})\}\boldsymbol{v}] > 0,$$

Thus we can see that $E[\boldsymbol{Z}\boldsymbol{Z}^{\intercal} f_{Y|\boldsymbol{Z}}(\boldsymbol{Z}^{\intercal}\boldsymbol{b})]$ is positive definite for any $\tau \in [\tau_L, \tau_U]$. This leads to the strictly monotonicity of $\boldsymbol{\mu}(\boldsymbol{b},\tau)$, which further suggests that an exact zero-crossing $\widetilde{\boldsymbol{\theta}}(\tau) = (\widetilde{\alpha}_0(\tau), \ldots, \widetilde{\beta}^{(J)}(\tau))^{\intercal}$, as long as it exists, is the unique solution to $\boldsymbol{\mu}(\boldsymbol{b},\tau) = 0$.

As stated in Lemma A3.1, suppose $Q_Y(\tau|\boldsymbol{X}_G) = Q_Y(\tau)$ almost surely for $\tau \in [\tau_L, \tau_U]$, we can write $Q_Y(\tau|\boldsymbol{X}_G) = \boldsymbol{Z}^{\intercal}\boldsymbol{\theta}_0(\tau)$ with $\boldsymbol{\theta}_0(\tau) = (Q_Y(\tau), 0, \ldots, 0)$. From the uniqueness of the solution to $\boldsymbol{\mu}(\boldsymbol{b},\tau) = 0$, we have $(\widetilde{\beta}^{(1)}(\tau), \ldots, \widetilde{\beta}^{(J)}(\tau))^{\intercal} = \boldsymbol{0}$. □

3.7.1.3 Asymptotic Properties Without Assuming the Linear Quantile Regression Model

We establish Theorem A3.1 and Theorem A3.2 for the uniform consistency and weak convergence of $\widehat{\boldsymbol{\theta}}(\tau)$ for $\tau \in [\tau_L, \tau_U] \subseteq (0,1)$:

Theorem A3.1. *Under regularity conditions 3.1-3.3, we have*

$$\lim_{n\to\infty} \sup_{\tau\in[\tau_L,\tau_U]} ||\widehat{\boldsymbol{\theta}}(\tau) - \widetilde{\boldsymbol{\theta}}(\tau)|| \to_p 0.$$

Theorem A3.2. *Under regularity conditions 3.1-3.3,* $\sqrt{n}(\widehat{\boldsymbol{\theta}}(\tau) - \widetilde{\boldsymbol{\theta}}(\tau))$ *converge weakly to a mean zero Gaussian process for* $\tau \in [\tau_L, \tau_U]$ *with covariance*

$$\boldsymbol{\Phi}(\tau', \tau) = \boldsymbol{A}\{\widetilde{\boldsymbol{\theta}}(\tau')\}^{-1} E\{\boldsymbol{\chi}(\tau')\boldsymbol{\chi}(\tau)^{\intercal}\}\boldsymbol{A}\{\widetilde{\boldsymbol{\theta}}(\tau)\}^{-1}.$$

The proofs of Theorem A3.1 and Theorem A3.2 closely resembles that in Peng and Fine (2009) and thus are omitted.

3.7.1.4 Proof of Theorem 3.1

Denote $\widehat{\boldsymbol{R}}(\tau) = (\widehat{R}^{(1)}(\tau), \ldots, \widehat{R}^{(J)}(\tau))$ with $\widehat{R}^{(j)}(\tau) = \widehat{\beta}^{(j)}(\tau)/\widehat{\sigma}_n^{(j)}(\tau)$.

Proof. Following similar lines of Peng and Fine (2009), we can justify the uniform consistency of the estimation for the variance covariance matrix via the sample-based procedure outlined in Section 2.2, i.e.

$$\sup_{\tau\in[\tau_L,\tau_U]} \|\boldsymbol{V}_n(\tau) - \boldsymbol{\Phi}(\tau,\tau)\| \to_p 0, \tag{3.5}$$

which implies

$$\sup_{\tau\in[\tau_L,\tau_U]} |\widehat{\sigma}_n^{(j)}(\tau) - \sigma^{(j)}(\tau)| \to_p 0. \tag{3.6}$$

Under the null hypothesis $H_{0,j}$, we have $\widetilde{\beta}^{(j)}(\tau) = 0$ for $\tau \in [\tau_L, \tau_U]$. This leads to

$$\begin{aligned} n^{1/2}\widehat{R}^{(j)}(\tau) &= \frac{n^{1/2}\{\widehat{\beta}^{(j)}(\tau) - \widetilde{\beta}^{(j)}(\tau)\}}{\widehat{\sigma}_n^{(j)}(\tau)} \\ &= \frac{n^{1/2}\{\widehat{\beta}^{(j)}(\tau) - \widetilde{\beta}^{(j)}(\tau)\}}{\sigma^{(j)}(\tau)} \left(\frac{\sigma^{(j)}(\tau)}{\widehat{\sigma}_n^{(j)}(\tau)} - 1\right) + \frac{n^{1/2}\{\widehat{\beta}^{(j)}(\tau) - \widetilde{\beta}^{(j)}(\tau)\}}{\sigma^{(j)}(\tau)} \end{aligned}$$

By Theorem A3.2, there is weak convergence of $\frac{n^{1/2}\{\widehat{\beta}^{(j)}(\tau)-\widetilde{\beta}^{(j)}(\tau)\}}{\widehat{\sigma}_n^{(j)}(\tau)}$ to a mean zero Gaussian process with covariance process

$$\Phi^{(j)}(\tau,\tau) = \frac{\mathbf{\Phi}^{(j+1,j+1)}(\tau',\tau)}{\sigma^{(j)}(\tau')\sigma^{(j)}(\tau)},$$

where $\mathbf{\Phi}^{(j+1,j+1)}(\tau',\tau)$ denotes the $j+1^{\text{th}}$ diagonal element of $\mathbf{\Phi}(\tau',\tau)$. Also, we have from Condition 3.4 and (3.6) that $\sup_{\tau\in[\tau_L,\tau_U]}\left|\frac{\sigma^{(j)}(\tau)}{\widehat{\sigma}_n^{(j)}(\tau)}-1\right| \to_p 1$. Applying the results in Theorem A3.2 and the Slutsky's Theorem (line 11 of Example 1.4.7 in Boucheron et al. (2013)), we have $n^{1/2}\widehat{R}^{(j)}(\tau) \to_d \mathcal{X}^{(j)}(\tau)$ in $l^\infty(\mathcal{F}_T)$, where $l^\infty(S)$ is the collection of all bounded functions $f: S \mapsto R$ for any index set S and $\mathcal{F}_T = \{\frac{\boldsymbol{\xi}_1^{(j)}(\boldsymbol{c},\tau)}{\sigma^{(j)}(\tau)}, \boldsymbol{c}\in R^2, \tau\in[\tau_L,\tau_U]\}$ with $\boldsymbol{\xi}_1(\boldsymbol{c},\tau) = \boldsymbol{A}(\boldsymbol{c})^{-1}\boldsymbol{Z}_1(I(Y_1 \le \boldsymbol{Z}_1^{\mathrm{T}}\boldsymbol{c}) - \tau)$. Then, by the extended continuous mapping theorem (Theorem 1.11.1 in van der Vaart et al. (1996)), we establish the limiting null distribution for $\widehat{T}_{inte}^{(j)}$ as

$$\widehat{T}_{inte}^{(j)} = \int_{\tau_L}^{\tau_U}\left|\frac{n^{1/2}\widehat{\beta}^{(j)}(\tau)}{\sqrt{\widehat{\sigma}_n^{(j)}(\tau)}}\right|^2 d\tau \to_d \int_{\tau_L}^{\tau_U}\{\mathcal{X}^{(j)}(\tau)\}^2 d\tau.$$

Under $H_{0,G}$, we have $\widetilde{\beta}^{(j)}(\tau) = 0$ for $\tau\in[\tau_L,\tau_U]$ and $j = 1,\ldots,J$. Given that J is finite and using the above results, we can directly obtain the limiting null distributions of the proposed unconditional indices using the continuous mapping theorem:

$$\widehat{T}_{UC} = \max_{j=1,\ldots,J}\{\widehat{T}_{inte}^{(j)}\} \to_d \max_{j=1,\ldots,J}\left\{\int_{\tau_L}^{\tau_U}\{\mathcal{X}^{(j)}(\tau)\}^2 d\tau\right\}.$$

□

3.7.1.5 Theorem A3.3 and the Proof

Theorem A3.3. *Suppose the regularity conditions 3.1-3.4 in Section 3.7 hold. Then* $\widehat{T}_{UC}$ *is consistent against the alternative hypothesis*

$$H_{a,G}: \textit{For some } j_1 \in \{1,\ldots,J\}, \textit{there exists } \tau \in [\tau_L, \tau_U] \textit{ such that } |\widetilde{\beta}^{(j_1)}(\tau)| > 0.$$

Proof. We first look into the asymptotic limit of $\widehat{T}_{UC}$ under the alternative hypothesis $H_{a,G}$. Given $H_{a,G}$ holds, there exists $\tau \in [\tau_L, \tau_U]$ such that $|\widetilde{\beta}^{(j_1)}(\tau)| > 0$ for some j_1. Using the fact that $(a-b)^2 - \frac{1}{2}a^2 + b^2 = \frac{1}{2}(a-2b)^2 \geq 0$ for any $a, b \in R$, we can show that

$$\begin{aligned}
\widehat{T}_{UC} \geq \widehat{T}_{inte}^{(j_1)} = & \int_{\tau_L}^{\tau_U} \left| \frac{n^{1/2}\widehat{\beta}^{(j_1)}(\tau)}{\widehat{\sigma}_n^{(j_1)}(\tau)} \right|^2 d\tau = \int_{\tau_L}^{\tau_U} \left| \frac{n^{1/2}\widetilde{\beta}^{(j_1)}(\tau)}{\widehat{\sigma}_n^{(j_1)}(\tau)} - \frac{n^{1/2}(\widetilde{\beta}^{(j_1)}(\tau) - \widehat{\beta}^{(j_1)}(\tau))}{\widehat{\sigma}_n^{(j_1)}(\tau)} \right|^2 d\tau \\
\geq & \frac{1}{2}\int_{\tau_L}^{\tau_U} \left| \frac{n^{1/2}\widetilde{\beta}^{(j_1)}(\tau)}{\widehat{\sigma}_n^{(j_1)}(\tau)} \right|^2 d\tau - \int_{\tau_L}^{\tau_U} \left| \frac{n^{1/2}(\widetilde{\beta}^{(j_1)}(\tau) - \widehat{\beta}^{(j_1)}(\tau))}{\widehat{\sigma}_n^{(j_1)}(\tau)} \right|^2 d\tau \\
\equiv & \widehat{T}_{inte,1}^{(j_1)} - \widehat{T}_{inte,2}^{(j_1)}
\end{aligned}$$

By the continuous mapping theorem, Condition 3.4, and (3.6), we have $n^{-1}\widehat{T}_{inte,1}^{(j_1)} \to_p \nu_0$, where $\nu_0 = \frac{1}{2}\int_{\tau_L}^{\tau_U} \left| \frac{\widetilde{\beta}^{(j_1)}(\tau)}{\sigma^{(j_1)}(\tau)} \right|^2 d\tau$, and

$$\widehat{T}_{inte,2}^{(j_1)} \to_d \int_{\tau_L}^{\tau_U} \{\mathcal{X}^{(j_1)}(\tau)\}^2 d\tau$$

and thus $O_p(1)$. For any $a > 0$, we can write

$$\begin{aligned}
\text{pr}(\widehat{T}_{UC} > a) \geq & \text{pr}(\widehat{T}_{inte}^{(j_1)} > a) \geq \text{pr}(n^{-1}\widehat{T}_{inte,1}^{(j_1)} > n^{-1}\widehat{T}_{inte,2}^{(j_1)} + a \cdot n^{-1}) \\
\geq & \text{pr}(n^{-1}\widehat{T}_{inte,1}^{(j_1)} > \nu_0/2) - \text{pr}(n^{-1}\widehat{T}_{inte,2}^{(j_1)} + a \cdot n^{-1} > \nu_0/2)
\end{aligned}$$

Under the alternative hypothesis $H_{a,G}$ and Condition 3.4, we have $\nu_0 > 0$, which suggests $\text{pr}(n^{-1}\widehat{T}_{inte,1}^{(j_1)} > \nu_0/2) \to \text{pr}(\nu_0 > \nu_0/2) = 1$ as $n \to \infty$. Also, for any $a > 0$,

we have $n^{-1}\widehat{T}^{(j_1)}_{inte,2}+a\cdot n^{-1}=o_p(1)$, thus $\text{pr}(n^{-1}\widehat{T}^{(j_1)}_{inte,2}+a\cdot n^{-1}>\nu_0/2)\to 0$ as $n\to\infty$. It then follows that $\text{pr}(\widehat{T}_{UC}>a)\to 1$ as $n\to\infty$. Let $a=c_{\alpha,UC}$, where $c_{\alpha,UC}$ is the α-level critical value determined upon the limiting null distribution of $\widehat{T}_{UC}$. We have

$$\text{pr}\left(\left|\widehat{T}_{UC}\right|>c_{\alpha,UC}\right)\to 1$$

as $n\to\infty$, which implies that $\widehat{T}_{UC}$ is consistent against $H_{a,G}$. This completes the Proof of Theorem A3.3. □

3.7.1.6 Justification of the Resampling Procedure

Firstly, as justified in the Proof of Theorem 3.1, for $j=1,\ldots,J$, there is weak convergence of $\frac{n^{1/2}\{\widehat{\beta}^{(j)}(\tau)-\widetilde{\beta}^{(j)}(\tau)\}}{\widehat{\sigma}^{(j)}_n(\tau)}$ to a mean zero Gaussian process with covariance process

$$\Phi^{(j)}(\tau,\tau)=\frac{\boldsymbol{\Phi}^{(j+1,j+1)}(\tau',\tau)}{\sigma^{(j)}(\tau')\sigma^{(j)}(\tau)},$$

where $\boldsymbol{\Phi}^{(j+1,j+1)}(\tau',\tau)$ denotes the $j+1^{\text{th}}$ diagonal element of $\boldsymbol{\Phi}(\tau',\tau)$.

Next, given that $\{\iota^b_i\}^n_{i=1}$ are i.i.d. random variables following a standard normal distribution. Conditional on the observed data $\{(Y_i,\boldsymbol{Z}_i)\}^n_{i=1}$, we have the asymptotic covariance matrix of $\frac{n^{-1/2}\sum^n_{i=1}\widehat{\boldsymbol{\xi}}^{(j)}_i(\tau)\iota^b_i}{\widehat{\sigma}^{(j)}_n(\tau)}$ for $j=1,\ldots,J$ as

$$E\left\{\frac{n^{-\frac{1}{2}}\sum^n_{i=1}\widehat{\boldsymbol{\xi}}^{(j)}_i(\tau')\iota^b_i}{\widehat{\sigma}^{(j)}_n(\tau')}\cdot\frac{n^{-\frac{1}{2}}\sum^n_{i=1}\widehat{\boldsymbol{\xi}}^{(j)}_i(\tau)\iota^b_i}{\widehat{\sigma}^{(j)}_n(\tau)}\middle|\{(Y_i,\boldsymbol{Z}_i)\}^n_{i=1}\right\}=\frac{n^{-1}\sum^n_{i=1}\widehat{\boldsymbol{\xi}}^{(j)}_i(\tau')\widehat{\boldsymbol{\xi}}^{(j)}_i(\tau)}{\widehat{\sigma}^{(j)}_n(\tau')\widehat{\sigma}^{(j)}_n(\tau)},$$

which converges in probability to $\Phi^{(j)}(\tau,\tau)$ due to the uniform consistency of $\widehat{\boldsymbol{\theta}}(\cdot)$ and sample-based variance covariance estimation procedure (Peng and Fine, 2009) and by the law of large numbers.

Following the arguments in Lin et al. (1993), conditional on $\{(Y_i,\boldsymbol{Z}_i)\}^n_{i=1}$, $\frac{n^{-1/2}\sum^n_{i=1}\widehat{\boldsymbol{\xi}}^{(j)}_i(\tau)\iota^b_i}{\widehat{\sigma}^{(j)}_n(\tau)}$ is zero-mean Gaussian with covariance function converge to the same limit as $\frac{n^{1/2}\{\widehat{\beta}^{(j)}(\tau)-\widetilde{\beta}^{(j)}(\tau)\}}{\widehat{\sigma}^{(j)}_n(\tau)}$ for $j=1,\ldots,J$. Given that J is finite, applying the extended continuous mapping

theorem as in the Proof of Theorem 3.1, we have $\widehat{T}_{UC,b}$ conditional on the observed data and T_{UC} are asymptotically equal under $H_{0,G}$. This completes the justification for using the resampling procedure in Section 2.2 to obtain the p values of the proposed tests.

3.7.2 Theoretical Proofs for Screening in Ultra-high Dimensional Setting

3.7.2.1 Some Necessary Notations and Regularity Conditions

For the unconditional screening procedures, we need to impose additional regularity conditions as follow:

Condition 3.5. *For some positive constant* α_0, $\min_{G_l \in M_{[\tau_L,\tau_U]}} n^{-1} T_{UC}^{(G_l)} > \alpha_0 n^{\zeta - 1/2}$.

Condition 3.6. *For some positive constant* α_0, $\min_{\{G_l : G_l \cap M_{[\tau_L,\tau_U]} \neq \varnothing\}} n^{-1} T_{UC}^{(G_l)} > \alpha_0 n^{\zeta - 1/2}$ *and* $\min_{\{\{r_m\} : r_m \in M_{[\tau_L,\tau_U]}\}} n^{-1} T_{UC}^{(\{r_m\})} > \alpha_0 n^{\zeta - 1/2}$.

For the conditional screening procedure, we similarly impose an additional regularity condition:

Condition 3.7. *For some positive constant* α_0, $\min_{G_{c,l} \in M_{[\tau_L,\tau_U]}^{(\mathcal{C})}} n^{-1} T_{C}^{(G_{c,l})} > \alpha_0 n^{\zeta - 1/2}$.

All these conditions, including 3.5, 3.6 and 3.7, assume that the signal for the variables in the active set is strong enough, where a larger ζ indicates a stronger signal.

3.7.2.2 Proofs for Theorem 3.2 and Theorem 3

We first introduce Lemma A3.3, which establish the exponential tail bound of $|\widehat{\beta}^{(j)}(\tau) - \widetilde{\beta}^{(j)}(\tau)|$ and $|\widehat{\sigma}_n^{(j)}(\tau) - \sigma^{(j)}(\tau)|$ for $\tau \in [\tau_L, \tau_U]$ and $j = 1, \ldots, J$ with fixed J.

Lemma A3.3. *Given that the regularity conditions 3.1-3.4 holds. For any* $c > 0$ *and* $0 < \zeta \leq 1/2$,

(A) there exists positive constant c_1 *such that*

$$pr(|\widehat{\beta}^{(j)}(\tau) - \widetilde{\beta}^{(j)}(\tau)| \geq cn^{\zeta-1/2}) \leq 3\exp(-c_1 n^{4\zeta-1});$$

(B) there exists positive constant c_2 *such that*

$$pr(|\widehat{\sigma}_n^{(j)}(\tau) - \sigma^{(j)}(\tau)| \geq cn^{\zeta-1/2}) \leq 3(J+1)\exp(-c_2 n^{4\zeta-1})$$

for any $\tau \in [\tau_L, \tau_U]$ *and* $r = 1, \ldots, p$ *as* n *is sufficiently large.*

Proof. Let $B_n(\boldsymbol{b}) = n^{-1}\sum_{i=1}^n \rho_\tau(Y_i - \boldsymbol{Z}_i^\intercal \boldsymbol{b})$ and $B(\boldsymbol{b}) = E[\rho_\tau(Y - \boldsymbol{Z}^\intercal \boldsymbol{b})]$. From the definition, $\widehat{\boldsymbol{\theta}}(\tau) = \arg\min_{\boldsymbol{b}\in R^{J+1}} B_n(\boldsymbol{b})$ and $\widetilde{\boldsymbol{\theta}}(\tau)$ is equivalent to the unique minimizer of $B(\boldsymbol{b})$. Applying the Lemma 2 of Hjort and Pollard (2011), we can derive from the convexity of the objective function $B(b)$ that for any $\delta > 0$,

$$\begin{aligned} & \mathrm{pr}(|\widehat{\beta}^{(j)}(\tau) - \widetilde{\beta}^{(j)}(\tau)| \geq \delta) \leq \mathrm{pr}(\|\widehat{\boldsymbol{\theta}}(\tau) - \widetilde{\boldsymbol{\theta}}(\tau)\| \geq \delta) \qquad (3.7) \\ \leq \; & \mathrm{pr}\Big(\sup_{\|\boldsymbol{b}-\widetilde{\boldsymbol{\theta}}(\tau)\|\leq\delta} |B_n(\boldsymbol{b}) - B(\boldsymbol{b})| \geq \frac{1}{2}\inf_{\|\boldsymbol{b}-\widetilde{\boldsymbol{\theta}}(\tau)\|=\delta} (B(\boldsymbol{b}) - B(\widetilde{\boldsymbol{\theta}}(\tau)))\Big). \end{aligned}$$

Consider $\boldsymbol{b} = (b^{(0)}, \ldots, b^{(J)})$ such that $\|\boldsymbol{b} - \widetilde{\boldsymbol{\theta}}(\tau)\| = b^{(j)} - \widetilde{\beta}^{(j)}(\tau) = cn^{\zeta-1/2} \equiv \delta$. Following Knight's identity (Knight,1998), which takes the form of $\rho_\tau(u-v) - \rho(u) = v[I(u<0) - \tau] + \int_0^v [I(u \leq s) - I(u \leq 0)]ds$ for $u, v \in R$, we have for $u^* = Y - \boldsymbol{Z}^\intercal \widetilde{\boldsymbol{\theta}}(\tau)$

and $v^* = X^{(j)} cn^{\zeta-1/2}$,

$$
\begin{aligned}
& B(\boldsymbol{b}) - B(\widetilde{\boldsymbol{\theta}}(\tau)) \\
= \; & E\{\rho_\tau(Y - \boldsymbol{Z}^\intercal \widetilde{\boldsymbol{\theta}}(\tau) - X^{(j)} cn^{\zeta-1/2}) - \rho_\tau(Y - \boldsymbol{Z}^\intercal \widetilde{\boldsymbol{\theta}}(\tau))\} \\
= \; & cn^{\zeta-1/2} E\{X^{(j)}[I(Y \le \boldsymbol{Z}^\intercal \widetilde{\boldsymbol{\theta}}(\tau)) - \tau]\} + E\{\int_0^{v^*} [I(Y \le \boldsymbol{Z}^\intercal \widetilde{\boldsymbol{\theta}}(\tau) + s) - I(Y \le \boldsymbol{Z}^\intercal \widetilde{\boldsymbol{\theta}}(\tau))] ds\} \\
= \; & cn^{\zeta-1/2} E\{X^{(j)}[F_{Y|\boldsymbol{X}_{G_l}}(\boldsymbol{Z}^\intercal \widetilde{\boldsymbol{\theta}}(\tau)) - \tau]\} + E\{\int_0^{v^*} [F_{Y|\boldsymbol{X}_{G_l}}(\boldsymbol{Z}^\intercal \widetilde{\boldsymbol{\theta}}(\tau) + s) - F_{Y|\boldsymbol{X}_{G_l}}(\boldsymbol{Z}^\intercal \widetilde{\boldsymbol{\theta}}(\tau))] ds\} \\
\equiv \; & W_1 + W_2.
\end{aligned}
$$

From the definition of $\widetilde{\boldsymbol{\theta}}(\tau)$, we have $|W_1| = cn^{\zeta-1/2} E\{X^{(j)}(\tau - \tau)\} = 0$. Also, for some ξ between $\boldsymbol{Z}^\intercal \widetilde{\boldsymbol{\theta}}(\tau)$ and $\boldsymbol{Z}^\intercal \widetilde{\boldsymbol{\theta}}(\tau) + s$, we can write

$$
|W_2| = E\{\int_0^{v^*} f_{Y|\boldsymbol{X}_{G_l}}(\xi) s ds\} = 1/2 f_{Y|\boldsymbol{X}_{G_l}}(\xi) E\{(cn^{\zeta-1/2} X^{(j)})^2\} = O(n^{2\zeta-1}).
$$

This indicates that there exists a positive constant γ_1 such that $\inf_{\|\boldsymbol{b} - \widetilde{\boldsymbol{\theta}}(\tau)\| = \delta}(B(\boldsymbol{b}) - B(\widetilde{\boldsymbol{\theta}}(\tau)) \ge \gamma_1 n^{2\zeta-1}$, which leads to

$$
\begin{aligned}
& \mathrm{pr}(|\widehat{\beta}^{(j)}(\tau) - \widetilde{\beta}^{(j)}(\tau)| \ge \delta) \le \mathrm{pr}(\sup_{\|\boldsymbol{b} - \widetilde{\boldsymbol{\theta}}(\tau)\| \le \delta} |B_n(\boldsymbol{b}) - B(\boldsymbol{b})| \ge \frac{\gamma_1}{2} n^{2\zeta-1}) \\
\le \; & \mathrm{pr}(|B_n(\widetilde{\boldsymbol{\theta}}(\tau)) - B(\widetilde{\boldsymbol{\theta}}(\tau))| \ge \frac{\gamma_1}{4} n^{2\zeta-1}) \\
& + \mathrm{pr}(\sup_{\|\boldsymbol{b} - \widetilde{\boldsymbol{\theta}}(\tau)\| \le \delta} |B_n(\boldsymbol{b}) - B_n(\widetilde{\boldsymbol{\theta}}(\tau)) - B(\boldsymbol{b}) + B(\widetilde{\boldsymbol{\theta}}(\tau))| \ge \frac{\gamma_1}{4} n^{2\zeta-1}) \\
\equiv \; & D_1 + D_2.
\end{aligned}
$$

Let $Q_i = \rho_\tau(Y_i - \boldsymbol{Z}_i^\intercal \widetilde{\boldsymbol{\theta}}(\tau))$. Then $B_n(\widetilde{\boldsymbol{\theta}}(\tau)) - B(\widetilde{\boldsymbol{\theta}}(\tau)) = n^{-1} \sum_{i=1}^n (Q_i - E\{Q_i\})$. From the conditions 3.1 and Condition 3.2, we learn that Q_i is bounded, i.e., $\sup_i |Q_i| < \infty$. Then there exist M, such that $|Q_i - E\{Q_i\}|$ is uniformly bounded above by M.

Applying Bernsterin's inequality, we have

$$\begin{aligned} D_1 &= \text{pr}\left(|B_n(\widetilde{\boldsymbol{\theta}}(\tau)) - B(\widetilde{\boldsymbol{\theta}}(\tau))| \geq \frac{\gamma_1}{4} n^{2\zeta-1}\right) = \text{pr}\left(|\sum_{i=1}^{n}(Q_i - E\{Q_i\})| \geq \frac{\gamma_1}{4} n^{2\zeta}\right) \\ &\leq 2\exp(-\frac{\gamma_1^2 n^{4\zeta-1}/32}{M^2 + M\gamma_1 n^{2\zeta-1}/12}) \leq 2\exp(-c_1 n^{4\zeta-1}) \end{aligned}$$

for some c_1. Then we work on D_2 by introducing Lemma A3.4.

Lemma A3.4 (Massart's concentration theorem, 2000)**.** *Let $W_1, \ldots, W_n$ be independent random variables and let $\mathcal{G}$ be a class of functions satisfying $a_{i,g} \leq g(W_i) \leq b_{i,g}$ for some real numbers $a_{i,g}$ and $b_{i,g}$, and for all $1 \leq i \leq n$ and $g \in \mathcal{G}$. Define $L^2 = \sup_{g\in\mathcal{G}} \sum_{i=1}^{n}(b_{i,g} - a_{i,g})^2/n$ and $Z = \sup_{g\in\mathcal{G}} |\sum_{i=1}^{n}(g(W_i) - E(g(W_i)))|$. Then for any positive t,*

$$pr(Z \geq EZ + t) \leq \exp[-\frac{nt^2}{2L^2}].$$

To apply Lemma A3.4, first we check the conditions. Let $H_i = \rho_\tau(Y_i - \boldsymbol{Z}_i^\intercal \boldsymbol{b}) - \rho_\tau(Y_i - \boldsymbol{Z}_i^\intercal \widetilde{\boldsymbol{\theta}}(\tau))$. From Knight's identity, we have

$$\begin{aligned} H_i &= \boldsymbol{Z}_i^\intercal(\boldsymbol{b} - \widetilde{\boldsymbol{\theta}}(\tau))[I(Y_i \leq \boldsymbol{Z}_i^\intercal \widetilde{\boldsymbol{\theta}}(\tau)) - \tau] \\ &\quad + \int_0^{\boldsymbol{Z}_i^\intercal(\boldsymbol{b}-\widetilde{\boldsymbol{\theta}}(\tau))} [I(Y_i - \boldsymbol{Z}_i^\intercal \widetilde{\boldsymbol{\theta}}(\tau) \leq s) - I(Y_i \leq \boldsymbol{Z}_i^\intercal \widetilde{\boldsymbol{\theta}}(\tau))]ds \end{aligned}$$

This leads to $|H_i| \leq 2|\boldsymbol{Z}_i^\intercal(\boldsymbol{b} - \widetilde{\boldsymbol{\theta}}(\tau))| = 2X_i^{(j)} c n^{\zeta-1/2} \leq c' n^{\zeta-1/2}$ for some positive constant c'.

Define $D^* = \sup_{\|\boldsymbol{b}-\widetilde{\boldsymbol{\theta}}(\tau)\| \leq c n^{\zeta-1/2}} n^{-1} |\sum_{i=1}^{n}(H_i - E\{H_i\})|$. Let $e_1, \ldots, e_n$ be a Rademacher sequence, which are i.i.d. taking values of ± 1 with probability $1/2$ independent of $H_1, \ldots, H_n$. We have from the symmetrization theorem (Lemma 2.3.1

in van der Vaart et al. (1996))

$$\begin{aligned} E(D^*) &= E\left\{ \sup_{\|\boldsymbol{b}-\widetilde{\boldsymbol{\theta}}(\tau)\|\le cn^{\zeta-1/2}} n^{-1}|\sum_{i=1}^{n}(H_i - E\{H_i\})| \right\} \\ &\le 2E\left\{ \sup_{\|\boldsymbol{b}-\widetilde{\boldsymbol{\theta}}(\tau)\|\le cn^{\zeta-1/2}} n^{-1}|\sum_{i=1}^{n} e_i H_i| \right\}. \end{aligned}$$

Applying the contraction theorem in Ledoux and Talagrand (2013), we have

$$E\left\{ \sup_{\|\boldsymbol{b}-\widetilde{\boldsymbol{\theta}}(\tau)\|\le cn^{\zeta-1/2}} n^{-1}|\sum_{i=1}^{n} e_i H_i| \right\} \le 2E\left\{ \sup_{\|\boldsymbol{b}-\widetilde{\boldsymbol{\theta}}(\tau)\|\le cn^{\zeta-1/2}} n^{-1}|\sum_{i=1}^{n} e_i \boldsymbol{Z}_i^{\tau}(\boldsymbol{b}-\widetilde{\boldsymbol{\theta}}(\tau))| \right\}. \tag{3.8}$$

From (3.8), Condition 3.3, and $e_i^2 = 1$ with probability 1, we may further derive that there exists some positive constant c'', such that

$$\begin{aligned} E(D^*) &\le 4E\left\{ \sup_{\|\boldsymbol{b}-\widetilde{\boldsymbol{\theta}}(\tau)\|\le cn^{\zeta-1/2}} n^{-1}|\sum_{i=1}^{n} e_i \boldsymbol{Z}_i^{\tau}(\boldsymbol{b}-\widetilde{\boldsymbol{\theta}}(\tau))| \right\} \le 4cn^{\zeta-1/2} E\|n^{-1}\sum_{i=1}^{n} e_i \boldsymbol{Z}_i\| \\ &\le 4cn^{\zeta-1/2}[E\|n^{-1}\sum_{i=1}^{n} e_i \boldsymbol{Z}_i\|^2]^{1/2} = 4cn^{\zeta-1/2}[n^{-2}E(\sum_{i=1}^{n} \boldsymbol{Z}_i^{\tau}\boldsymbol{Z}_i)]^{1/2} \le c''n^{\zeta-1}. \end{aligned}$$

Take $t = \gamma_1 n^{2\zeta-1}/4 - c''n^{\zeta-1}$ and $L^2 = 4(c')^2 n^{2\zeta-1}$. Applying Lemma A3.4,

$$\begin{aligned} D_2 &= \mathrm{pr}(D^* \ge \gamma_1 n^{2\zeta-1}/4) = \mathrm{pr}(D^* \ge E(D^*) + (\gamma_1 n^{2\zeta-1}/4 - E(D^*))) \\ &\le \mathrm{pr}(D^* \ge E(D^*) + (\gamma_1 n^{2\zeta-1}/4 - c''n^{\zeta-1})) \\ &\le \exp\left(-\frac{n(\gamma_1 n^{2\zeta-1}/4 - c''n^{\zeta-1})^2}{8(c')^2 n^{2\zeta-1}}\right) \le \exp(-c_2 n^{2\zeta}) \end{aligned}$$

for some positive constant c_2 as n is sufficiently large. Given $0 < \zeta < 1/2$, we have $4\zeta - 1 < 2\zeta$. This indicates that there exists positive constant c_1 and c_2 such that

$$\mathrm{pr}(|\widehat{\beta}^{(j)}(\tau) - \widetilde{\beta}^{(j)}(\tau)| \ge cn^{\zeta-1/2}) \le 2\exp(-c_1 n^{4\zeta-1}) + \exp(-c_2 n^{2\zeta}) \le 3\exp(-c_1 n^{4\zeta-1})$$

for any $\tau \in [\tau_L, \tau_U]$ and $r = 1, \ldots, p$. This completes the Proof of (A).

Let $R_n(\boldsymbol{b}) = n^{-1}\sum_{i=1}^{n}[\rho_\tau(Y_i - \boldsymbol{Z}_i^\intercal \boldsymbol{b}) + n^{-1/2}u \cdot \boldsymbol{e}_{n,k}(\tau)^\intercal \boldsymbol{b}], k = 1, \ldots, J+1$, where $\|n^{-1/2}u \cdot \boldsymbol{e}_{n,k}(\tau)\| = O(n^{-1/2})$. From the definition, $\widetilde{\boldsymbol{c}}_{n,k}(\tau) = \arg\min_{\boldsymbol{b}\in R^2} R_n(\boldsymbol{b})$. Using similar arguments as the Proof of (A), we can derive that for any $c > 0$ and $0 < \zeta \leq 1/2$, there exists positive constant $m_1^{(k)}$ such that

$$\mathrm{pr}(\|\widetilde{\boldsymbol{c}}_{n,k}(\tau) - \widetilde{\boldsymbol{\theta}}(\tau)\| \geq cn^{\zeta-1/2}) \leq 3\exp(-m_1^{(k)}n^{4\zeta-1}).$$

for any $\tau \in [\tau_L, \tau_U]$ and $k = 1, \ldots, J+1$ as n is sufficiently large.

Denote $\boldsymbol{d}_{n,k}(\tau) = \{\widetilde{\boldsymbol{c}}_{n,k}(\tau) - \widehat{\boldsymbol{\theta}}(\tau)\}$. Using similar arguments as in Peng and Fine (2009), we can derive $\sup_\tau \|\widetilde{\boldsymbol{c}}_{n,k}(\tau) - \widetilde{\boldsymbol{\theta}}(\tau)\| \to 0$, a.s., and

$$\boldsymbol{S}_n\{\widetilde{\boldsymbol{c}}_{n,k}(\tau), \tau\} - \boldsymbol{S}_n\{\widehat{\boldsymbol{\theta}}(\tau), \tau\} = \boldsymbol{A}\{\boldsymbol{\theta}_k^*(\tau)\} \cdot n^{1/2}\{\widetilde{\boldsymbol{c}}_{n,k}(\tau) - \widehat{\boldsymbol{\theta}}(\tau)\} = \boldsymbol{A}\{\boldsymbol{\theta}_k^*(\tau)\} \cdot n^{1/2}\boldsymbol{d}_{n,k}(\tau),$$

where $\boldsymbol{\theta}_k^*(\tau)$ is between $\widetilde{\boldsymbol{c}}_{n,k}(\tau)$ and $\widehat{\boldsymbol{\theta}}(\tau)$. As $\boldsymbol{S}_n(\widetilde{\boldsymbol{c}}_{n,k}(\tau), \tau) \approx u\boldsymbol{e}_{n,k}(\tau)$ and $\boldsymbol{S}_n\{\widehat{\boldsymbol{\theta}}(\tau), \tau\} \approx 0$, we have

$$n^{1/2}\boldsymbol{d}_{n,k}(\tau) - \boldsymbol{A}\{\widetilde{\boldsymbol{\theta}}(\tau)\}^{-1}u\boldsymbol{e}_{n,k}(\tau) \approx \{\boldsymbol{A}\{\boldsymbol{\theta}_k^*(\tau)\}^{-1} - \boldsymbol{A}\{\widetilde{\boldsymbol{\theta}}(\tau)\}^{-1}\}u\boldsymbol{e}_{n,k}(\tau).$$

As $\widetilde{\boldsymbol{c}}_{n,k}(\tau)$ and $\widehat{\boldsymbol{\theta}}(\tau)$ both belong to $\mathcal{B}(\rho_0)$ with probability 1 as n is large enough, we can conclude that $\boldsymbol{\theta}_k^*(\tau)$ belongs to $\mathcal{B}(\rho_0)$ with probability 1. By Condition 3.2 and Condition 3.3, we learn that $\boldsymbol{A}(\boldsymbol{b})$ is Lipschitz continuous and uniformly bounded both above and below for any $\boldsymbol{b} \in \mathcal{B}(\rho_0)$. Then there exists positive constant $C_{\boldsymbol{A},1}^{(k)}$ and $C_{\boldsymbol{A},2}^{(k)}$, such that $\|\boldsymbol{A}\{\boldsymbol{\theta}_k^*(\tau)\}^{-1} - \boldsymbol{A}\{\widetilde{\boldsymbol{\theta}}(\tau)\}^{-1}\| \leq C_{\boldsymbol{A},1}^{(k)}\|\boldsymbol{\theta}_k^*(\tau) - \widetilde{\boldsymbol{\theta}}(\tau)\|$ and $\|\boldsymbol{A}\{\boldsymbol{\theta}_k^*(\tau)\}^{-1}\| \leq C_{\boldsymbol{A},2}^{(k)}$. Then we have for sufficiently large n,

$$\begin{aligned} &\|n^{1/2}\boldsymbol{d}_{n,k}(\tau) - \boldsymbol{A}\{\widetilde{\boldsymbol{\theta}}(\tau)\}^{-1}u\boldsymbol{e}_{n,k}(\tau)\| \leq 2\|\{\boldsymbol{A}\{\boldsymbol{\theta}_k^*(\tau)\}^{-1} - \boldsymbol{A}\{\widetilde{\boldsymbol{\theta}}(\tau)\}^{-1}\}u\boldsymbol{e}_{n,k}(\tau)\| \\ \leq\ & 2C_{\boldsymbol{A},1}^{(k)}\|\boldsymbol{\theta}_k^*(\tau) - \widetilde{\boldsymbol{\theta}}(\tau)\|\|u\boldsymbol{e}_{n,k}(\tau)\| \leq 2C_{\boldsymbol{A},1}^{(k)}\|\widetilde{\boldsymbol{c}}_{n,k}(\tau) - \widetilde{\boldsymbol{\theta}}(\tau)\|\|u\boldsymbol{e}_{n,k}(\tau)\|. \end{aligned}$$

Using the fact that $n^{-1/2}u \cdot e_{n,k}(\tau) = O(n^{-1/2})$, almost surely for $k = 1, \ldots, J+1$, we have $2C^{(k)}_{\boldsymbol{A},1}\|\widetilde{\boldsymbol{c}}_{n,k}(\tau) - \widetilde{\boldsymbol{\theta}}(\tau)\|\|u\boldsymbol{e}_{n,k}(\tau)\| = o(1)$ and $\|\boldsymbol{A}\{\widetilde{\boldsymbol{\theta}}(\tau)\}^{-1}u\boldsymbol{e}_{n,k}(\tau)\| = O(1)$. All the above lead to

$$\begin{aligned} & \|n^{1/2}\boldsymbol{d}_{n,k}(\tau) + \boldsymbol{A}\{\widetilde{\boldsymbol{\theta}}(\tau)\}^{-1}u\boldsymbol{e}_{n,k}(\tau)\| \\ \leq\ & \|n^{1/2}\boldsymbol{d}_{n,k}(\tau) - \boldsymbol{A}\{\widetilde{\boldsymbol{\theta}}(\tau)\}^{-1}u\boldsymbol{e}_{n,k}(\tau)\| + 2\|\boldsymbol{A}\{\widetilde{\boldsymbol{\theta}}(\tau)\}^{-1}u\boldsymbol{e}_{n,k}(\tau)\| \\ \leq\ & 3\|\boldsymbol{A}\{\widetilde{\boldsymbol{\theta}}(\tau)\}^{-1}u\boldsymbol{e}_{n,k}(\tau)\| \end{aligned}$$

as n is large enough. Then we can write

$$\begin{aligned} & \mathrm{pr}(|\widehat{\sigma}^{(j)}_n(\tau) - \sigma^{(j)}(\tau)| \geq cn^{\zeta-1/2}) \leq \mathrm{pr}(\|\widehat{\boldsymbol{\Phi}}(\tau,\tau) - \boldsymbol{\Phi}(\tau,\tau)\| \geq cn^{\zeta-1/2}) \\ \leq\ & \mathrm{pr}(\|nu^{-2}\boldsymbol{D}_n(\tau)\boldsymbol{D}_n(\tau)^{\intercal} - \boldsymbol{A}\{\widetilde{\boldsymbol{\theta}}(\tau)\}^{-1}\boldsymbol{E}_n(\tau)\boldsymbol{E}_n(\tau)^{\intercal}\{\boldsymbol{A}\{\widetilde{\boldsymbol{\theta}}(\tau)\}^{-1}\}^{\intercal}\| \geq cn^{\zeta-1/2}) \\ \leq\ & \mathrm{pr}(u^{-2}\|\sum_{k=1}^{J+1} n\boldsymbol{d}_{n,k}(\tau)\boldsymbol{d}_{n,k}(\tau)^{\intercal} - u^2\boldsymbol{A}\{\widetilde{\boldsymbol{\theta}}(\tau)\}^{-1}\boldsymbol{e}_{n,k}(\tau)\boldsymbol{e}_{n,k}(\tau)^{\intercal}\{\boldsymbol{A}\{\widetilde{\boldsymbol{\theta}}(\tau)\}^{-1}\}^{\intercal}\| \geq cn^{\zeta-1/2}) \\ \leq\ & \mathrm{pr}(\sum_{k=1}^{J+1} u^{-2}\|n\boldsymbol{d}_{n,k}(\tau)\boldsymbol{d}_{n,k}(\tau)^{\intercal} - u^2\boldsymbol{A}\{\widetilde{\boldsymbol{\theta}}(\tau)\}^{-1}\boldsymbol{e}_{n,k}(\tau)\boldsymbol{e}_{n,k}(\tau)^{\intercal}\{\boldsymbol{A}\{\widetilde{\boldsymbol{\theta}}(\tau)\}^{-1}\}^{\intercal}\| \geq cn^{\zeta-1/2}) \\ \leq\ & \sum_{k=1}^{J+1} \mathrm{pr}(u^{-2}\|n\boldsymbol{d}_{n,k}(\tau)\boldsymbol{d}_{n,k}(\tau)^{\intercal} - u^2\boldsymbol{A}\{\widetilde{\boldsymbol{\theta}}(\tau)\}^{-1}\boldsymbol{e}_{n,k}(\tau)\boldsymbol{e}_{n,k}(\tau)^{\intercal}\{\boldsymbol{A}\{\widetilde{\boldsymbol{\theta}}(\tau)\}^{-1}\}^{\intercal}\| \geq \frac{cn^{\zeta-1/2}}{J+1}) \\ \leq\ & \sum_{k=1}^{J+1} \mathrm{pr}(u^{-2}\|n^{1/2}\boldsymbol{d}_{n,k}(\tau) + \boldsymbol{A}\{\widetilde{\boldsymbol{\theta}}(\tau)\}^{-1}u\boldsymbol{e}_{n,k}(\tau)\| \cdot \|\{n^{1/2}\boldsymbol{d}_{n,k}(\tau) - \boldsymbol{A}\{\widetilde{\boldsymbol{\theta}}(\tau)\}^{-1}u\boldsymbol{e}_{n,k}(\tau)\}^{\intercal}\| \geq \frac{cn^{\zeta-1/2}}{J+1}) \\ \leq\ & \sum_{k=1}^{J+1} \mathrm{pr}(3u^{-2}\|\boldsymbol{A}\{\widetilde{\boldsymbol{\theta}}(\tau)\}^{-1}u\boldsymbol{e}_n(\tau)\| \cdot \|n^{1/2}\boldsymbol{d}_{n,k}(\tau) - \boldsymbol{A}\{\widetilde{\boldsymbol{\theta}}(\tau)\}^{-1}u\boldsymbol{e}_{n,k}(\tau)\| \geq \frac{cn^{\zeta-1/2}}{J+1}) \\ \leq\ & \sum_{k=1}^{J+1} \mathrm{pr}(3u^{-2}C^{(k)}_{\boldsymbol{A},2}\|u\boldsymbol{e}_{n,k}(\tau)\| \cdot 2C^{(k)}_{\boldsymbol{A},1}\|\widetilde{\boldsymbol{c}}_{n,k}(\tau) - \widetilde{\boldsymbol{\theta}}(\tau)\|\|u\boldsymbol{e}_{n,k}(\tau)\| \geq \frac{cn^{\zeta-1/2}}{J+1}) \\ & (\|\boldsymbol{e}_{n,k}(\tau)\| = 1 \text{ as the norm of eigenvectors of real symmetric matrices is always 1.}) \\ \leq\ & \sum_{k=1}^{J+1} \mathrm{pr}(6 \max_{k=1,\ldots,J+1}\{C^{(k)}_{\boldsymbol{A},2}\} \cdot \max_{k=1,\ldots,J+1}\{C^{(k)}_{\boldsymbol{A},1}\}\|\widetilde{\boldsymbol{c}}_{n,k}(\tau) - \widetilde{\boldsymbol{\theta}}(\tau)\| \geq \frac{cn^{\zeta-1/2}}{J+1}) \\ \leq\ & \sum_{k=1}^{J+1} \mathrm{pr}(\|\widetilde{\boldsymbol{c}}_{n,k}(\tau) - \widetilde{\boldsymbol{\theta}}(\tau)\| \geq c^* n^{\zeta-1/2}) \text{ (Let } c^* = c/\{6(J+1)\max_k\{C^{(k)}_{\boldsymbol{A},2}\}\max_k\{C^{(k)}_{\boldsymbol{A},1}\}\}) \\ \leq\ & \sum_{k=1}^{J+1} 3\exp(-m^{(k)}_1 n^{4\zeta-1}) \leq 3(J+1)\exp(-c_2 n^{4\zeta-1}) \end{aligned}$$

for some positive constant $c_2 = \min_{k=1,\ldots,J+1}\{m_1^{(k)}\}$ and sufficiently large n. This completes the Proof of (B). □

We further introduce Lemma A3.5, which establish the exponential bound for $\sup_{\tau\in[\tau_L,\tau_U]}|\widehat{\beta}^{(j)}(\tau)-\widetilde{\beta}^{(j)}(\tau)|$ and $\sup_{\tau\in[\tau_L,\tau_U]}|\widehat{\sigma}_n^{(j)}(\tau)-\sigma^{(j)}(\tau)|$.

Lemma A3.5. *Given that the regularity conditions 3.1-3.4 holds. For any $c>0$ and $0<\zeta\leq 1/2$, (i) there exists positive constant c_3 and η_3 such that*

$$pr\left(\sup_{\tau\in[\tau_L,\tau_U]}|\widehat{\beta}^{(j)}(\tau)-\widetilde{\beta}^{(j)}(\tau)|>cn^{\zeta-1/2}\right)\leq c_3\exp(-\eta_3 n^{4\zeta-1}-\log(n^{\zeta-1/2}));$$

(ii) there exists positive constant c_4 and η_4 such that

$$pr\left(\sup_{\tau\in[\tau_L,\tau_U]}|\widehat{\sigma}_n^{(j)}(\tau)-\sigma^{(j)}(\tau)|>cn^{\zeta-1/2}\right)\leq c_4\exp(-\eta_4 n^{4\zeta-1}-\log(n^{\zeta-1/2})).$$

Proof. We first prove (i). Note that by Condition 3.2, $\widetilde{\beta}^{(j)}(\tau)$ is uniformly bounded and Lipschitz continuous. Then there exist $w_0\geq 0$, such that $|\widetilde{\beta}^{(j)}(\tau_a)-\widetilde{\beta}^{(j)}(\tau_b)|\leq w_0|\tau_a-\tau_b|$. Given $0<\delta<\min\{1,\frac{3w_0}{(\tau_U-\tau_L)}\}$, we can define a grid partition for $[\tau_L,\tau_U]$ as $\tau_L=\tau_0<\tau_1<\cdots<\tau_{N_\delta}=\tau_U$ with $N_\delta=\lfloor\frac{6w_0(\tau_U-\tau_L)}{\delta}\rfloor$, where $\lfloor x\rfloor$ denotes the largest integer smaller than x. With this partition, the grid would be $|\tau_k-\tau_{k-1}|\leq\frac{\delta}{3w_0}$. Then we have $|\widetilde{\beta}^{(j)}(\tau_k)-\widetilde{\beta}^{(j)}(\tau_{k-1})|\leq\frac{\delta}{3}$.

We learn from the definition for $\widehat{\beta}^{(j)}(\tau_k)$ and $\widetilde{\beta}^{(j)}(\tau_k)$ that they are non-decreasing. Suppose $|\widehat{\beta}^{(j)}(\tau_k)-\widetilde{\beta}^{(j)}(\tau_k)|\leq\frac{\delta}{3}$ and $|\widehat{\beta}^{(j)}(\tau_{k-1})-\widetilde{\beta}^{(j)}(\tau_{k-1})|\leq\frac{\delta}{3}$, then for $\tau_{k-1}\leq x\leq\tau_k$, we have

$$\widehat{\beta}^{(j)}(x)-\widetilde{\beta}^{(j)}(x)\leq\widehat{\beta}^{(j)}(\tau_k)-\widetilde{\beta}^{(j)}(\tau_k)+\widetilde{\beta}^{(j)}(\tau_k)-\widetilde{\beta}^{(j)}(x)\leq\frac{2\delta}{3}<\delta.$$

The other direction can be shown by the same arguments. Then we have $|\widehat{\beta}^{(j)}(x)-\widetilde{\beta}^{(j)}(x)|<\delta$ for any $\tau_{k-1}\leq x\leq\tau_k$. Thus if $\sup_{\tau\in[\tau_L,\tau_U]}|\widehat{\beta}^{(j)}(\tau)-\widetilde{\beta}^{(j)}(\tau)|>\delta$, there

exist some $0 \leq k \leq N_\delta$ such that $|\widehat{\beta}^{(j)}(\tau_k) - \widetilde{\beta}^{(j)}(\tau_k)| > \frac{\delta}{3}$. Let $\delta = cn^{\zeta-1/2}$. Applying Lemma A3.3 (A), there exist some positive constant η_3 s.t.

$$\begin{aligned} & \text{pr}\left(\sup_{\tau\in[\tau_L,\tau_U]} |\widehat{\beta}^{(j)}(\tau) - \widetilde{\beta}^{(j)}(\tau)| > cn^{\zeta-1/2}\right) \leq N_\delta \text{pr}\left(|\widehat{\beta}^{(j)}(\tau_k) - \widetilde{\beta}^{(j)}(\tau_k)| > \frac{c}{3}n^{\zeta-1/2}\right) \\ \leq \quad & 18w_0(\tau_U - \tau_L)\exp(-\eta_3 n^{4\zeta-1} - \log(cn^{\zeta-1/2})) \equiv c_3\exp(-\eta_3 n^{4\zeta-1} - \log(n^{\zeta-1/2})), \end{aligned}$$

where $c_3 = 18w_0(\tau_U - \tau_L)/c$. This completes the Proof of (i).

By Condition 3.2, Condition 3.3 and Condition 3.4, we learn that $\sigma^{(j)}(\tau)$ is also uniformly bounded and Lipschitz continuous. Thus, using similar arguments as those for (i) and applying Lemma A3.3 (B), we can show that there exist some positive constant c_4 and η_4 such that

$$\text{pr}\left(\sup_{\tau\in[\tau_L,\tau_U]} |\widehat{\sigma}_n^{(j)}(\tau) - \sigma^{(j)}(\tau)| > cn^{\zeta-1/2}\right) \leq c_4\exp(-\eta_4 n^{4\zeta-1} - \log(n^{\zeta-1/2})).$$

for sufficiently large n. This completes the Proof of (ii). □

for Theorem 3.2. From the definition, we can write

$$\begin{aligned} |n^{-1}\widehat{T}_{UC}^{(G_l)} - n^{-1}T_{UC}^{(G_l)}| \quad &= \quad \left| n^{-1}\max_{j\in G_l}\int_{\tau_L}^{\tau_U}\left|\frac{\sqrt{n}\widehat{\beta}^{(j)}(\tau)}{\widehat{\sigma}_n^{(j)}(\tau)}\right|^2 d\tau - n^{-1}\max_{j\in G_l}\int_{\tau_L}^{\tau_U}\left|\frac{\sqrt{n}\widetilde{\beta}^{(j)}(\tau)}{\sigma^{(j)}(\tau)}\right|^2 d\tau\right| \\ &\leq \quad \max_{j\in G_l}\left|\int_{\tau_L}^{\tau_U}\left|\frac{\widehat{\beta}^{(j)}(\tau)}{\widehat{\sigma}_n^{(j)}(\tau)}\right|^2 d\tau - \int_{\tau_L}^{\tau_U}\left|\frac{\widetilde{\beta}^{(j)}(\tau)}{\sigma^{(j)}(\tau)}\right|^2 d\tau\right| \end{aligned}$$

We can further write

$$
\begin{aligned}
& \int_{\tau_L}^{\tau_U} \left| \frac{\widehat{\beta}^{(j)}(\tau)}{\widehat{\sigma}_n^{(j)}(\tau)} \right|^2 d\tau - \int_{\tau_L}^{\tau_U} \left| \frac{\widetilde{\beta}^{(j)}(\tau)}{\sigma^{(j)}(\tau)} \right|^2 d\tau \\
= & \int_{\tau_L}^{\tau_U} \left| \frac{[\widehat{\beta}^{(j)}(\tau) - \widetilde{\beta}^{(j)}(\tau)]}{\widehat{\sigma}_n^{(j)}(\tau)} \right|^2 d\tau + 2 \int_{\tau_L}^{\tau_U} \frac{\widetilde{\beta}^{(j)}(\tau)}{\sigma^{(j)}(\tau)} \frac{\widehat{\beta}^{(j)}(\tau) - \widetilde{\beta}^{(j)}(\tau)}{\widehat{\sigma}_n^{(j)}(\tau)} d\tau \\
& + \int_{\tau_L}^{\tau_U} \left(\frac{\{\sigma^{(j)}(\tau)\}^2}{\{\widehat{\sigma}_n^{(j)}(\tau)\}^2} - 1 \right) \left| \frac{\widetilde{\beta}^{(j)}(\tau)}{\sigma^{(j)}(\tau)} \right|^2 d\tau \\
& + 2 \int_{\tau_L}^{\tau_U} \left(\frac{\sigma^{(j)}(\tau)}{\widehat{\sigma}_n^{(j)}(\tau)} - 1 \right) \frac{\widetilde{\beta}^{(j)}(\tau)}{\sigma^{(j)}(\tau)} \frac{\widehat{\beta}^{(j)}(\tau) - \widetilde{\beta}^{(j)}(\tau)}{\widehat{\sigma}_n^{(j)}(\tau)} d\tau \\
\equiv & A_{j,1} + A_{j,2} + A_{j,3} + A_{j,4}
\end{aligned}
$$

From the uniform consistency and weak convergence of $\widehat{\beta}^{(j)}(\tau)$, and uniform consistency of $\widehat{\sigma}_n^{(j)}(\tau)$, we have $A_{j,4} = o(n^{-1/2}) = o(n^{\zeta - 1/2})$, which implies $A_{j,4} \leq \frac{c}{2} n^{\zeta - 1/2}$ as n is large enough. Therefore, for sufficiently large n, we have

$$
\begin{aligned}
& \mathrm{pr}(\max_{1 \leq l \leq L} |n^{-1} \widehat{T}_{UC}^{(G_l)} - n^{-1} T_{UC}^{(G_l)}| \geq c n^{\zeta - 1/2}) \\
\leq & \mathrm{pr}\left(\max_{1 \leq l \leq L} \max_{j \in G_l} \left| \int_{\tau_L}^{\tau_U} \left| \frac{\widehat{\beta}^{(j)}(\tau)}{\widehat{\sigma}_n^{(j)}(\tau)} \right|^2 d\tau - \int_{\tau_L}^{\tau_U} \left| \frac{\widetilde{\beta}^{(j)}(\tau)}{\sigma^{(j)}(\tau)} \right|^2 d\tau \right| \geq c n^{\zeta - 1/2} \right) \\
= & \mathrm{pr}\left(\max_{1 \leq j \leq p} \left| \int_{\tau_L}^{\tau_U} \left| \frac{\widehat{\beta}^{(j)}(\tau)}{\widehat{\sigma}_n^{(j)}(\tau)} \right|^2 d\tau - \int_{\tau_L}^{\tau_U} \left| \frac{\widetilde{\beta}^{(j)}(\tau)}{\sigma^{(j)}(\tau)} \right|^2 d\tau \right| \geq c n^{\zeta - 1/2} \right) \\
\leq & \mathrm{pr}(\max_{1 \leq j \leq p} (|A_{j,1}| + |A_{j,2}| + |A_{j,3}|) \geq \frac{c}{2} n^{\zeta - 1/2}) \\
\leq & \sum_{k=1}^{3} \mathrm{pr}(\max_{1 \leq j \leq p} (|A_{j,k}|) \geq \frac{c}{6} n^{\zeta - 1/2}).
\end{aligned}
$$

This suggests that in order to obtain the exponential tail bound for $\mathrm{pr}(\max_{1 \leq l \leq L} |n^{-1} \widehat{T}_{UC}^{(G_l)} - n^{-1} T_{UC}^{(G_l)}| \geq c n^{\zeta - 1/2})$, we need to derive the bound for $\mathrm{pr}(\max_{1 \leq j \leq p} (|A_{j,k}|) \geq \frac{c}{6} n^{\zeta - 1/2}), k = 1, 2, 3$.

We then work on $\mathrm{pr}(\max_{1 \leq j \leq p} (|A_{j,k}|) \geq \frac{c}{6} n^{\zeta - 1/2})$ for $k = 1, 2, 3$ applying Lemma A3.5.

The exponential tail bound for $\mathbf{pr(\max_{1\le j\le p}(|A_{j,1}|) \ge \frac{c}{6}n^{\zeta-1/2})}$

By definition, $A_{j,1} = \int_{\tau_L}^{\tau_U} \left|\frac{[\widehat{\beta}^{(j)}(\tau)-\widetilde{\beta}^{(j)}(\tau)]}{\widehat{\sigma}_n^{(j)}(\tau)}\right|^2 d\tau$. Due to the consistency of $\widehat{\beta}^{(j)}(\tau)$ and by Condition 3.4, we can derive that for sufficient large n, there exists positive constant π_1 and π_2, such that $|\widehat{\beta}^{(j)}(\tau) - \widetilde{\beta}^{(j)}(\tau)| \le 3\pi_1$ and $\widehat{\sigma}_n^{(j)}(\tau) \ge \pi_2/2$, then we have

$$
\begin{aligned}
& \mathrm{pr}(\max_{1\le j\le p}(|A_{j,1}|) \ge \frac{c}{4}n^{\zeta-1/2}) \\
= \;& \mathrm{pr}\left(\max_{1\le j\le p}\left(\int_{\tau_L}^{\tau_U}\left|\frac{[\widehat{\beta}^{(j)}(\tau) - \widetilde{\beta}^{(j)}(\tau)]}{\widehat{\sigma}_n^{(j)}(\tau)}\right|^2 d\tau\right) \ge \frac{c}{6}n^{\zeta-1/2}\right) \\
\le \;& \mathrm{pr}(\max_{1\le j\le p}\sup_{\tau\in[\tau_L,\tau_U]}\left|\frac{[\widehat{\beta}^{(j)}(\tau) - \widetilde{\beta}^{(j)}(\tau)]}{\widehat{\sigma}_n^{(j)}(\tau)}\right|^2 \ge \frac{c}{6(\tau_U-\tau_L)}n^{\zeta-1/2}) \\
\le \;& \mathrm{pr}(\max_{1\le j\le p}\sup_{\tau\in[\tau_L,\tau_U]}|\widehat{\beta}^{(j)}(\tau) - \widetilde{\beta}^{(j)}(\tau)| \ge \frac{c\pi_2^2}{72\pi_1(\tau_U-\tau_L)}n^{\zeta-1/2}) \\
\le \;& p\cdot\mathrm{pr}(\sup_{\tau\in[\tau_L,\tau_U]}|\widehat{\beta}^{(j)}(\tau) - \widetilde{\beta}^{(j)}(\tau)| \ge c^* n^{\zeta-1/2})
\end{aligned}
$$

with $c^* = \frac{c\pi_2^2}{72\pi_1(\tau_U-\tau_L)}$. Applying Lemma A3.5 (A), there exists positive constant c_3^* and η_3^*, such that

$$
\begin{aligned}
\mathrm{pr}(\max_{1\le j\le p}(|A_{j,1}|) \ge \frac{c}{6}n^{\zeta-1/2}) \;&\le\; p\cdot\mathrm{pr}(\sup_{\tau\in[\tau_L,\tau_U]}|\widehat{\beta}^{(j)}(\tau) - \widetilde{\beta}^{(j)}(\tau)| \ge c^* n^{\zeta-1/2}) \\
&\le\; pc_3^*\exp(-\eta_3^* n^{4\zeta-1} - \log(n^{\zeta-1/2})).
\end{aligned}
$$

The exponential tail bound for $\mathbf{pr(\max_{1\le j\le p}(|A_{j,2}|) \ge \frac{c}{6}n^{\zeta-1/2})}$

Similarly, for $A_{j,2} = 2\int_{\tau_L}^{\tau_U} \frac{\widetilde{\beta}^{(j)}(\tau)}{\sigma^{(j)}(\tau)} \frac{\widehat{\beta}^{(j)}(\tau)-\widetilde{\beta}^{(j)}(\tau)}{\widehat{\sigma}_n^{(j)}(\tau)} d\tau$, we can write

$$\begin{aligned}
& \mathrm{pr}(\max_{1\le j\le p}(|A_{j,2}|) \ge \frac{c}{6}n^{\zeta-1/2}) \\
= \;& \mathrm{pr}\left(\max_{1\le j\le p}\left(2\int_{\tau_L}^{\tau_U} \frac{\widetilde{\beta}^{(j)}(\tau)}{\sigma^{(j)}(\tau)} \frac{\widehat{\beta}^{(j)}(\tau)-\widetilde{\beta}^{(j)}(\tau)}{\widehat{\sigma}_n^{(j)}(\tau)} d\tau\right) \ge \frac{c}{6}n^{\zeta-1/2}\right) \\
\le \;& \mathrm{pr}(\max_{1\le j\le p} \sup_{\tau\in[\tau_L,\tau_U]} |\widehat{\beta}^{(j)}(\tau) - \widetilde{\beta}^{(j)}(\tau)| \ge \frac{c\pi_2^2}{24\pi_1(\tau_U-\tau_L)}n^{\zeta-1/2}) \\
\le \;& p\cdot \mathrm{pr}(\sup_{\tau\in[\tau_L,\tau_U]} |\widehat{\beta}^{(j)}(\tau) - \widetilde{\beta}^{(j)}(\tau)| \ge c^{**}n^{\zeta-1/2})
\end{aligned}$$

with $c^{**} = \frac{c\pi_2^2}{24\pi_1(\tau_U-\tau_L)}$. Applying Lemma A3.5, there exists positive constant c_3^{**} and η_3^{**}, such that

$$\begin{aligned}
\mathrm{pr}(\max_{1\le j\le p}(|A_{j,2}|) \ge \frac{c}{6}n^{\zeta-1/2}) \;&\le\; p\cdot \mathrm{pr}(\sup_{\tau\in[\tau_L,\tau_U]} |\widehat{\beta}^{(j)}(\tau) - \widetilde{\beta}^{(j)}(\tau)| \ge c^{**}n^{\zeta-1/2}) \\
&\le\; pc_3^{**}\exp(-\eta_3^{**}n^{4\zeta-1} - \log(n^{\zeta-1/2}))
\end{aligned}$$

for sufficient large n.

The exponential tail bound for $\mathbf{pr(\max_{1\le j\le p}(|A_{j,3}|) \ge \frac{c}{6}n^{\zeta-1/2})}$

Then, for $A_{j,3} = \int_{\tau_L}^{\tau_U} \left(\frac{\{\sigma^{(j)}(\tau)\}^2}{\{\widehat{\sigma}_n^{(j)}(\tau)\}^2} - 1\right) \left|\frac{\widetilde{\beta}^{(j)}(\tau)}{\sigma^{(j)}(\tau)}\right|^2 d\tau$. Due to the consistency of $\widehat{\beta}^{(j)}(\tau)$, there exists π_3 such that $|\widehat{\beta}^{(j)}(\tau) + \widetilde{\beta}^{(j)}(\tau)| \le 3\pi_3$ as n is sufficiently large. Then we can write

$$\begin{aligned}
& \mathrm{pr}(\max_{1\le j\le p}(|A_{j,3}|) \ge \frac{c}{6}n^{\zeta-1/2}) \\
= \;& \mathrm{pr}\left(\max_{1\le j\le p}\left(\int_{\tau_L}^{\tau_U}\left(\frac{\{\sigma^{(j)}(\tau)\}^2}{\{\widehat{\sigma}_n^{(j)}(\tau)\}^2} - 1\right)\left|\frac{\widetilde{\beta}^{(j)}(\tau)}{\sigma^{(j)}(\tau)}\right|^2 d\tau\right) \ge \frac{c}{6}n^{\zeta-1/2}\right) \\
\le \;& \mathrm{pr}(\max_{1\le j\le p} \sup_{\tau\in[\tau_L,\tau_U]} |\widehat{\sigma}_n^{(j)}(\tau) - \sigma^{(j)}(\tau)| \ge \frac{c\pi_2^3}{36\pi_1^2\pi_3(\tau_U-\tau_L)}n^{\zeta-1/2}) \\
\le \;& p\cdot \mathrm{pr}(\sup_{\tau\in[\tau_L,\tau_U]} |\widehat{\sigma}_n^{(j)}(\tau) - \sigma^{(j)}(\tau)| \ge c'n^{\zeta-1/2})
\end{aligned}$$

with $c' = \frac{c\pi_2^3}{36\pi_1^2\pi_3(\tau_U-\tau_L)}$. Applying Lemma A3.5, there exists positive constant c_4^* and

η_4^*, such that

$$\begin{aligned}\mathrm{pr}(\max_{1\le j\le p}(|A_{j,3}|)\ge\frac{c}{6}n^{\zeta-1/2}) &\le p\cdot\mathrm{pr}(\sup_{\tau\in[\tau_L,\tau_U]}|\widehat{\sigma}_n^{(j)}(\tau)-\sigma^{(j)}(\tau)|\ge c'n^{\zeta-1/2})\\ &\le pc_4^*\exp(-\eta_4^*n^{4\zeta-1}-\log(n^{\zeta-1/2}))\end{aligned}$$

for sufficient large n.

As a result, we can show that

$$\begin{aligned}&\mathrm{pr}(\max_{1\le l\le L}|n^{-1}\widehat{T}_{UC}^{(G_l)}-n^{-1}T_{UC}^{(G_l)}|\ge cn^{\zeta-1/2})\\ \le\ &\mathrm{pr}\left(\max_{1\le j\le p}\left|\int_{\tau_L}^{\tau_U}\left|\frac{\widehat{\beta}^{(j)}(\tau)}{\widehat{\sigma}_n^{(j)}(\tau)}\right|^2d\tau-\int_{\tau_L}^{\tau_U}\left|\frac{\widetilde{\beta}^{(j)}(\tau)}{\sigma^{(j)}(\tau)}\right|^2d\tau\right|\ge cn^{\zeta-1/2}\right)\\ \le\ &pc_3^*\exp(-\eta_3^*n^{4\zeta-1}-\log(n^{\zeta-1/2}))+pc_3^{**}\exp(-\eta_3^{**}n^{4\zeta-1}-\log(n^{\zeta-1/2}))\\ &+pc_4^*\exp(-\eta_4^*n^{4\zeta-1}-\log(n^{\zeta-1/2}))\\ \le\ &p\upsilon\exp(-\eta n^{4\zeta-1}-\log(n^{\zeta-1/2}))\end{aligned}$$

for sufficiently large n and some positive constant υ and η. This completes the Proof of Theorem 3.2. □

for Theorem 3. This can be shown follow the same arguments as in the Proof of Theorem 3.2 replacing $\widehat{T}_{UC}^{(G_l)}$ with $\widehat{T}_{C}^{(G_{c,l})}$, thus are omitted. □

3.7.2.3 Proofs for Corollary 3.1, Corollary 3.2 and Corollary 3.3

for Corollary 3.1. Given condition 3.5, for $\text{pr}(M_{[\tau_L,\tau_U]} \subseteq \widehat{M}_{[\tau_L,\tau_U]})$, we can write

$$\begin{aligned}
&\text{pr}(M_{[\tau_L,\tau_U]} \subseteq \widehat{M}_{[\tau_L,\tau_U]}) \geq \text{pr}(\min_{G_l \in M_{[\tau_L,\tau_U]}} n^{-1}\widehat{T}_{UC}^{(G_l)} \geq \nu_n) \\
\geq\ & \text{pr}(\min_{G_l \in M_{[\tau_L,\tau_U]}} n^{-1}T_{UC}^{(G_l)} - \max_{G_l \in M_{[\tau_L,\tau_U]}} |n^{-1}\widehat{T}_{UC}^{(G_l)} - n^{-1}T_{UC}^{(G_l)}| \geq \nu_n) \\
\geq\ & 1 - \text{pr}(\max_{G_l \in M_{[\tau_L,\tau_U]}} |n^{-1}\widehat{T}_{UC}^{(G_l)} - n^{-1}T_{UC}^{(G_l)}| \geq \min_{G_l \in M_{[\tau_L,\tau_U]}} n^{-1}T_{UC}^{(G_l)} - \nu_n) \\
\geq\ & 1 - \text{pr}(\max_{G_l \in M_{[\tau_L,\tau_U]}} |n^{-1}\widehat{T}_{UC}^{(G_l)} - n^{-1}T_{UC}^{(G_l)}| \geq \frac{\alpha_0}{2}n^{\zeta-1/2}).
\end{aligned}$$

Applying Theorem 3.2, there exists positive constant a_1 and b_1 such that

$$\text{pr}(\max_{G_l \in M_{[\tau_L,\tau_U]}} |n^{-1}\widehat{T}_{UC}^{(G_l)} - n^{-1}T_{UC}^{(G_l)}| \geq \frac{\alpha_0}{2}n^{\zeta-1/2}) \leq S_{[\tau_L,\tau_U]} \cdot a_1 \exp(-b_1 n^{4\zeta-1} - \log(n^{\zeta-1/2})).$$

Under the assumption that $S_{[\tau_L,\tau_U]} = o(n)$, we have $S_{[\tau_L,\tau_U]} \cdot a_1 \exp(-b_1 n^{4\zeta-1} - \log(n^{\zeta-1/2}))$ converges to zero as n goes to infinity. This completes the Proof of the sure screening property described in Corollary 3.1. □

for Corollary 3.2. Define the set of relevant covariates in two steps as

$$M_{[\tau_L,\tau_U]}^{G_1} = \{G_l :\ 1 \leq l \leq L \text{ and there exists } \tau \in [\tau_L, \tau_U] \text{ such that } Q_\tau(Y|\boldsymbol{X}) \text{ depends on } X_{G_l}\}$$

$$M_{[\tau_L,\tau_U]}^{G_2} = \{\{r_m\} :\ 1 \leq m \leq M \text{ and there exists } \tau \in [\tau_L, \tau_U] \text{ such that } Q_\tau(Y|\boldsymbol{X}) \text{ depends on } X_{\{r_m\}}\}.$$

Remarkably, under the proposed screening procedure, we have $M_{[\tau_L,\tau_U]}^{G_2} = M_{[\tau_L,\tau_U]}$ and $|M_{[\tau_L,\tau_U]}^{G_2}| = S_{\tau_L,\tau_U}$. Also, it is easy to see that $\{G_l :\ 1 \leq l \leq L, G_l \in M_{[\tau_L,\tau_U]}^{G_1}\} = \{G_l : 1 \leq l \leq L, G_l \cap M_{[\tau_L,\tau_U]} \neq \varnothing\}$, which suggests $|M_{[\tau_L,\tau_U]}^{G_1}| \leq S_{\tau_L,\tau_U}$. Denote $\widehat{M}_{[\tau_L,\tau_U]}^{G_1} = \{G_l : 1 \leq l \leq L,\ w_{1,l} \geq \nu_{n,1}\}$ and $\widehat{M}_{[\tau_L,\tau_U]}^{G_2} = \{r_m : 1 \leq m \leq M,\ w_{2,m} \geq \nu_{n,2}\}$. Then we can write

$$\text{pr}(M_{[\tau_L,\tau_U]} \subseteq \widehat{M}_{[\tau_L,\tau_U]}^{G}) = \text{pr}(M_{[\tau_L,\tau_U]}^{G_1} \subseteq \widehat{M}_{[\tau_L,\tau_U]}^{G_1})\text{pr}(M_{[\tau_L,\tau_U]}^{G_2} \subseteq \widehat{M}_{[\tau_L,\tau_U]}^{G_2})$$

Using similar lines as in the Proof of Corollary 3.1, under condition 3.6, we have

$$\begin{aligned}\mathrm{pr}(M^{G_1}_{[\tau_L,\tau_U]} \subseteq \widehat{M}^{G_1}_{[\tau_L,\tau_U]}) &\geq \mathrm{pr}(\min_{G_l \in M^{G_1}_{[\tau_L,\tau_U]}} n^{-1}\widehat{T}^{(G_l)}_{UC} \geq \nu_{n,1}) = \mathrm{pr}(\min_{G_l \cap M_{[\tau_L,\tau_U]} \neq \varnothing} n^{-1}\widehat{T}^{(G_l)}_{UC} \geq \nu_{n,1}) \\ &\geq 1 - \mathrm{pr}(\max_{G_l \cap M_{[\tau_L,\tau_U]} \neq \varnothing} |n^{-1}\widehat{T}^{(G_l)}_{UC} - n^{-1}T^{(G_l)}_{UC}| \geq \frac{\alpha_0}{2} n^{\zeta-1/2})\end{aligned}$$

and

$$\begin{aligned}\mathrm{pr}(M^{G_2}_{[\tau_L,\tau_U]} \subseteq \widehat{M}^{G_2}_{[\tau_L,\tau_U]}) &\geq \mathrm{pr}(\min_{r_m \in M_{[\tau_L,\tau_U]}} n^{-1}\widehat{T}^{(\{r_m\})}_{UC} \geq \nu_{n,2}) \\ &\geq 1 - \mathrm{pr}(\max_{r_m \in M_{[\tau_L,\tau_U]}} |n^{-1}\widehat{T}^{(\{r_m\})}_{UC} - n^{-1}T^{(\{r_m\})}_{UC}| \geq \frac{\alpha_0}{2} n^{\zeta-1/2}).\end{aligned}$$

Applying Theorem 3.2, we can conclude that there exist positive constant a_2^* and b_2^* such that

$$\begin{aligned}&\mathrm{pr}(\max_{G_l \cap M_{[\tau_L,\tau_U]} \neq \varnothing} |n^{-1}\widehat{T}^{(G_l)}_{UC} - n^{-1}T^{(G_l)}_{UC}| \geq \frac{\alpha_0}{2} n^{\zeta-1/2}) \\ \leq\ & |M^{G_1}_{[\tau_L,\tau_U]}| \cdot a_2^* \exp(-b_2^* n^{4\zeta-1} - \log(n^{\zeta-1/2})) \leq S_{\tau_L,\tau_U} \cdot a_2^* \exp(-b_2^* n^{4\zeta-1} - \log(n^{\zeta-1/2}));\end{aligned}$$

and exist positive constant a_2^{**} and b_2^{**} such that

$$\begin{aligned}&\mathrm{pr}(\max_{r_m \in M_{[\tau_L,\tau_U]}} |n^{-1}\widehat{T}^{(\{r_m\})}_{UC} - n^{-1}T^{(\{r_m\})}_{UC}| \geq \frac{\alpha_0}{2} n^{\zeta-1/2}) \\ \leq\ & |M^{G_2}_{[\tau_L,\tau_U]}| \cdot a_2^{**} \exp(-b_2^{**} n^{4\zeta-1} - \log(n^{\zeta-1/2})) = S_{\tau_L,\tau_U} \cdot a_2^{**} \exp(-b_2^{**} n^{4\zeta-1} - \log(n^{\zeta-1/2})).\end{aligned}$$

These lead to

$$\begin{aligned}&\mathrm{pr}(M_{[\tau_L,\tau_U]} \subseteq \widehat{M}^{G}_{[\tau_L,\tau_U]}) \\ \geq\ & \{1 - S_{\tau_L,\tau_U} \cdot a_2^* \exp(-b_2^* n^{4\zeta-1} - \log(n^{\zeta-1/2}))\}\{1 - S_{\tau_L,\tau_U} \cdot a_2^{**} \exp(-b_2^{**} n^{4\zeta-1} - \log(n^{\zeta-1/2}))\} \\ \geq\ & 1 - S_{\tau_L,\tau_U} \cdot a_2 \exp(-b_2 n^{4\zeta-1} - \log(n^{\zeta-1/2}))\end{aligned}$$

with $a_2 = a_2^* + a_2^{**}$ and $b_2 = \min\{b_2^*, b_2^{**}\}$. As $S_{[\tau_L,\tau_U]} = o(n)$, we can show that

$\text{pr}(M_{[\tau_L,\tau_U]} \subseteq \widehat{M}^G_{[\tau_L,\tau_U]})$ converges to one as n goes to infinity. This completes the Proof of Corollary 3.2. □

for Corollary 3.3. For $\text{pr}(M^{(\mathcal{C})}_{[\tau_L,\tau_U]} \subseteq \widehat{M}^{(\mathcal{C})}_{[\tau_L,\tau_U]})$, using similar lines as in the Proof of Corollary 3.1 and given condition 3.7, we have

$$\begin{aligned}\text{pr}(M^{(\mathcal{C})}_{[\tau_L,\tau_U]} \subseteq \widehat{M}^{(\mathcal{C})}_{[\tau_L,\tau_U]}) &\geq \text{pr}(\min_{G_{c,l}\in M^{(\mathcal{C})}_{[\tau_L,\tau_U]}} n^{-1}\widehat{T}_C^{(G_{c,l})} \geq \nu_{c,n}) \\ &\geq 1 - \text{pr}(\max_{G_{c,l}\in M^{(\mathcal{C})}_{[\tau_L,\tau_U]}} |n^{-1}\widehat{T}_C^{(G_{c,l})} - n^{-1}T_C^{(G_{c,l})}| \geq \frac{\alpha_0}{2}n^{\zeta-1/2}).\end{aligned}$$

Applying Theorem 3, there exists positive constant a_3 and b_3 such that

$$\text{pr}(\max_{G_{c,l}\in M^{(\mathcal{C})}_{[\tau_L,\tau_U]}} |n^{-1}\widehat{T}_C^{(G_{c,l})} - n^{-1}T_C^{(G_{c,l})}| \geq \frac{\alpha_0}{2}n^{\zeta-1/2}) \leq S^{(\mathcal{C})}_{[\tau_L,\tau_U]} \cdot a_3 \exp(-b_3 n^{4\zeta-1} - \log(n^{\zeta-1/2})).$$

Assume $S^{(\mathcal{C})}_{[\tau_L,\tau_U]} = o(n)$, we have $S^{(\mathcal{C})}_{[\tau_L,\tau_U]} \cdot a_3 \exp(-b_3 n^{4\zeta-1} - \log(n^{\zeta-1/2}))$ converge to zero as n goes to infinity. This completes the Proof of for Corollary 3.3. □

3.7.3 Some Computational Considerations

To calculate $\{\widehat{T}^{(j)}_{inte}, j = 1, \ldots, J\}$, we adopt the grid method as in Cui and Peng (2022). Specifically, we approximate $\widehat{R}^{(j)}(\tau) = \widehat{\beta}^{(j)}(\tau)/\widehat{\sigma}^{(j)}_n(\tau)$ as a piecewise-constant function on $\mathcal{G} \equiv \{\tau_l\}_{l=0}^{N^*}$, a finely determined grid with $\tau_0 = \tau_L, \tau_{N^*} = \tau_U$ and grid size $\max_{1\leq l\leq N^*}\{\tau_l - \tau_{l-1}\} = o(n^{-1/2})$, and then calculate

$$\widehat{T}^{(j)}_{inte} = n\sum_{l=1}^{N^*} \widehat{R}^{(j)2}(\tau_l)(\tau_l - \tau_{l-1}).$$

We also extend the empirical strategy in Cui and Peng (2022) to determine the choice of u for the estimation of $\boldsymbol{V}_n(\tau)$. The following describe the algorithm developed to select u among a set of candidate values, $\mathcal{U} = \{u_1, \ldots, u_U\}$, for multivariate setting:

(S1.a) For each $u \in \mathcal{U}$, calculate $\widehat{R}^{(j)}(\tau;u) \equiv \widehat{\beta}^{(j)}(\tau)/\widehat{\sigma}_n^{(j)}(\tau;u), j = 1,\ldots,J$ for $\tau \in \mathcal{G}$, where $\widehat{\sigma}_n^{(j)}(\tau;u)$ denotes the $\widehat{\sigma}_n^{(j)}(\tau)$ computed with adjust constant u.

(S1.b) Calculate $\widehat{R}^{\star}(u) = \max_{\tau\in\mathcal{G}} \max_{1\le j\le J} \widehat{R}^{(j)}(\tau;u) - \text{median}_{\tau\in\mathcal{G}} \max_{1\le j\le J} \widehat{R}^{(j)}(\tau;u)$ for each $u \in \mathcal{U}$.

(S1.c) Calculate $\widehat{R}^{\dagger}(u) = \max_{\tau\in\mathcal{G}} \max\{\boldsymbol{V}_n(\tau;u)\} - \min_{\tau\in\mathcal{G}} \min\{\boldsymbol{V}_n(\tau;u)\}$, where $\boldsymbol{V}_n(\tau;u)$ is $\boldsymbol{V}_n(\tau)$ computed with adjust constant u and $\max\{\boldsymbol{V}_n(\tau;u)\}$ (or $\min\{\boldsymbol{V}_n(\tau;u)\}$) denotes the largest (or smallest) component of the matrix $\boldsymbol{V}_n(\tau;u)$.

(S1.d) Assign a large positive value to $A^{[0]}$ and $B^{[0]}$, say 10^5. Set $k = 1$ and $u^{[0]} = \max_{1\le k\le U} u_k + 1$.

(i) If $\widehat{R}^{\star}(k) < A^{[k-1]}$ and $\widehat{R}^{\dagger}(k) < B^{[k-1]}$, then let $A^{[k]} = \widehat{R}^{\star}(k)$, $B^{[k]} = \widehat{R}^{\dagger}(k)$, and $u^{[k]} = k$. Otherwise, let $A^{[k]} = A^{[k-1]}$, $B^{[k]} = B^{[k-1]}$, and $u^{[k]} = u^{[k-1]}$.

(ii) Increase k by 1 and go back to (i) until $k > U$.

(S1.e) If $|u^{[U]}| < |u_U| + 1$, then choose u as $u^{[U]}$. Otherwise, no appropriate u can be selected from $\mathcal{U}$.

Chapter 4

Non-parametric Testing for Survival Data With Time-dependent Covariates

4.1 The Proposed Testing Framework

4.1.1 Formulation of the Testing Problem

In this part, we focus on the problem to evaluate the association between a univariate time dependent covariate (continuous, discrete, or a mixture of both) and the time-to-event outcome. Let T be a continuous random variable representing the survival time. Denote $\bar{Z} = \{Z(t), 0 \leq t < \infty\}$ as a time dependent covariate process, where $Z(t)$ is a univariate covariate at $t \geq 0$. Let C denote the censoring time. Define $X = \min(T, C)$ and $\delta = I(T \leq C)$ as the observed survival time and the event status, respectively.

To deal with the time dependent covariates, we adopt the idea to assessing the overall effects over a pre-determined set of landmark times. Specifically, at each landmark time t_0, we assess the effect on the event-free individuals, and then summarize across all landmark times. Let $M(t_0) = I(X > t_0)$ denote the indicator on whether an individual survive up to time t_0. Let $T^r(t_0)$ be the residual survival time, and $V(t_0)$ be a pre-specified functional form of $\bar{Z}(t_0) = \{Z(t), 0 \leq t \leq t_0\}$, the covariate process up to t_0. Define $C^r(t_0) = C - t_0$ as the residual censoring time. The observed residual survival time is then defined as $X^r(t_0) = \min(T^r(t_0), C^r(t_0))$. We assume $(T^r(t_0), V(t_0)) \perp C \mid M(t_0) = 1$ for any $t_0 \in \mathcal{T}$. The observed data at time t_0 is then noted as $\{X_i^r(t_0), \delta_i, V_i(t_0)\}_{\{i; M_i(t_0)=1, i=1\ldots,n\}}$.

Denote $\mathcal{T}$ as a finite set of the landmark time points of interest, and $\Delta = [\tau_L, \tau_U]$ as the interval of quantile levels of interest for $T^r(t_0)$. Based on the problem of interest, we may formulate it as a hypothesis testing problem to test for

$$H_{0,\mathcal{T}} : Q_{T^r(t_0)|V(t_0),M(t_0)=1}(\tau_1) = Q_{T^r(t_0)|M(t_0)=1}(\tau_1) \text{ for } \tau_1 \in \Delta, t_0 \in \mathcal{T} \tag{4.1}$$

versus its alternative. Here $Q_{T^r(t_0)|M(t_0)=1}(\tau_1)$ denotes the τ_1th quantile of $T^r(t_0)$

conditional on $M(t_0) = 1$, and $Q_{T^r(t_0)|V(t_0),M(t_0)=1}(\tau_1)$ denotes the τ_1th quantile of $T^r(t_0)$ conditional on $V(t_0)$ and $M(t_0) = 1$.

4.1.2 The Proposed Test Statistics

Consider the simplest case with $\tau_L = \tau_U = \tau_1$ and $\mathcal{T} = \{t_1\}$, i.e. the target is to test whether $Q_{T^r(t_1)|V(t_1),M(t_1)=1}(\tau_1) = Q_{T^r(t_1)|M(t_1)=1}(\tau_1)$. Since $T^r(t_1)$ is continuous, we have the uniqueness of $Q_{T^r(t_1)|V(t_1)=v,M(t_1)=1}(\tau_1)$ for any v in the support of $V(t_1)$, noted as $\Omega_{V(t_1)}$. Following simple algebra, we can show that

$$\begin{aligned} & Q_{T^r(t_1)|V(t_1),M(t_1)=1}(\tau_1) = Q_{T^r(t_1)|M(t_1)=1}(\tau_1) \\ \Leftrightarrow \quad & \int_{\Omega_{V(t_1)}} \frac{c^2(\tau_1, v; t_1)}{\tau_1(1-\tau_1)F_2(v;t_1)(1-F_2(v;t_1))} d\mu_2(v) = 0, \end{aligned} \tag{4.2}$$

where $c(\tau_1, v; \cdot) = Cov\{I(F_1(T^r(\cdot); \cdot) \le \tau_1), I(V(\cdot) \le v) | M(\cdot) = 1\}$, μ_2 is a measure, and $F_1(s; \cdot) = P(T^r(\cdot) \le s | M(\cdot) = 1)$ and $F_2(v; \cdot) = P(V(\cdot) \le v | M(\cdot) = 1)$ represent the cumulative distribution function of $T^r(\cdot)$ and $V(\cdot)$, respectively, conditional on $M(\cdot) = 1$.

Motivated from (4.2), we define two indices to test for $H_{0,\mathcal{T}}$ in (4.1):

$$\begin{aligned} q_{max}(T, \bar{Z}; \Delta, \mathcal{T}) &= \max_{t_0 \in \mathcal{T}} \int_\Delta \int_{\Omega_{V(t_0)}} \frac{c^2(\tau_1, v; t_0)}{\tau_1(1-\tau_1)F_2(v;t_0)(1-F_2(v;t_0))} d\mu_1(\tau_1) d\mu_2(v); \\ q_{sum}(T, \bar{Z}; \Delta, \mathcal{T}) &= \sum_{t_0 \in \mathcal{T}} \int_\Delta \int_{\Omega_{V(t_0)}} \frac{c^2(\tau_1, v; t_0)}{\tau_1(1-\tau_1)F_2(v;t_0)(1-F_2(v;t_0))} d\mu_1(\tau_1) d\mu_2(v). \end{aligned}$$

From the definition, we can learn that $q_{max}(T, \bar{Z}; \Delta, \mathcal{T})$ measures the maximal covariate effect across $\mathcal{T}$ and $q_{sum}(T, \bar{Z}; \Delta, \mathcal{T})$ measures the total covariate effect over $\mathcal{T}$. It is remarkable that with time independent covariate, i.e. $Z(t) = Z(t_0)$ for any t, these indices reduce to

$$q(T, Z(t_0); \Delta) = \int_\Delta \int_{\Omega_{V(t_0)}} \frac{c^2(\tau_1, v; t_0)}{\tau_1(1-\tau_1)F_2(v;t_0)(1-F_2(v;t_0))} d\mu_1(\tau_1) d\mu_2(v).$$

When we have continuous covariate, we can show that $q(T, Z(t_0); \Delta)$ is equivalent to the index suggested in Zhu et al. (2018). This is discussed with more details later in Section 4.1.4.

To estimate the indices, the key insight is to obtain the estimator for $F_1(s; t_0)$, $F_2(v; t_0)$ and $c(\tau_1, v; t_0)$, which can be written as

$$\begin{aligned} F_1(s;t_0) &= \frac{P(T^r(t_0) \le s, M(t_0)=1)}{\pi_{t_0}}, \quad F_2(v;t_0) = \frac{P(V(t_0) \le v, M(t_0)=1)}{\pi_{t_0}}, \\ c(\tau_1, v; t_0) &= \frac{P(F_1(T^r(t_0);t_0) \le \tau_1, V(t_0) \le v, M(t_0)=1)}{\pi_{t_0}} - \\ & \quad \frac{P(F_1(T^r(t_0);t_0) \le \tau_1, M(t_0)=1)}{\pi_{t_0}} \cdot \frac{P(V(t_0) \le v, M(t_0)=1)}{\pi_{t_0}} \end{aligned}$$

where $\pi_{t_0} = P(M(t_0) = 1)$. When there is no censoring, they can be straight forwardly estimated based on the empirical distribution:

$$\begin{aligned} F_{n,1}(s;t_0) &= \frac{1}{n}\sum_{i=1}^{n} \frac{I(T_i^r(t_0) \le s, M_i(t_0)=1)}{\widehat{\pi}_{t_0}}, \\ F_{n,2}(v;t_0) &= \frac{1}{n}\sum_{i=1}^{n} \frac{I(V_i(t_0) \le v, M_i(t_0)=1)}{\widehat{\pi}_{t_0}}, \\ \widehat{c}(\tau_1, v; t_0) &= \frac{1}{n}\sum_{i=1}^{n} \left\{ \frac{I(F_{n,1}(T_i^r(t_0);t_0) \le \tau_1, V_i(t_0) \le v, M_i(t_0)=1)}{\widehat{\pi}_{t_0}} - \right. \\ & \quad \left. \frac{I(F_{n,1}(T_i^r(t_0);t_0) \le \tau_1, M_i(t_0)=1)}{\widehat{\pi}_{t_0}} \cdot \frac{I(V_i(t_0) \le v, M_i(t_0)=1)}{\widehat{\pi}_{t_0}} \right\} \end{aligned}$$

where $\widehat{\pi}_{t_0} = \frac{1}{n}\sum_{i=1}^{n} I(M_i(t_0) = 1)$.

Under the non-informative censoring assumption, to deal with censoring, we consider to conduct estimation following the idea discussed in Lin et al. (1993). Specifically, we utilize the fact that

$$\begin{aligned} P(T^r(t_0) > q_{1,t_0}(\tau_1) | M(t_0)=1) &= \frac{P(X^r(t_0) > q_{1,t_0}(\tau_1), M(t_0)=1)}{\pi_{t_0} G^r_{C,t_0}(q_{1,t_0}(\tau_1))} \\ P(T^r(t_0) > q_{1,t_0}(\tau_1), V(t_0) \le v | M(t_0)=1) &= \frac{P(X^r(t_0) > q_{1,t_0}(\tau_1), V(t_0) \le v, M(t_0)=1)}{\pi_{t_0} G^r_{C,t_0}(q_{1,t_0}(\tau_1))}, \end{aligned}$$

where $q_{1,t_0}(\tau_1) = Q_{T^r(t_0)|M(t_0)=1}(\tau_1)$, $\pi_{t_0} = P(M(t_0) = 1)$ and $G^r_{C,t_0}(\cdot) = P(C^r(t_0) > \cdot | M(t_0) = 1)$ is the survival distribution of the residual censoring time C^r conditional on $M_i(t_0) = 1$. These suggest the estimators

$$\begin{aligned}
F_{n,1}(\widehat{q}_{1,t_0}(\tau_1)) &= 1 - \frac{1}{n}\sum_{i=1}^{n} \frac{I(X_i^r(t_0) > \widehat{q}_{1,t_0}(\tau_1), M_i(t_0) = 1)}{\widehat{G}^r_{C,t_0}(\widehat{q}_{1,t_0}(\tau_1))\widehat{\pi}_{t_0}}, \\
F_{n,2}(v) &= \frac{1}{n}\sum_{i=1}^{n} \frac{I(V_i(t_0) \le v, M_i(t_0) = 1)}{\widehat{\pi}_{t_0}}, \\
F_{n,12}(\widehat{q}_{1,t_0}(\tau_1), v) &= F_{n,2}(v) - \frac{1}{n}\sum_{i=1}^{n} \frac{I(X_i^r(t_0) > \widehat{q}_{1,t_0}(\tau_1), V_i(t_0) \le v, M_i(t_0) = 1)}{\widehat{G}^r_{C,t_0}(\widehat{q}_{1,t_0}(\tau_1))\widehat{\pi}_{t_0}}, \\
\widehat{c}(\tau_1, v; t_0) &= F_{n,12}(\widehat{q}_{1,t_0}(\tau_1), v) - F_{n,1}(\widehat{q}_{1,t_0}(\tau_1)) \cdot F_{n,2}(v)
\end{aligned}$$

where $\widehat{q}_{1,t_0}(\tau_1) = \widehat{Q}_{T^r(t_0)|M(t_0)=1}(\tau_1)$, $\widehat{\pi}_{t_0} = \frac{1}{n}\sum_{i=1}^{n} I(M_i(t_0) = 1)$ and $\widehat{G}^r_{C,t_0}(\cdot)$ is the Kaplan-Meier estimator for the survival distribution of the conditional residual censoring time.

Let $\mathcal{A}_2(t_0) = \{a_1 = -\infty, a_2, \ldots, a_{n_2(t_0)}\}$ denote a set with $a_2, \ldots, a_{n_2(t_0)}$ as the unique values within the set $\{V_i(t_0)\}_{\{i;M_i=1,i=1\ldots,n\}}$ in an increasing order with $n_2(t_0) \le n$. Motivating from the fact that the empirical distributions are step functions, the general estimators for the proposed test statistics are defined as

$$\begin{aligned}
&\widehat{q}_{max}(T, \bar{Z}; \Delta, \mathcal{T}) = \max_{t_0 \in \mathcal{T}} \widehat{q}(T^r(t_0), V(t_0); \Delta) \\
= &\max_{t_0 \in \mathcal{T}} \sum_{j_1=1}^{n} \sum_{j_2=1}^{n_2(t_0)-1} \int_{\Delta \cap [\frac{j_1-1}{n}, \frac{j_1}{n})} \int_{[a_{j_2}, a_{j_2+1})} \frac{\widehat{c}^2(\tau_1, v; t_0)}{\tau_1(1-\tau_1)F_{n,2}(v)(1-F_{n,2}(v))} d\mu_1(\tau_1) d\mu_2(v) \\
= &\max_{t_0 \in \mathcal{T}} \sum_{j_1=1}^{n} \sum_{j_2=1}^{n_2(t_0)-1} \widehat{c}^2(\frac{j_1}{n}, a_{j_2}; t_0) \int_{\Delta \cap [\frac{j_1-1}{n}, \frac{j_1}{n})} \frac{d\mu_1(\tau_1)}{\tau_1(1-\tau_1)} \int_{[a_{j_2}, a_{j_2+1})} \frac{d\mu_2(v)}{F_{n,2}(v)(1-F_{n,2}(v))},
\end{aligned}$$

$$\widehat{q}_{sum}(T,\bar{Z};\Delta,\mathcal{T}) = \sum_{t_0\in\mathcal{T}} \widehat{q}(T^r(t_0),V(t_0);\Delta)$$

$$= \sum_{t_0\in\mathcal{T}}\sum_{j_1=1}^{n}\sum_{j_2=1}^{n_2(t_0)-1} \int_{\Delta\cap[\frac{j_1-1}{n},\frac{j_1}{n})}\int_{[a_{j_2},a_{j_2+1})} \frac{\widehat{c}^2(\tau_1,v;t_0)}{\tau_1(1-\tau_1)F_{n,2}(v)(1-F_{n,2}(v))}d\mu_1(\tau_1)d\mu_2(v)$$

$$= \sum_{t_0\in\mathcal{T}}\sum_{j_1=1}^{n}\sum_{j_2=1}^{n_2(t_0)-1} \widehat{c}^2(\frac{j_1}{n},a_{j_2};t_0)\int_{\Delta\cap[\frac{j_1-1}{n},\frac{j_1}{n})}\frac{d\mu_1(\tau_1)}{\tau_1(1-\tau_1)}\int_{[a_{j_2},a_{j_2+1})}\frac{d\mu_2(v)}{F_{n,2}(v)(1-F_{n,2}(v))}.$$

Notably, in the circumstance of interest, we set μ_1 as a Lebesgue measure (i.e $\mu_1(\tau_1) = \tau_1$), and μ_2 as a counting measure (i.e $\mu_2(v) = I(v \geq a_{j_2})$ for $v \in [a_{j_2}, a_{j_2+1}), j_2 = 1,\ldots,n_2(t_0)-1$). This leads to

$$\int_{[a,b)}\frac{1}{\tau_1(1-\tau_1)}d\mu_1(\tau_1) = \{\log(b)-\log(1-b)\} - \{\log(a)-\log(1-a)\} \text{ for any } a,b,$$
$$\int_{[a_{j_2},a_{j_2+1})}\frac{1}{F_{n,2}(v)(1-F_{n,2}(v))}d\mu_2(v) = \frac{1}{F_{n,2}(a_{j_2})(1-F_{n,2}(a_{j_2}))}.$$

In Theorem 4.1, we establish the limit null distribution of the proposed test statistics.

Theorem 4.1. *Given that Conditions 4.1 and 4.2 holds. Assume the density function f_1 and its first derivation with respect to τ are bounded away from zero and infinity on Δ. Under null hypothesis $H_{0,\mathcal{T}}$, we have*

$$n\widehat{q}_{max}(T,\bar{Z};\Delta,\mathcal{T}) \to_d \max_{t_0\in\mathcal{T}}\int_\Delta\int_{\Omega_{V(t_0)}}\chi^2(\tau_1,v;t_0)d\mu_1(\tau_1)d\mu_2(v),$$
$$n\widehat{q}_{sum}(T,\bar{Z};\Delta,\mathcal{T}) \to_d \sum_{t_0\in\mathcal{T}}\int_\Delta\int_{\Omega_{V(t_0)}}\chi^2(\tau_1,v;t_0)d\mu_1(\tau_1)d\mu_2(v),$$

where $\chi(\tau_1,v;t_0)$ is as defined in the Appendix Section 4.5.2.

We also investigate the asymptotic behavior of the proposed test statistics under a general class of alternative hypotheses. The findings are stated in Theorem 4.2.

Theorem 4.2. *Given that Conditions 4.1 and 4.2 holds. Assume the density function f_1 and its first derivation with respect to τ are bounded away from zero and infinity*

on Δ.

(A) $\widehat{q}_{max}(T, \bar{Z}; \Delta, \mathcal{T})$ *is consistent against the alternative hypothesis*

$$H_{a,max}: \quad q_{max}(T, \bar{Z}; \Delta, \mathcal{T}) > 0.$$

(B) $\widehat{q}_{sum}(T, \bar{Z}; \Delta, \mathcal{T})$ *is consistent against the alternative hypothesis*

$$H_{a,sum}: \quad q_{sum}(T, \bar{Z}; \Delta, \mathcal{T}) > 0.$$

The results of Theorem 4.2 indicate that the test statistics have power approaching to 1 (as n goes to ∞) under alternative cases subject to very mild constraints. Given the smoothness of $q_{1,t_0}(\tau_1)$, a general scenario that ensures the consistency of both statistics can be described as

> $\widetilde{H}_a$: There exists an interval $[\tau_1, \tau_2] \subseteq [\tau_L, \tau_U]$ for $v \in \Omega_{V(t_0)}$ and $t_k \in \mathcal{T}$ such that $|c(\tau, v; t_k)| > 0$ for $\tau \in [\tau_1, \tau_2]$.

This suggests that the proposed tests are powerful even when it only influences a segment of the outcome distribution for certain v and t_0. This feature is conceptually appealing for handling a dynamic covariate effect. The detailed proofs for Theorems 4.1 and 4.2 can be found in Appendix Sections 4.5.2 and 4.5.3.

4.1.3 The Proposed Resampling Procedure

Under $H_{0,\mathcal{T}}$, as presented in the Proof of Theorem 4.1, we can show that given t_0,

$$\begin{aligned} \frac{\widehat{c}(\tau_1, v; t_0)}{\sqrt{F_{n,2}(v)(1 - F_{n,2}(v))}} &= \frac{F_{n,12}(\widehat{q}_{1,t_0}(\tau_1), v) - F_{n,1}(\widehat{q}_{1,t_0}(\tau_1)) \cdot F_{n,2}(v)}{\sqrt{F_{n,2}(v)(1 - F_{n,2}(v))}} \\ &= -\frac{1}{n} \sum_{i=1}^{n} \xi_i(\tau_1, v; t_0) + o_p(n^{-1/2}) \end{aligned}$$

holds uniformly for any (τ_1, v), where

$$\begin{aligned}\xi_i(\tau_1, v; t_0) &= \frac{I(M_i(t_0)=1)}{\pi_{t_0}\sqrt{F_2(v)(1-F_2(v))}} \times \\ &\{I[X_i^r(t_0) > q_{1,t_0}(\tau_1)][I(V_i(t_0)\le v) - F_2(v)]/G^r_{C,t_0}(q_{1,t_0}(\tau_1)) \\ &-(1-\tau_1)[I(V_i(t_0)\le v) - F_2(v)]\}.\end{aligned}$$

This motivates us to consider the following resampling procedure:

(a) Generate B independent sets of $\{\iota_i^b\}_{i=1}^n$, where $\{\iota_i^b\}_{i=1}^n$ are independent random variables from a standard normal distribution and $b = 1, 2, \ldots, B$.

(b) Compute the estimates for the influence function as

$$\begin{aligned}\widehat{\xi}_i(\tau_1, v; t_0) &= \frac{I(M_i(t_0)=1)}{\widehat{\pi}_{t_0}\sqrt{F_{n,2}(v)(1-F_{n,2}(v))}} \times \\ &\{[I(X_i^r(t_0) > \widehat{q}_{1,t_0}(\tau_1), V_i(t_0)\le v) - F_{n,2}(v) I(X_i^r(t_0) > \widehat{q}_{1,t_0}(\tau_1))]/ \\ &G^r_{C,t_0}[\widehat{q}_{1,t_0}(\tau_1)] - (1-\tau_1) I[V_i(t_0)\le v] + (1-\tau_1)F_{n,2}(v)\}.\end{aligned}$$

(c) For $b = 1, 2, \ldots, B$, calculate

$$\begin{aligned}n\widehat{q}_{max,b}(T^r, \bar{Z}; \Delta, \mathcal{T}) &= \max_{t_0\in\mathcal{T}} \sum_{j_1=1}^{n} \sum_{j_2=1}^{n_2(t_0)-1} \{n^{-1/2}\sum_{l=1}^{n} \widehat{\xi}_i(\frac{j_1}{n}, a_{j_2}; t_0)\iota_i^b\}^2 \\ &\times \int_{\Delta\cap[\frac{j_1-1}{n}, \frac{j_1}{n})} \frac{1}{\tau_1(1-\tau_1)} d\mu_1(\tau_1), \\ n\widehat{q}_{sum,b}(T^r, \bar{Z}; \Delta, \mathcal{T}) &= \sum_{t_0\in\mathcal{T}} \sum_{j_1=1}^{n} \sum_{j_2=1}^{n_2(t_0)-1} \{n^{-1/2}\sum_{l=1}^{n} \widehat{\xi}_i(\frac{j_1}{n}, a_{j_2}; t_0)\iota_i^b\}^2 \\ &\times \int_{\Delta\cap[\frac{j_1-1}{n}, \frac{j_1}{n})} \frac{1}{\tau_1(1-\tau_1)} d\mu_1(\tau_1).\end{aligned}$$

(d) The p value is calculated as

$$p_\cdot = \frac{1}{B}\sum_{b=1}^{B} I[n\widehat{q}_{\cdot,b}(T^r, \bar{Z}; \Delta, \mathcal{T}) > n\widehat{q}_\cdot(T^r, \bar{Z}; \Delta, \mathcal{T})].$$

The resampling procedure presented above is subject to easy implementation based on moment estimators. Similarly ideas were used by other authors, for example, Lin et al. (1993), Li and Peng (2014) and Cui and Peng (2022). The rigorous theoretical justification for the presented resampling procedure is provided in Section 4.5.4.

4.1.4 Connection with Zhu et al. (2018)'s Index

Remarkably, in the case when we have continuous time independent covariate without ties, the proposed test statistics reduce to

$$\widehat{q}(T, Z(t_0); \Delta) = \sum_{j_1=1}^{n} \sum_{j_2=1}^{n_2(t_0)-1} \int_{\Delta \cap [\frac{j_1-1}{n}, \frac{j_1}{n})} \int_{[a_{j_2}, a_{j_2+1})} \frac{\widehat{c}^2(\tau_1, v; t_0)}{\tau_1(1-\tau_1)F_{n,2}(v)(1-F_{n,2}(v))} . d\mu_1(\tau_1) d\mu_2(v).$$

Due to the fact that

$$\begin{aligned} & \int_{\Omega_{V(t_0)}} \frac{\mathrm{cov}^2\{I[T \le Q_T(\tau_1)], I[Z(t_0) \le v]\}}{\tau_1(1-\tau_1)F_2(v)[1-F_2(v)]} dv = 0 \\ \Leftrightarrow \quad & \int_{[0,1)} \frac{\mathrm{cov}^2\{I[T \le Q_T(\tau_1)], I[Z(t_0) \le Q_{Z(t_0)}(\tau_2)]\}}{\tau_1(1-\tau_1)\tau_2(1-\tau_2)} d\tau_2 = 0, \end{aligned}$$

we can easily show that the reduced proposed test statistics is equivalent to the index suggested in Zhu et al. (2018).

Moreover, to obtain the p values, Zhu et al. (2018) suggested a simulation-based procedure to determine the critical value, which takes the following steps:

(i) Generate $Y^*_{i,k}$ independently from uniform distribution for $i = 1, \ldots, n$ and $k = 1, 2$.

(ii) Re-estimate $q^c(Y_1, Y_2; \Delta, \mathcal{I}_2)$ based on $\{Y^*_{i,1}, Y^*_{i,2}\}_{i=1}^n$.

(iii) Repeat this procedure for B times and set c_α to be the upper α quantile of the estimates of $q^c(Y_1, Y_2; \Delta, \mathcal{I}_2)$ obtained from the randomly generated

samples.

It is easy to see from the definition of the index that given different Δ and $\mathcal{I}_2$, the critical values are different. Moreover, the critical values are subject to change when we have discrete covariate or time dependent covariates with different sets of landmark times. These suggest that we need to simulate the critical value for every single case, which introduces additional computation thus not desirable. Alternatively, the proposed resampling procedure is more general, and easy to implement with low computational cost.

4.2 Numerical Studies

4.2.1 Simulations to Compare the Proposed Method with Zhu et al. (2018)

We first conduct some simulations to compare the Proposed Method with Zhu et al. (2018) in the case with continuous outcome without censoring and continuous covariate. As discussed in Section 4.1.4, the proposed test statistics reduce to the equivalent form of the index suggested in Zhu et al. (2018) in this case. The main difference is the method used to conduct the inference. Specifically, we propose a resampling procedure to simulate the limiting null distribution, while Zhu et al. (2018) utilize a simulation-based procedure.

To compare the two methods, we consider the same simulation settings as the Zhu et al. (2018)'s paper:

Set-up I: $Y_1 = A\{Y_2^2 I(Y_2 > 0) + \widetilde{Y}_2^2 I(Y_2 \leq 0)\} + \varepsilon$.

Set-up II: $Y_1 = \exp(AY_2^2)\varepsilon$.

Set-up III: $Y_1 = AY_2^2 + \varepsilon$.

Table 4.1: Empirical rejection rates based on 1000 replicates with continuous outcome without censoring and continuous covariate. The upper numbers show the results based on the proposed test, note as IQI; and the numbers in the below brackets show the results for the test suggested in Zhu et al. (2018), noted as AQI.

Set-up	n	method	A=0	A=1	A=2
	100	IQI	0.049	0.929	0.973
		AQI	(0.050)	(0.954)	(0.986)
	200	IQI	0.043	1.000	1.000
		AQI	(0.048)	(1.000)	(1.000)
	100	IQI	0.049	0.884	0.976
		AQI	(0.050)	(1.000)	(1.000)
	200	IQI	0.043	1.000	1.000
		AQI	(0.048)	(1.000)	(1.000)
	100	IQI	0.049	1.000	1.000
		AQI	(0.050)	(1.000)	(1.000)
	200	IQI	0.043	1.000	1.000
		AQI	(0.048)	(1.000)	(1.000)

Here ε, Y_2 and $\widetilde{Y}_1$ are generated independently from the standard Cauchy distribution. A considered include 0,1 and 2. When $A = 0$, Y_1 and Y_2 are independent in all three models. When $A \neq 0$, there is partial effect on $\tau_1 \in (0.5, 1)$ for Set-up I, partial effect on $\tau_1 \in (0, 0.5) \cup (0.5, 1)$ for Set-up II, and constant effect across (0,1) for Set-up III. The sample size is set as $n = 50$, 100 or 200, and the significance level is set to be $\alpha = 0.05$.

Table 4.1 summarizes the results based on 1000 replicates. From the table, we notice that both the proposed method and Zhu et al. (2018) reserve the correct size of 0.05 under the null cases even with relative small sample size of 50. When the sample size is small, say $n = 100$, there is more difference between the empirical power based on Zhu et al. (2018) and those obtained with the proposed procedure. Meanwhile, we notice that the difference become smaller as the sample size increases. This can be explained by the fact that the estimation of the influence curve tends to have more variance when the sample size is small, leading to relatively lower power. Overall, we can conclude that the proposed procedure is comparable to Zhu et al. (2018).

4.2.2 Simulations Studies for Time-to-event Outcome with Time Independent Covariate

Next, we conduct extensive simulation studies to investigate the finite-sample performance of the proposed resampling-based testing procedures for time-to-event outcome with time independent covariates. We consider six set-ups where T and $\widetilde{Z}$ follow different relationships. In all set-ups, we generate $\widetilde{Z}$ from $Uniform(0,1)$ and generate censoring time C from $Uniform(U_L, U_U)$, where U_L and U_U are properly specified to produce 0%, 15% or 30% censoring. Let $\Phi(\cdot)$ denote the cumulative distribution function of the standard normal distribution. The six simulation set-ups are described as follows.

Set-up I: Generate T such that $Q_\tau\{\log(T)\} = \Phi^{-1}(\tau)$. Set $(U_L, U_U) = (2, 3.8)$ to produce 15% censoring, and set $(U_L, U_U) = (1, 2.5)$ to produce 30% censoring.

Set-up II: Generate T such that $Q_\tau\{\log(T)\} = 0.2X + \Phi^{-1}(\tau)$. Set $(U_L, U_U) = (2.5, 3.9)$ to produce 15% censoring and set $(U_L, U_U) = (1.2, 2.8)$ to produce 30% censoring.

Set-up III: Generate T such that $Q_\tau\{\log(T)\} = 0.5X + \Phi^{-1}(\tau)$. Set $(U_L, U_U) = (2.7, 4.9)$ to produce 15% censoring, and set $(U_L, U_U) = (1.5, 3)$ to produce 30% censoring.

Set-up IV: Generate T such that $Q_\tau\{\log(T)\} = l_4(\tau)X + \Phi^{-1}(\tau)$, where $l_4(\tau)$ is as plotted in Figure 2.1. Set $(U_L, U_U) = (2, 3.9)$ to produce 15% censoring, and set $(U_L, U_U) = (1, 2.5)$ to produce 30% censoring.

Set-up V: Generate T such that $Q_\tau\{\log(T)\} = l_5(\tau)X + \Phi^{-1}(\tau)$, where $l_5(\tau)$ is as plotted in Figure 2.1. Set $(U_L, U_U) = (5.2, 6.5)$ to produce 15% censoring, and set $(U_L, U_U) = (1.5, 3.5)$ to produce 30% censoring.

Set-up VI: Generate T such that $Q_\tau\{\log(T)\} = l_6(\tau)X + \Phi^{-1}(\tau)$, where $l_6(\tau)$ is as plotted in Figure 2.1. Set $(U_L, U_U) = (3.5, 5.5)$ to produce 15% censoring, and set $(U_L, U_U) = (1.1, 3.5)$ to produce 30% censoring.

In Figure 4.1, we plot the true coefficient function $\beta_0^{(1)}(\tau)$ for each set-up. It is remarkable that these are the same set-ups as Cui and Peng (2022). It is easy to see that Set-up I represents a null case, where $\widetilde{Z}$ has no effect on any quantile of T. Set-up II and III are two set-ups where $\widetilde{Z}$ has nonzero constant effects over all $\tau \in [0.1, 0.6]$. The constant effect in Set-up II has a magnitude of 0.2, which is smaller than that in Set-up III, which is 0.5. In Set-ups IV, V, and (VII), $\widetilde{Z}$ has a dynamic effect varying across different τ's. More specifically, $\widetilde{Z}$ has a partial effect over the τ-interval $[0.1, 0.49]$ in Set-up IV. In Set-up V, the magnitude of $\widetilde{Z}$'s effect is symmetric around 0.5, while the sign of the effect is opposite for $\tau < 0.5$ and for $\tau > 0.5$, and the effect equals 0 at $\tau = 0.5$. In Set-up VI, the τ-varying effect pattern of $\widetilde{Z}$ is similar to that in Set-up V except that there is a small interval around 0.5 where $\widetilde{Z}$ has no effect in Set-upVI.

We compare the proposed method, noted as "IQI", with the Wald test based on the Cox PH model, denoted by "CPH (Wald)", the Wald test based on the locally concerned quantile regression that focuses on $\tau = 0.4, 0.5$, or 0.6, denoted by "CQR (Wald)", as well as the tests based on working linear quantile regression model in Cui and Peng (2022), noted as "GST" and "GIT". To implement CQR (Wald), we adopt Peng and Huang (2008)'s estimates with variance estimated by bootstrapping. The resampling size used for both CQR (Wald) and the proposed testing procedures is set as 2500. For all the methods, we consider sample sizes 200 and 400.

In Table 4.2, we report the empirical rejection rates based on 1000 simulations. The results in Set-up I show that the proposed IQI, and the existing tests, GST, GIT, CQR (Wald) and CPH (Wald), have empirical sizes quite close to the nominal level 0.05. The GST yields relatively larger empirical type I errors as compared to the

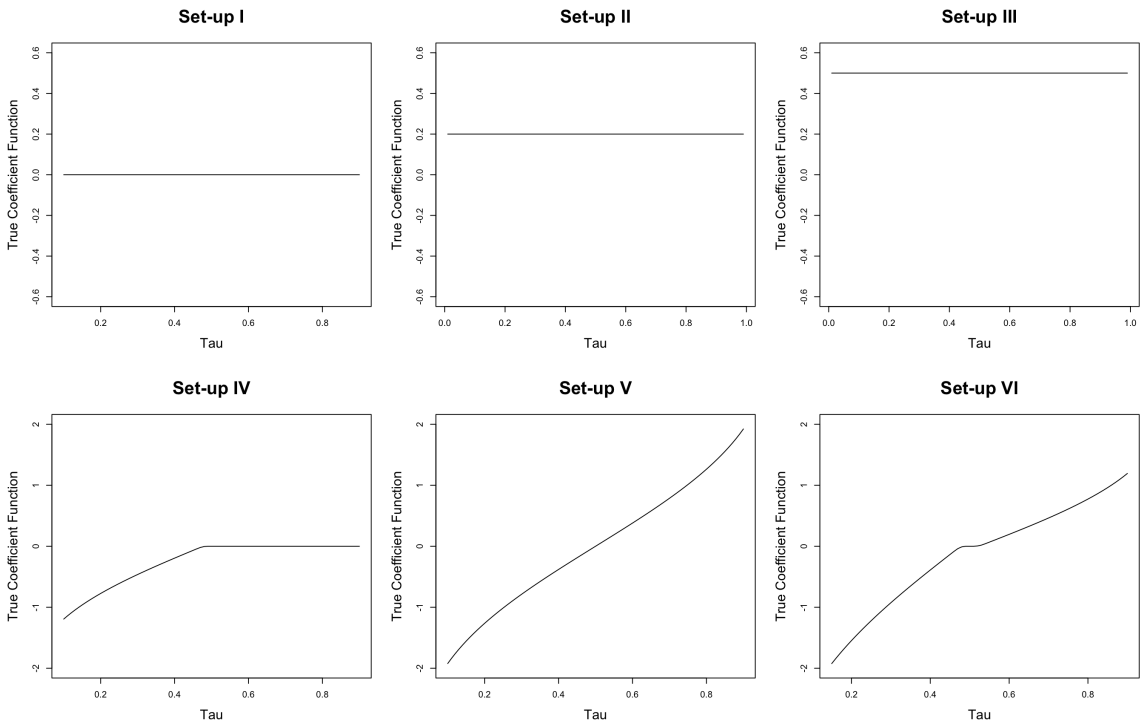


Figure 4.1: The true coefficient function for all simulation set-ups.

other tests, especially when the sample size is not large. Comparably, the proposed IQI is more robust and the empirical size is close to 0.05 with relative small sample size of 200.

When the quantile effect of $\widetilde{Z}$ is constant over τ (i.e. Set-ups II and III), CPH (Wald) has comparable empirical power as compared to the proposed IQI. These observations suggest that in the trivial constant effect cases, the proposed test have comparable power with the traditional Cox regression based tests. In both Set-ups II and III, the locally concerned CQR (Wald) consistently yields lower empirical power than the globally concerned methods, IQI, GIT and GST. This reflects the power benefit resulted from integrating information on covariate effects on different quantiles, rather than focusing on the covariate effect on a single quantile.

In Set-ups IV, V, and VI, the effect of $\widetilde{Z}$ is τ-varying, reflecting its dynamic association with T. In these cases, CPH (Wald), which assumes a constant covariate effect, can have poor power to detect the dynamic effect of $\widetilde{Z}$ (e.g. 14.0% empirical

power in Set-up V with $n = 400$ in the presence of 30% censoring), while the proposed IQI can yield much higher power (e.g. >75% power in Set-up V with $n = 400$ in the presence of 30% censoring). The locally concerned CQR (Wald) can have higher power than CPH (Wald) when the targeted quantile level is within the τ-region where $\beta_0^{(1)}(\tau)$ is non-zero. When the targeted quantile level is outside the τ-region with non-zero effect, such as $\tau = 0.6$ in Set-up IV or $\tau = 0.5$ in Set-ups V and VI, the CQR (Wald) has even poorer power compared to CPH (Wald). This is well expected because these cases may serve as the null cases for the locally concerned CQR (Wald). This confirms that CQR (Wald) is inadequate to capture the meaningful effect of $\widetilde{Z}$ that is manifested at non-targeted quantiles.

We compare the simulation results across settings that are only differed by the censoring distribution. For each relationship between $\widetilde{Z}$ and T specified by Set-ups I-VI, we consider three different censoring distributions to yield 0%, 15%, and 30% censoring. From our comparisons, we find that quantile based tests, including IQI, GST, GIT and CQR (Wald), demonstrate small variations in empirical powers as the censoring rate (or distribution) changes. In cases with a constant covariate effect, the Cox regression based test, CPH (Wald), also has similar performance among settings with different censoring rates. However, in Set-up V, where the covariate effect is not constant over τ, CPH (Wald) has reasonably good power when there is no censoring or only 15% censoring, but its performance deteriorates considerably when the censoring rate is increased to 30%. We have a similar observation for CPH (Wald) in Set-up VI. A reasonable interpretation of these observations is that the capacity to detect a dynamic effect can be weakened by incorrectly assuming a constant proportional hazard effect and can be further attenuated by the missing data from censoring.

In summary, our simulation results demonstrate the robust satisfactory performance of the proposed method for detecting a time independent covariate of either a constant or dynamic effect on censored outcome.

Table 4.2: Empirical rejection rates based on 1000 replicates with censored outcome and continuous covariate.

Set-up	n	IQI	GST	GIT	CQR(Wald)			CPH (Wald)
		$\Delta = (0.1, 0.6)$			$\tau = 0.4$	$\tau = 0.5$	$\tau = 0.6$	
0% censoring								
I	200	0.051	0.098	0.070	0.055	0.052	0.056	0.048
	400	0.057	0.093	0.075	0.069	0.064	0.060	0.047
II	200	0.100	0.215	0.156	0.104	0.108	0.108	0.121
	400	0.178	0.275	0.216	0.162	0.156	0.139	0.183
III	200	0.446	0.541	0.478	0.344	0.374	0.337	0.456
	400	0.712	0.790	0.771	0.589	0.595	0.590	0.745
IV	200	0.183	0.378	0.250	0.074	0.045	0.049	0.060
	400	0.387	0.656	0.476	0.101	0.056	0.055	0.049
V	200	0.387	0.618	0.452	0.106	0.057	0.121	0.428
	400	0.759	0.939	0.828	0.169	0.071	0.165	0.737
VI	200	0.437	0.729	0.543	0.095	0.047	0.088	0.228
	400	0.836	0.971	0.898	0.154	0.048	0.097	0.446
15% censoring								
I	200	0.051	0.100	0.073	0.066	0.062	0.057	0.049
	400	0.061	0.091	0.078	0.072	0.072	0.066	0.051
II	200	0.100	0.234	0.167	0.117	0.131	0.117	0.115
	400	0.175	0.275	0.214	0.155	0.153	0.150	0.178
III	200	0.441	0.566	0.485	0.359	0.401	0.360	0.450
	400	0.705	0.786	0.772	0.585	0.592	0.576	0.722
IV	200	0.180	0.377	0.254	0.097	0.060	0.053	0.063
	400	0.396	0.652	0.478	0.116	0.065	0.063	0.067
V	200	0.383	0.653	0.464	0.143	0.070	0.118	0.260
	400	0.756	0.937	0.827	0.208	0.071	0.153	0.458
VI	200	0.438	0.731	0.552	0.149	0.062	0.086	0.125
	400	0.836	0.971	0.896	0.198	0.055	0.095	0.201
30% censoring								
I	200	0.042	0.171	0.095	0.062	0.060	0.048	0.047
	400	0.053	0.110	0.085	0.069	0.074	0.065	0.056
II	200	0.094	0.302	0.186	0.115	0.122	0.105	0.122
	400	0.168	0.305	0.221	0.152	0.156	0.138	0.188
III	200	0.427	0.681	0.539	0.360	0.393	0.322	0.432
	400	0.697	0.828	0.791	0.585	0.590	0.534	0.703
IV	200	0.156	0.440	0.271	0.101	0.061	0.044	0.056
	400	0.370	0.668	0.480	0.115	0.065	0.062	0.085
V	200	0.384	0.799	0.573	0.135	0.069	0.103	0.092
	400	0.758	0.960	0.846	0.206	0.068	0.135	0.140
VI	200	0.425	0.803	0.587	0.148	0.063	0.077	0.053
	400	0.823	0.978	0.903	0.199	0.052	0.082	0.064

4.2.3 Simulations Studies for Time-to-event Outcome with Time Dependent Covariate

4.2.3.1 Data Generation Steps and Simulation Set-ups

It is not direct to generate varying covariate effect based on the time dependent cox model. We propose to generate the residual survival time sequentially based on the quantile regression model.

Consider a sequence of time $\{t_1, t_2, \ldots, t_L\} = \{0, 0.5, \ldots, 5\}$. We first generate the covariate process $\bar{Z}_i(t)$ on $t_j \in \{t_1, t_2, \ldots, t_J\}$ from a pre-specified distribution. Then, to simulate the survival time, the basic idea is that we generate the residual survival time for each t_j given that the $T > t_j$, and then truncate the survival time with t_{j+1}. Specifically, for each individual, we take the following steps:

(a) Initialize $j = 1$. Generate the conditional quantile $\tau_j \sim Uniform(0, 1)$.

(b) Generate the log residual survival time $\log(T_j - t_j)$ given that there is no event at t_j, from the model with $\boldsymbol{V}(t_j) = \boldsymbol{Z}(t_j)$:

$$Q_{\log(T_j - t_j)}\{\tau_j | \bar{\boldsymbol{Z}}(t_j)\} = \boldsymbol{V}(t_j)^{\intercal} \xi_j(\tau_j, t_j) + \Phi^{-1}(\tau_j). \tag{4.3}$$

(c) If $T_j \leq t_{j+1}$ or $j = J$, then $T = T_j$ and stop; otherwise, set $j = j + 1$ and go to step (b).

We then generate the log censoring time C from the Uniform distribution. This leads to $\delta = I(T \leq C)$ and $X = \min\{T, C\}$.

Moreover, to more easily generate the varying covariate effects, we consider an intermediate true coefficient function $\eta(\tau)$ for the covariate effect we set for $\boldsymbol{V}(t_j)$. Specifically, given the time sequence $\{t_1, t_2, \ldots\}$, we take the following steps:

(a) Start from $t_1 = 0$. At the baseline, the model is

$$Q_{\log T_1}\{\tau_1|\bar{\boldsymbol{Z}}(t_1)\} = \boldsymbol{V}(t_1)^\intercal \xi_1(\tau_1, t_1) + \Phi^{-1}(\tau_1), \tau_1 \in (0,1), \tag{4.4}$$

where $\xi_1(\tau_1, t_1) = \eta(\tau_1)$. From the model (4.4), given $\boldsymbol{V}(t_1)$, we can solve the equation $\log t_2 - \boldsymbol{V}(t_1)^\intercal \xi_1(\tau, t_1) - \Phi^{-1}(\tau) = 0$ for τ and denote the unique solution as τ_2^*. With the data generation procedure, we will truncate the time with t_2 and stop if $T_1 \le t_2$, i.e., $\tau_1 \le \tau_2^*$. Define $\tau_C^{(2)} = \tau_2^*$.

(b) Next we consider t_2. Given event free at t_2, the model is

$$Q_{\log(T_2 - t_2)}\{\tau_2|\bar{\boldsymbol{Z}}(t_2)\} = \boldsymbol{V}(t_2)^\intercal \xi_2(\tau_2, t_2) + \Phi^{-1}(\tau_2), \tau_2 \in (0,1), \tag{4.5}$$

where $\xi_2(\tau_2, t_2) = \eta\{\tau_C^{(2)} + (1 - \tau_C^{(2)})\tau_2\}$. Similarly, from the model (4.5), given $\boldsymbol{V}(t_2)$, we can solve the equation $\log(t_3 - t_2) - \boldsymbol{V}(t_2)^\intercal \xi_2(\tau, t_2) - \Phi^{-1}(\tau) = 0$ for τ and denote the unique solution as τ_3^*. With the data generation procedure, we will truncate the time with t_3 and stop if $T_2 \le t_3$, i.e., $\tau_2 \le \tau_3^*$. Define $\tau_C^{(3)} = \tau_C^{(2)} + (1 - \tau_C^{(2)})\tau_2^*$. Note that we can directly obtain $\tau_C^{(3)}$ by solving

$$\begin{aligned} 0 &= \log(t_3 - t_2) - \boldsymbol{V}(t_2)^\intercal \xi_0(\tau_2^*, t_2) - \Phi^{-1}(\tau_2^*) \\ &= \log(t_3 - t_2) - \boldsymbol{V}(t_2)^\intercal \eta\{\tau_C^{(2)} + (1 - \tau_C^{(2)})\tau_2^*\} - \Phi^{-1}(\tau_2^*) \\ &= \log(t_3 - t_2) - \boldsymbol{V}(t_2)^\intercal \eta\{\tau_C^{(3)}\} - \Phi^{-1}\left(\frac{\tau_C^{(3)} - \tau_C^{(2)}}{1 - \tau_C^{(2)}}\right) \end{aligned}$$

(c) Suppose for t_j, we can similarly obtain $\tau_C^{(j+1)}$ by solving

$$\log(t_{j+1} - t_j) - \boldsymbol{V}(t_j)^\intercal \eta\{\tau_C^{(j+1)}\} - \Phi^{-1}\left(\frac{\tau_C^{(j+1)} - \tau_C^{(j)}}{1 - \tau_C^{(j)}}\right) = 0.$$

Then for t_{j+1}, given event free at t_{j+1}, the model is

$$Q_{\log(T_{j+1}-t_{j+1})}\{\tau_{j+1}|\bar{\boldsymbol{Z}}(t_{j+1})\} = \boldsymbol{V}(t_{j+1})^{\intercal}\xi_{j+1}(\tau_{j+1}, t_{j+1}) + \Phi^{-1}(\tau_{j+1}), \tau_{j+1} \in (0,1), \tag{4.6}$$

where $\xi_{j+1}(\tau_{j+1}, t_{j+1}) = \eta\{\tau_C^{(j+1)} + (1-\tau_C^{(j+1)})\tau_{j+1}\}$. From the model (4.6), given $\boldsymbol{V}(t_{j+1})$, we can solve the equation $\log(t_{j+2} - t_{j+1}) - \boldsymbol{V}(t_{j+1})^{\intercal}\xi_{j+1}(\tau, t_{j+1}) - \Phi^{-1}(\tau) = 0$ for τ and denote the unique solution as τ_{j+2}^*. With the data generation procedure, we will truncate the time with t_{j+2} and stop if $T_{j+1} \leq t_{j+2}$, i.e., $\tau_{j+1} \leq \tau_{j+2}^*$. Define $\tau_C^{(j+2)} = \tau_C^{(j+1)} + (1 - \tau_C^{(j+1)})\tau_{j+2}^*$. We can directly obtain $\tau_C^{(j+2)}$ by solving

$$\log(t_{j+2} - t_{j+1}) - \boldsymbol{V}(t_{j+1})^{\intercal}\eta\{\tau_C^{(j+2)}\} - \Phi^{-1}\left(\frac{\tau_C^{(j+2)} - \tau_C^{(j+1)}}{1 - \tau_C^{(j+1)}}\right) = 0.$$

The simulation set-ups we consider include:

Set-up I: The null case without covariate effects;

Set-up II: The case with constant covariate effects of magnitude 1;

Set-up IV: The case with partial covariate effect with symmetric effects on both the lower and upper quantile level;

Set-up V: The case with partial covariate effect on both the lower and upper quantile level with larger effects on the lower quantile level.

In Figure 4.2, we plot the intermediate function $\eta(\tau)$ we used for each set-up. In this simulation study, the censoring rate is about 30% through choosing parameter for the Uniform distribution.

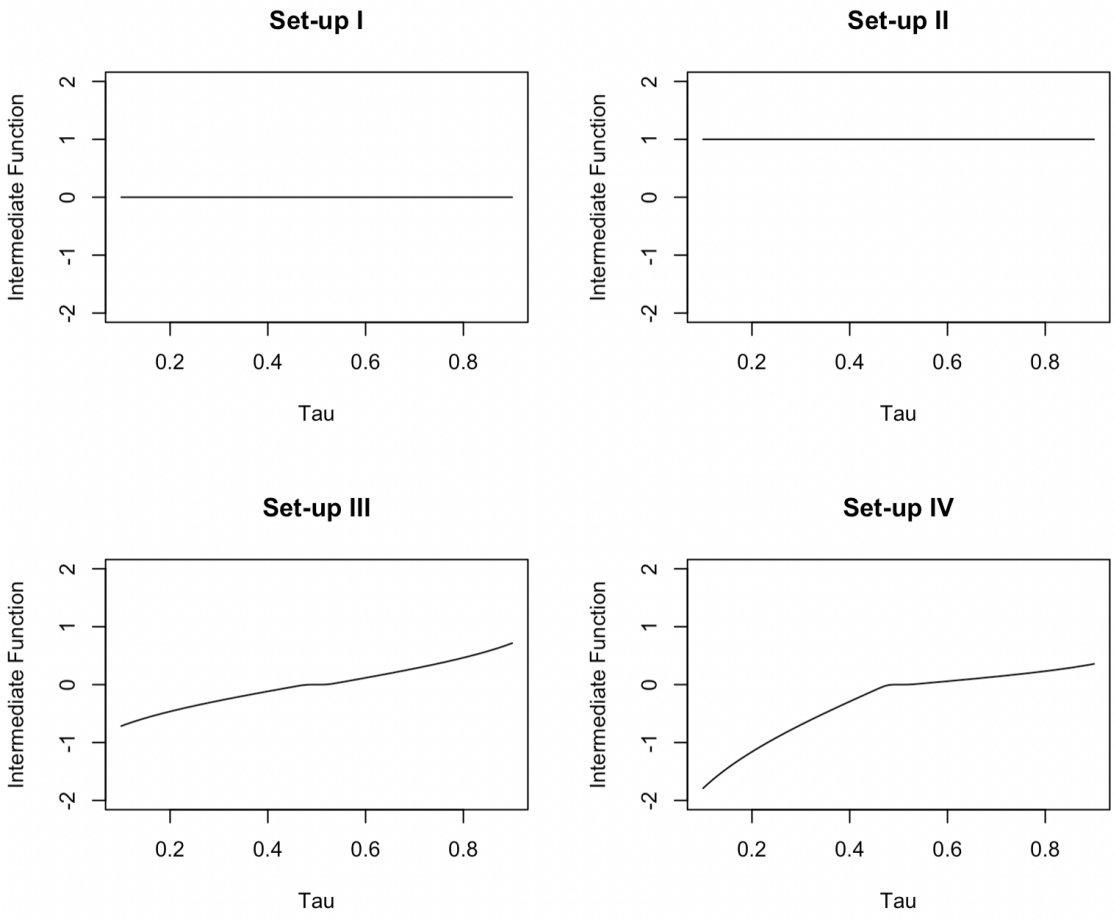


Figure 4.2: The intermediate function $\eta(\tau)$ used for all simulation set-ups.

4.2.3.2 Simulation Results

We first consider the case with continuous time dependent covariates. Specifically, we generate the covariate process $\bar{Z}_i(t)$ on $t_j \in \{t_1, t_2, \ldots, t_J\}$ from the normal distribution, i.e., $Z_i(t_j) \sim N(\mu_i, 0.25)$ with $\mu_i \sim Uniform(-0.5, 0.5)$. For each generated data, we compare the proposed methods based on $\widehat{q}_{max}(T, \bar{Z}; \Delta, \mathcal{T})$ and $\widehat{q}_{sum}(T, \bar{Z}; \Delta, \mathcal{T})$, noted respectively as "IQI(MAX)" and "IQI(SUM)", with the Wald test based on the time dependent Cox PH model, denoted by "CPH (Wald)". The sample size we considered is 200. For the set of landmark time points considered, we consider $R_1 = \{0, 2\}$, $R_2 = \{0, 1, 2\}$ and $R_3 = \{0, 0.5, 1, 1.5, 2\}$.

Table 4.3 summarize the empirical rejection rates based on 1000 replicates. The results in Set-up I show that the proposed IQI(MAX), IQI(SUM) and CPH (Wald), all have empirical sizes quite close to the nominal level 0.05 with relatively small sample size of 200. When the quantile effect of $\bar{Z}(t)$ is constant over τ (i.e. Set-ups II), CPH (Wald) has comparable empirical power as compared to the proposed IQI(MAX) and IQI(SUM). These observations suggest that in the trivial constant effect cases, the proposed test have comparable power with the traditional Cox regression based tests. In Set-ups III and IV, the effect of $\bar{Z}(t)$ is τ-varying, reflecting its dynamic association with T. In these cases, CPH (Wald), which assumes a constant covariate effect, can have poor power to detect the dynamic effect of $\bar{Z}(t)$ (e.g. 5.3% empirical power in Set-up IV), while the proposed IQI(MAX) and IQI(SUM) can yield much higher power (e.g. >80% power with landmark time set R_3 in Set-up IV). Moreover, we notice that the power tends to increase with more landmark time points considered. This can be explained by the fact that we are able to utilize more information from the observed data by considering more landmark time points to capture the dynamic effect. This results in a gain in statistical power, as more effects can be detected. Finally, it is worth mentioning that the two proposed methods, IQI(MAX) and IQI(SUM), have comparable performance with IQI(SUM) tends to

Table 4.3: Empirical rejection rates based on 1000 replicates for sample size of 200 with censored outcome and continuous time dependent covariate.

Set-up	Method	Landmark Time Set		
		R_1	R_2	R_3
I	IQI(MAX)	0.057	0.057	0.054
	IQI(SUM)	0.057	0.055	0.051
	CPH(Wald)	0.049	0.049	0.049
II	IQI(MAX)	1.000	1.000	1.000
	IQI(SUM)	1.000	1.000	1.000
	CPH(Wald)	1.000	1.000	1.000
III	IQI(MAX)	0.302	0.460	0.494
	IQI(SUM)	0.357	0.485	0.541
	CPH(Wald)	0.097	0.097	0.097
IV	IQI(MAX)	0.616	0.770	0.822
	IQI(SUM)	0.691	0.807	0.861
	CPH(Wald)	0.053	0.053	0.053

be more powerful in most alternative cases. As a result, these two methods can be used interchangeably.

Next, we investigate the case with binary time dependent covariates. Specifically, we generate the covariate process $\bar{Z}_i(t)$ on $t_j \in \{t_1, t_2, \ldots, t_J\}$ from the binary distribution, i.e., $Z_i(t_j) \sim Bernoulli(0.5)$. For each generated data, we compare the three methods, the proposed "IQI(MAX)", "IQI(SUM)" and "CPH (Wald)". For the set of landmark time points considered, we consider $R_1 = \{0, 2\}$, $R_2 = \{0, 1, 2\}$ and $R_3 = \{0, 0.5, 1, 1.5, 2\}$. The sample size we considered is 200, 400 and 800.

Table 4.4 summarize the empirical rejection rates based on 1000 replicates. From the table, we have similar findings as the cases with continuous time dependent covariates. Additionally, we look at different sample sizes in this case. From the table, we can conclude that as the sample size increases, the empirical size is getting closer to the nominal level of 0.05 under the null case (i.e. Set-up I), and the empirical power if get closer to 1 under the alternative cases (i.e. Set-up II-IV).

Finally, we investigate the case with factor time dependent covariates. Specifically, we generate the covariate process $\bar{Z}_i(t)$ on $t_j \in \{t_1, t_2, \ldots, t_J\}$ from the multinomial

Table 4.4: Empirical rejection rates based on 1000 replicates with censored outcome and binary time dependent covariate.

Set-up	n	Time Set	IQI(MAX)	IQI(SUM)	CPH(Wald)
		R_1	0.066	0.066	0.044
	200	R_2	0.065	0.066	0.043
		R_3	0.066	0.072	0.044
		R_1	0.051	0.053	0.062
I	400	R_2	0.047	0.050	0.062
		R_3	0.056	0.046	0.062
		R_1	0.049	0.049	0.051
	800	R_2	0.045	0.054	0.051
		R_3	0.051	0.045	0.051
		R_1	1.000	1.000	1.000
	200	R_2	1.000	1.000	1.000
		R_3	1.000	1.000	1.000
		R_1	1.000	1.000	1.000
II	400	R_2	1.000	1.000	1.000
		R_3	1.000	1.000	1.000
		R_1	1.000	1.000	1.000
	800	R_2	1.000	1.000	1.000
		R_3	1.000	1.000	1.000
		R_1	0.233	0.353	0.164
	200	R_2	0.218	0.265	0.164
		R_3	0.211	0.221	0.164
		R_1	0.538	0.728	0.275
III	400	R_2	0.497	0.586	0.275
		R_3	0.477	0.475	0.275
		R_1	0.933	0.979	0.488
	800	R_2	0.913	0.952	0.488
		R_3	0.907	0.923	0.488
		R_1	0.299	0.435	0.192
	200	R_2	0.274	0.311	0.192
		R_3	0.252	0.225	0.192
		R_1	0.667	0.817	0.259
IV	400	R_2	0.634	0.698	0.259
		R_3	0.593	0.520	0.260
		R_1	0.986	0.994	0.472
	800	R_2	0.983	0.983	0.472
		R_3	0.976	0.929	0.472

distribution, i.e., $Z_i(t_j) \sim Multinomial(1, (0.25, 0.25, 0.25, 0.25))$ with levels "A", "B", "C" and "D". For each generated data, we apply the proposed "IQI(MAX)", "IQI(SUM)" with different inherent orders. Specifically, if the order is specified as "A" < "B" < "C" < "D", we will set "A" = 1, "B" = 2, "C" = 3 and "D" = 4. In this case, under the formulation of the test statistics, "D" is the reference level and the test statistics capture the overall covariate effects for "A", "B" and "C". The proposed method is then compared with "CPH (Wald)". For the set of landmark time points considered, we consider $R_1 = \{0, 2\}$, $R_2 = \{0, 1, 2\}$ and $R_3 = \{0, 0.5, 1, 1.5, 2\}$. The sample size we considered is 200, 400 and 800.

Table 4.5 summarize the empirical rejection rates based on 1000 replicates. From the table, we have similar findings as the previous cases with either continuous or binary time dependent covariates. Furthermore, when factor covariates are considered, we observe that the results can vary depending on the inherent order. For instance, the empirical power with landmark time set R_3 in Set-up IV is 50.4% and 66.8% respectively for "IQI(MAX)" and "IQI(SUM)" with "D" < "C" < "B" < "A", which are larger than the power of 28.9% and 45.9% respectively for "IQI(MAX)" and "IQI(SUM)" with "A" < "B" < "C" < "D". This may be due to different reference levels for the covariates, i.e. the overall covariate effects for "B", "C" and "D" may be stronger than that for "A", "B" and "C".

All these simulation results suggest that the proposed method can robustly and effectively detect the dynamic effect of time dependent covariates on censored outcome.

4.3 Real Example with FIRST Data

To illustrate the proposed testing framework, we applied the proposed method to 140 infants with cystic fibrosis (CF) after diagnosis from newborn screening from the

Table 4.5: Empirical rejection rates based on 1000 replicates with censored outcome and binary time dependent covariate.

Set-up	n	Time Set	"A" < "B" < "C" < "D"		"D" < "C" < "B" < "A"		
			IQI(MAX)	IQI(SUM)	IQI(MAX)	IQI(SUM)	CPH(Wald)
		R_1	0.054	0.048	0.054	0.048	0.040
	200	R_2	0.048	0.042	0.055	0.054	0.040
		R_3	0.043	0.049	0.050	0.059	0.040
		R_1	0.056	0.054	0.048	0.050	0.042
I	400	R_2	0.058	0.051	0.057	0.054	0.042
		R_3	0.057	0.050	0.049	0.052	0.042
		R_1	0.060	0.059	0.055	0.050	0.056
	800	R_2	0.057	0.059	0.052	0.049	0.056
		R_3	0.054	0.055	0.048	0.037	0.056
		R_1	1.000	1.000	1.000	1.000	1.000
	200	R_2	1.000	1.000	1.000	1.000	1.000
		R_3	1.000	1.000	1.000	1.000	1.000
		R_1	1.000	1.000	1.000	1.000	1.000
II	400	R_2	1.000	1.000	1.000	1.000	1.000
		R_3	1.000	1.000	1.000	1.000	1.000
		R_1	1.000	1.000	1.000	1.000	1.000
	800	R_2	1.000	1.000	1.000	1.000	1.000
		R_3	1.000	1.000	1.000	1.000	1.000
		R_1	0.299	0.471	0.438	0.650	0.181
	200	R_2	0.424	0.611	0.431	0.701	0.182
		R_3	0.416	0.599	0.449	0.705	0.182
		R_1	0.604	0.898	0.888	0.975	0.363
III	400	R_2	0.842	0.975	0.889	0.987	0.362
		R_3	0.853	0.968	0.902	0.989	0.363
		R_1	0.971	0.999	0.999	1.000	0.621
	800	R_2	1.000	1.000	1.000	1.000	0.621
		R_3	0.999	1.000	1.000	1.000	0.621
		R_1	0.256	0.438	0.574	0.764	0.233
	200	R_2	0.319	0.491	0.536	0.750	0.232
		R_3	0.289	0.459	0.504	0.668	0.232
		R_1	0.555	0.868	0.958	0.994	0.447
IV	400	R_2	0.700	0.936	0.955	0.997	0.447
		R_3	0.701	0.914	0.942	0.984	0.447
		R_1	0.978	1.000	1.000	1.000	0.810
	800	R_2	0.996	1.000	1.000	1.000	0.810
		R_3	0.996	1.000	1.000	1.000	0.810

Feeding Infants Right.. from the Start (FIRST) study. A question of interest in the study is to investigate whether there exists a relationship between feeding patterns and susceptibility to infection in CF infants. In this dataset, time to infection, noted as T, is censored in about 4.3% infants due to the end of study time as 12 months for each infant. The covariate process, noted as $\bar{Z} = \{Z(t), 0 \leq t < \infty\}$, is the feeding information (B: breastmilk feeding; P: partial breastmilk feeding; F: formula feeding) collected about bi-weekly for the first month, monthly from 1 to 6 months, and bi-monthly from 6 to 12 months.

Motivated from the scientific needs, we considered two ways to elaborate the feeding information. One method is to consider the feeding variable:

$$V_1(\{Z(t), 0 < t \leq t_0\}) = Z(t_0), \quad t_0 \geq 0,$$

which represents the most recent feeding information. Investigating the relationship between $V_1(\{Z(t), 0 < t \leq t_0\})$ and T can help determine the instant effect of feeding. For the choice of landmark time, we considered either the visit times $\mathcal{T}_1 = \{0, 0.5, 1, 2, 3, 4, 5\}$ or equally grid $\mathcal{T}_2 = \{0, 0.5, \ldots, 5\}$. The other approach is to consider

$$V_2(\{Z(t), 0 < t \leq t_0\}) = \frac{\int_{\max(d, t_0 - t_l)}^{t_0} Z(t-d)dt}{\min(t_0 - d, t_l)}, \quad t_0 \geq d$$

where we set $d = 1$ and $t_l = 3$. In this way, we are interested in the delayed 1 month cumulative effect with a duration of up to 3 months. The landmark time set was then either $\mathcal{T}_1 = \{2, 3, 4, 5\}$ or $\mathcal{T}_2 = \{1.5, 2, \ldots, 5\}$.

We started from some exploratory analysis. We first conducted the univariate analysis to fit the censored quantile regression model between the variables of each feeding type (B,P or F) and the outcome at each landmark time. Figure 4.4 shows the plots for the predicted distribution of the residual time for the two "sample" case

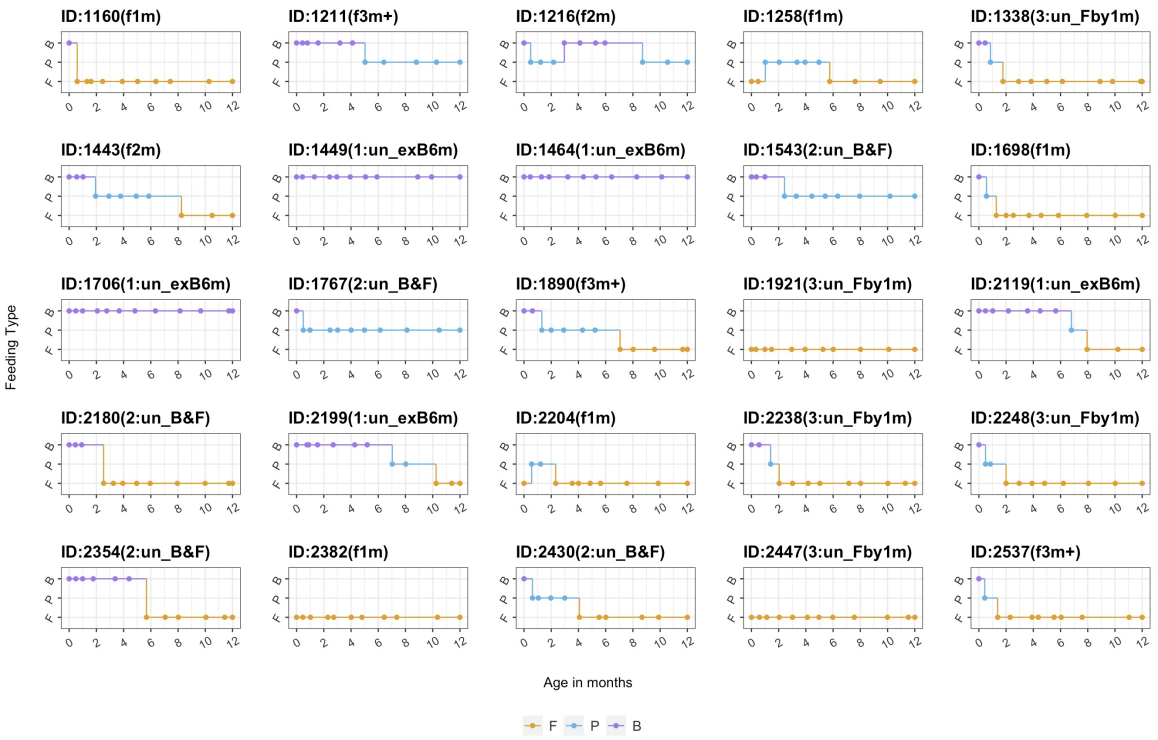


Figure 4.3: The feeding information over time for the first 25 infants.

with $X = 1$ and $X = 0$, where X represents the indicator of whether the feeding type observed at the landmark time t_0 is B (or P, or F). Figure 4.5 shows the plots for the predicted distribution of the residual time for the two "sample" case with $X = 1$ and $X = 0$, where X represents the duration of B (or P, or F) up to 3 months at $t_0 - 1$ for a landmark time t_0. From the figures, we can see some discrepancy between the two lines in the plots for F at the later landmark time, for example $t_0 = 3, 4$ and 5. We also conducted the multivariate analysis to fit the censored quantile regression model with feeding variables treated as factor with B as the reference level and multivariate covariate (P and F) for the first and second method, respectively. Figure 4.7 shows the plots for the p values. From the figure, we also notice some significant results for F at the later landmark time, for example $t_0 = 3$ and 5 in the instant effect case, and $t_0 = 2, 3, 4$ and 5 for the cumulative delayed effect case. In general, these figures suggest that formula feeding has a significant protective effect against infection in CF infants at later time points.

Next, we applied the proposed method to the data. As a comparison, we also applied the time dependent Cox model. The p values are reported in Table 4.6. From the table, it is evident that the proposed method is capable of detecting the significant effects related to formula feeding, whereas the time dependent Cox model fail to detect them. To view more details of the effect based on the proposed test statistics, we further suggest to look at $\frac{c(\tau_1,v;t_0)}{\sqrt{\tau_1(1-\tau_1)F_2(v;t_0)(1-F_2(v;t_0))}}$, which reflects the interval quantile correlation with positive value suggesting a protective effect. Moreover, we may use $\frac{c^2(\tau_1,v;t_0)}{\tau_1(1-\tau_1)F_2(v;t_0)(1-F_2(v;t_0))}$ to gain some insight into the magnitude of the effect. The plots for these two indices over time are presented as Figure 4.7. The figure suggests that formula feeding may have a protective effect against infection in later months, while breastmilk feeding may have some negative effects. This is consistent findings as Figures 4.4 and 4.5. In order to gain a deeper understanding of the cause of these effects, we conducted a subgroup analysis on the unfortified and fortified groups in addition to analyzing the entire population. From the p values in Table 4.6, we can see that for the delayed cumulative effect case, there are significant effects for breastmilk feeding in the unfortified groups, while not in the fortified groups. This suggests that the negative effect of breastmilk feeding could potentially be caused by inadequate fortification of the feedings.

4.4 Remarks

In this work, we develop a new testing framework that can help determine outcome-relevant covariates with univariate time dependent covariate. The proposed methods sensibly adopt a global perspective that examines covariate effects over a continuum of outcome quantiles across a set of landmark times, and model-free. Our numerical studies strongly support the advantages of the proposed methods over the time dependent Cox model, particularly in data settings with dynamic covariate effects.

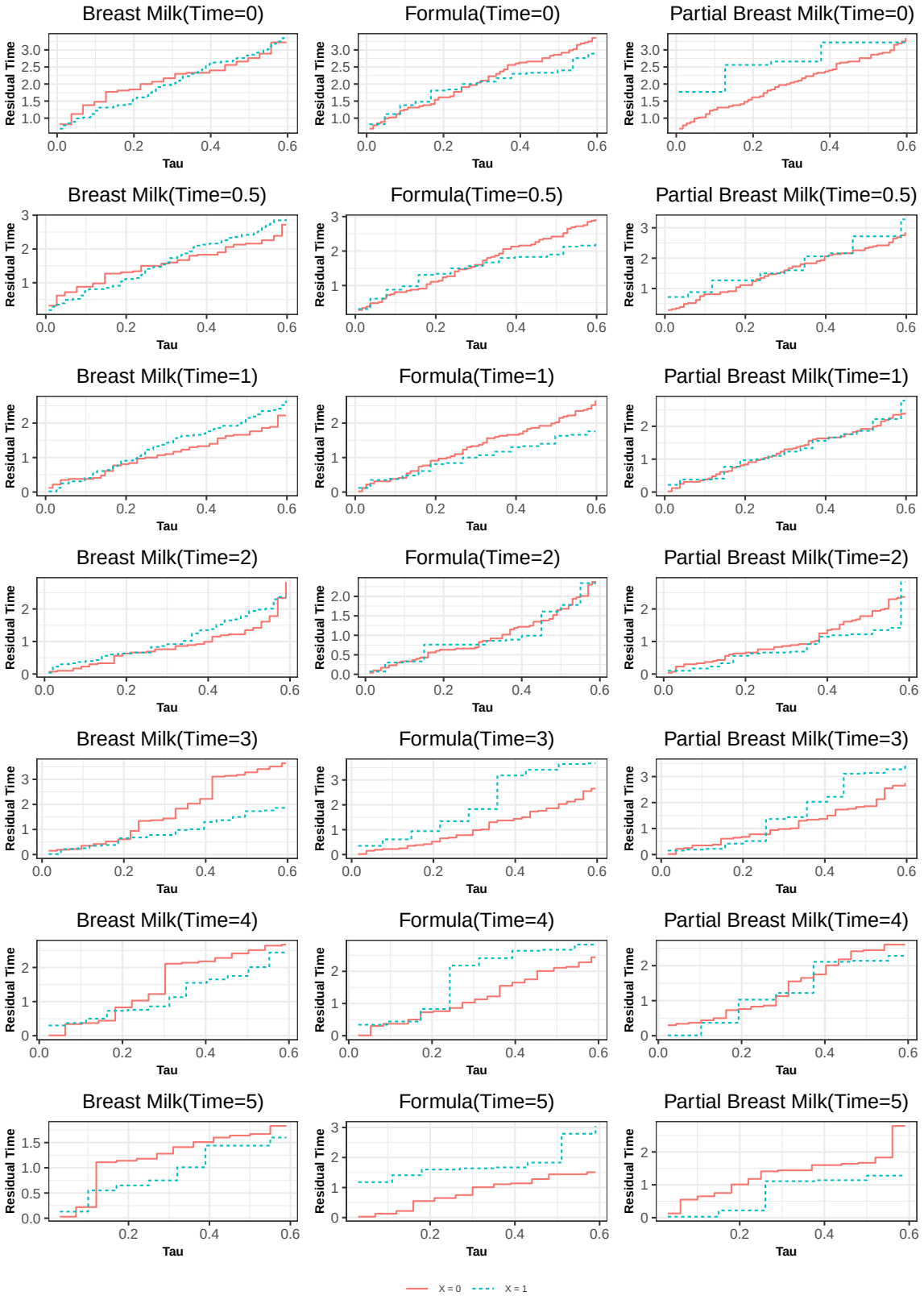


Figure 4.4: Univariate Exploratory Analysis: Plots for a "two-sample" case based on fitting CQR at different landmark time points. In these cases, the feeding variable represents the instant effect.

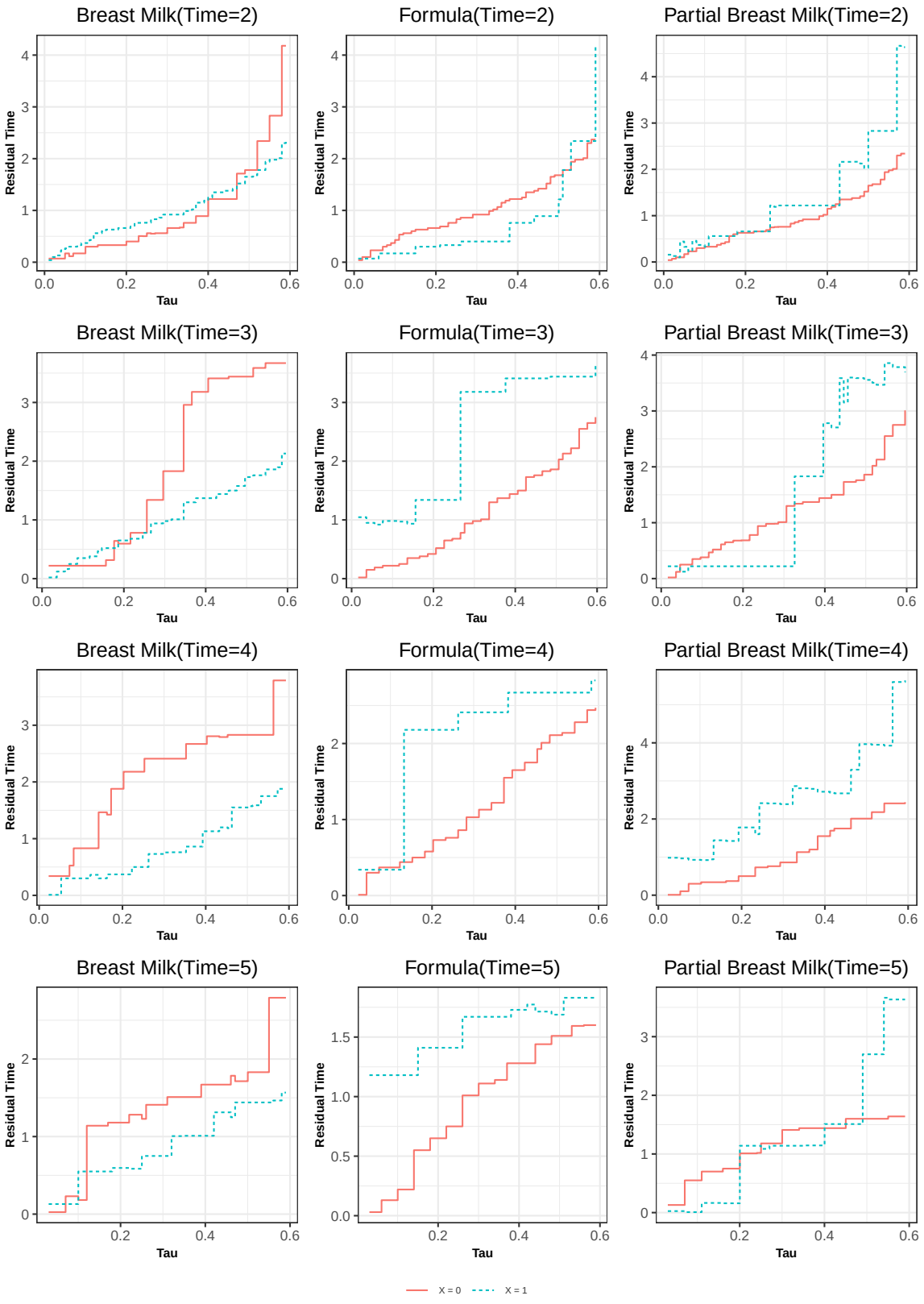


Figure 4.5: Univariate Exploratory Analysis: Plots for a "two-sample" case based on fitting CQR at different landmark time points. In these cases, the feeding variable represents the delayed cumulative effect up to 3 months with $d = 1$ month.

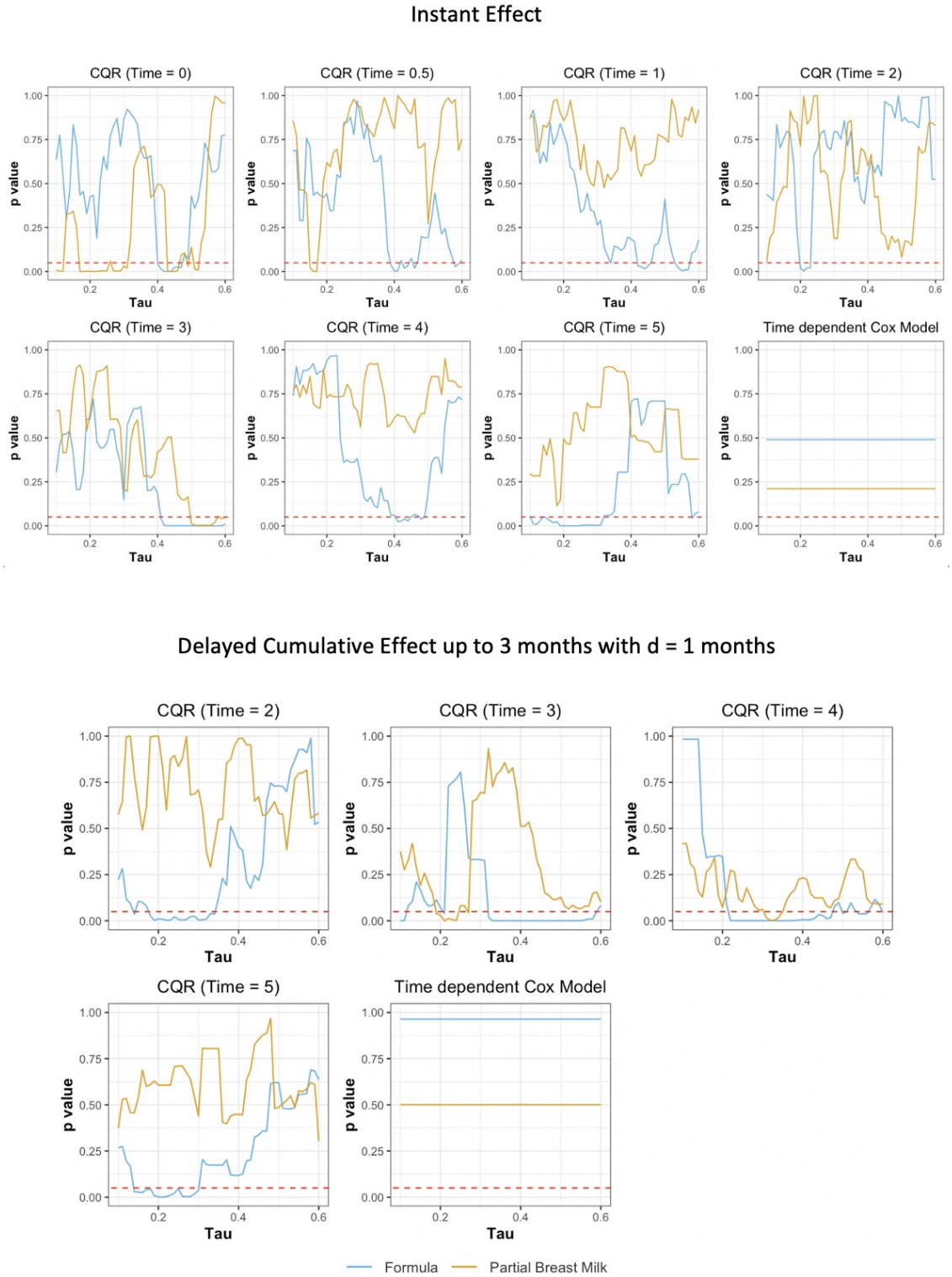
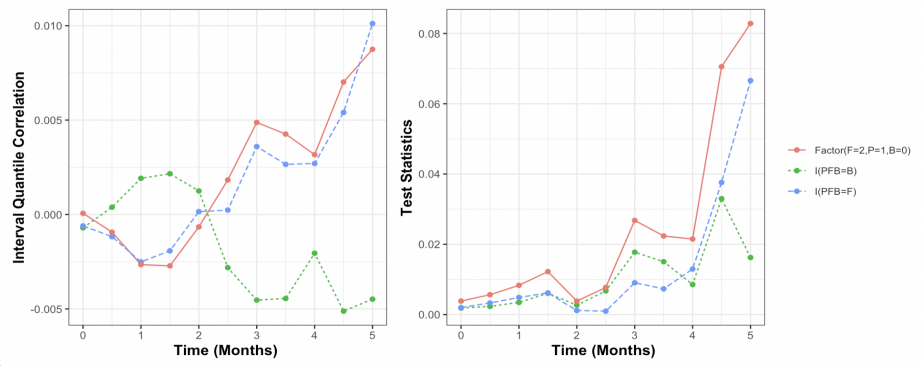
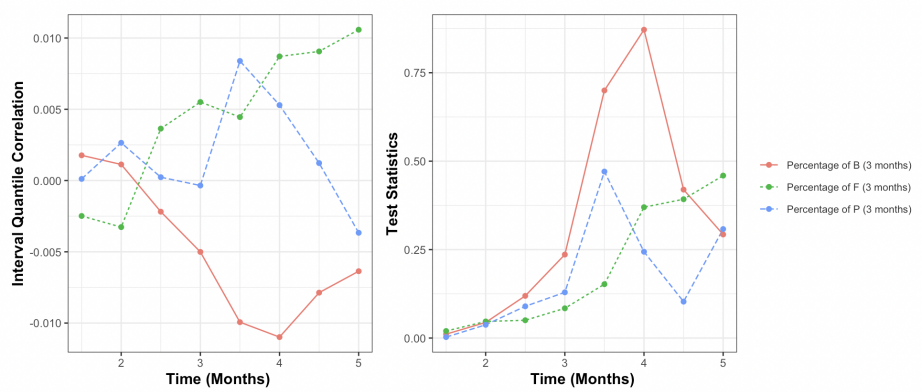


Figure 4.6: Multivariate Exploratory Analysis: Plots for the p values with censored quantile regression (CQR) at different landmark time points, and the time dependent Cox model.



(a) Instant Effect



(b) Cumulative Delayed Effect

Figure 4.7: Proposed Method: Plots for the interval quantile correlation and the test statistics at different time points.

Table 4.6: The p-values based on the proposed method and the time dependent Cox model with different $\mathcal{T}$.

Group	Variable	IQI(MAX)		IQI(SUM)		CPH(Wald)
		$\mathcal{T}_1$	$\mathcal{T}_2$	$\mathcal{T}_1$	$\mathcal{T}_2$	
Instant effect: $V_1(\{\bar{Z}(t), 0 < t \leq t_0\}) = Z(t_0)$						
All	I(PFB=B)	0.473	0.235	0.374	0.216	0.237
	I(PFB=F)	0.009	0.018	0.044	0.066	0.819
	Factor(Ref:B)	0.040	0.070	0.093	0.090	0.407
Unfortified	I(PFB=B)	0.159	0.068	0.159	0.073	0.462
	I(PFB=F)	0.050	0.046	0.065	0.043	0.869
	Factor(Ref:B)	0.111	0.028	0.084	0.044	0.602
Fortified	I(PFB=B)	0.471	0.597	0.552	0.678	0.352
	I(PFB=F)	0.088	0.107	0.260	0.376	0.790
	Factor(Ref:B)	0.240	0.299	0.391	0.551	0.481
Delayed cumulative effect: $V_2(\{\bar{Z}(t), 0 < t \leq t_0\}) = \frac{\int_{\max(d, t_0 - t_l)}^{t_0} Z(t-d)dt}{\min(t_0 - d, t_l)}, \quad t_0 \geq d$						
All	Percentage of B ($t_l = 3, d = 1$)	0.036	0.048	0.039	0.031	0.756
	Percentage of F ($t_l = 3, d = 1$)	0.034	0.056	0.012	0.025	0.807
	Percentage of P ($t_l = 3, d = 1$)	0.427	0.225	0.334	0.278	0.477
Unfortified	Percentage of B ($t_l = 3, d = 1$)	0.015	0.020	0.018	0.010	0.353
	Percentage of F ($t_l = 3, d = 1$)	0.047	0.063	0.053	0.062	0.860
	Percentage of P ($t_l = 3, d = 1$)	0.153	0.180	0.059	0.048	0.198
Fortified	Percentage of B ($t_l = 3, d = 1$)	0.763	0.891	0.746	0.840	0.939
	Percentage of F ($t_l = 3, d = 1$)	0.076	0.114	0.056	0.081	0.770
	Percentage of P ($t_l = 3, d = 1$)	0.375	0.412	0.204	0.296	0.608

Meanwhile, it is worth mentioning that this is a univariate approach that does not allow for adjusting for imbalances in baseline covariates. The extension to generalize the proposed method to multivariate setting is not straightforward, which may worth further exploration.

4.5 Appendix

Denote $\approx$ as asymptotic equivalence uniformly.

4.5.1 Some Lemmas and the proofs

4.5.1.1 Lemma A4.1 and the proof

We state the following necessary regularity condition.

Condition 4.1. *π_{t_0} is bounded below by zero, i.e.* $\min_{t_0 \in \mathcal{T}} \pi_{t_0} > 0$.

Condition 4.2. *$G^r_{C,t_0}(q_{1,t_0}(\tau))$ is uniformly bounded below by zero for $\tau \in \Delta$, i.e.* $\inf_{\tau\in\Delta} G^r_{C,t_0}(q_{1,t_0}(\tau)) > 0$.

Lemma A4.1. *Assume the density function f_1 and its first derivation with respect to τ are bounded away from zero and infinity on Δ. Under Conditions 4.1 and 4.2, we have with probability one*

$$\limsup_{n\to\infty} \pm \left\{ \frac{\widehat{q}_{1,t_0}(\tau_1) - q_{1,t_0}(\tau_1) - [\tau_1 - F_{n,1}(q_{1,t_0}(\tau_1))]/f_1(q_{1,t_0}(\tau_1))}{n^{-3/4}(\log\log n)^{3/4}} \right\} = \frac{2^{5/4}}{3^{3/4}} \frac{\tau_1^{1/2}(1-\tau_1)^{1/2}}{f_1(q_{1,t_0}(\tau_1))}.$$

Proof. Given that $\{X_i^r(t_0), M_i(t_0), \delta_i\}_{i=1}^n$ are i.i.d. random variables. To show Lemma A4.1, the key point is to first show that following similar idea of Bahadur representation (Bahadur, 1966), $\widehat{q}_{1,t_0}(\tau_1)$ can be written as

$$\widehat{q}_{1,t_0}(\tau_1) = q_{1,t_0}(\tau_1) + \frac{\frac{1}{n}\sum_{i=1}^n \frac{I[X_i^r(t_0) > q_{1,t_0}(\tau_1), M_i(t_0)=1]}{\widehat{G}^r_{C,t_0}(q_{1,t_0}(\tau_1))\widehat{\pi}_{t_0}} - (1-\tau_1)}{f_1(q_{1,t_0}(\tau_1))} + R_n(\tau_1),$$

where $R_n(\tau_1) = O_p[n^{-3/4}(\log n)^{1/2}(\log\log n)^{1/4}]$.

Denote

$$W_{n,1}(s) = [F_{n,1}(s) - F_{n,1}(q_{1,t_0}(\tau_1))] - [F_1(s) - F_1(q_{1,t_0}(\tau_1))].$$

Lemma A4.2. *Let $\{a_n\}_{n=1}^\infty$ be a sequence of positive constants such that $a_n = O(n^{-1/2}(\log n)^{1/2}(\log\log n)^{1/4})$. Define $\mathcal{B}_{\tau_1} = [q_{1,t_0}(\tau_1) - a_n, q_{1,t_0}(\tau_1) + a_n]$. Then under the assumptions of Lemma A4.1, we have*

$$\sup_{s\in\mathcal{B}_{\tau_1}} |W_{n,1}(s)| = O_p[n^{-3/4}(\log n)^{1/2}(\log\log n)^{1/4}].$$

almost surely as $n \to \infty$.

Proof of Lemma A4.2. The proof of Lemma A4.2 closely resembles the lines in the proof of Lemma 1 in Bahadur (1966) with $F_{n,1}$ and F_1 involved only by utilizing the

fact that $F_{n,1}$ and F_1 are non-decreasing in s, thus are omitted. □

Let $\{\tau_{k_{n_{t_0}}}\}_{n_{t_0}=1}^{\infty}$ denote a positive sequence such that $0 < \tau_{k_{n_{t_0}}} \leq 1$ for each $n_{t_0} = \sum_{l=1}^{n} I(l \in \{i; M_i(t_0) = 1, \delta_i = 1\})$ and

$$\tau_{k_{n_{t_0}}} = \tau_1 + o(n^{-1/2}(\log n)^{1/2}(\log\log n)^{1/4}))$$

For each n_{t_0}, let $U_{n_{t_0},\tau_1}, \ldots, U_{n_{t_0},\tau_{n_{t_0}}}$ be the values of $\{X_i^r(t_0)\}_{\{i;M_i(t_0)=1,\delta_i=1\}}$ sorted in ascending order, and let

$$V_{n_{t_0}}(\tau_1) = U_{n_{t_0},\tau_{k_{n_{t_0}}}}.$$

Lemma A4.3. *Under the assumptions of Lemma A4.1, we have*

$$V_{n_{t_0}}(\tau_1) = q_{1,t_0}(\tau_1) + [\tau_{k_{n_{t_0}}} - F_{n,1}(q_{1,t_0}(\tau_1))]/f_1(q_{1,t_0}(\tau_1)) + O_p[n^{-3/4}(\log n)^{1/2}(\log\log n)^{1/4}].$$

almost surely as $n \to \infty$.

Proof. The proof of Lemma A4.3 closely resembles the lines in the proof of Lemma 3 in Bahadur (1966), thus are omitted. □

Thus we have shown that

$$\widehat{q}_{1,t_0}(\tau_1) = q_{1,t_0}(\tau_1) + \frac{\tau_1 - F_{n,1}(q_{1,t_0}(\tau_1))}{f_1(q_{1,t_0}(\tau_1))} + R_n(\tau_1),$$

where $R_n(\tau_1) = O_p[n^{-3/4}(\log n)^{1/2}(\log\log n)^{1/4}]$.

Next, since the first derivation of f_1 is bounded and $f_1(q_{1,t_0}(\tau_1)) > 0$, following similar lines as Kiefer (1967), we have

$$\limsup_{n\to\infty} \pm \left\{ \frac{\widehat{q}_{1,t_0}(\tau_1) - q_{1,t_0}(\tau_1) - [\tau_1 - F_{n,1}(q_{1,t_0}(\tau_1))]/f_1(q_{1,t_0}(\tau_1))}{n^{-3/4}(\log\log n)^{3/4}} \right\} = \frac{2^{5/4}}{3^{3/4}} \frac{\tau_1^{1/2}(1-\tau_1)^{1/2}}{f_1(q_{1,t_0}(\tau_1))}$$

almost surely for either choice of sign.

□

4.5.1.2 Lemma A4.4 and the proof

Define $\mathcal{F}_1 = \{\widehat{q}_{1,t_0}(\tau_1) : |\widehat{q}_{1,t_0}(\tau_1) - q_{1,t_0}(\tau_1)| \leq c_1 n^{-1/2}\}$ and $\mathcal{F}_{12} = \{(\widehat{q}_{1,t_0}(\tau_1), v) : |\widehat{q}_{1,t_0}(\tau_1) - q_{1,t_0}(\tau_1)| \leq c_{12} n^{-1/2}\}$. Let $\delta \geq \sup_{\mathcal{F}_1} \text{var}\{V(t_0)I(T^r \leq q) - V(t_0)I(T^r \leq q_{1,t_0}(\tau_1))\}$. Remarkly,

$$\begin{aligned} & \text{var}\{V(t_0)I(T^r \leq q) - V(t_0)I(T^r \leq q_{1,t_0}(\tau_1)) \mid M(t_0) = 1\} \\ \leq \;& c_1 n^{-1/2} \sup_{\mathcal{A}_2(t_0)} \sup_{s \in \mathcal{F}_1} f_{T^r|V(t_0),M(t_0)=1}(s) E\{[V(t_0)]^2\}. \end{aligned}$$

Lemma A4.4. *Given that Conditions 4.1 and 4.2 holds. For $\varepsilon > 0$ and $\alpha \geq 1$, we have*

$$\begin{aligned} & pr\{\sup_{\mathcal{F}_{12}} n^{1/2} |[F_{n,12}(\widehat{q}_{1,t_0}(\tau_1), v) - F_{n,12}(q_{1,t_0}(\tau_1), v)] - [F_{12}(\widehat{q}_{1,t_0}(\tau_1), v) - F_{12}(q_{1,t_0}(\tau_1), v)]| \geq 8\varepsilon\} \\ \leq \;& 192 n^2 \varepsilon^{-4} \exp\{-\varepsilon^2/(128\delta^2)\} + 256\delta^{-8} \exp(-n\delta^2), \end{aligned} \tag{4.7}$$

and

$$\begin{aligned} & pr\{\sup_{\mathcal{F}_1} n^{1/2} |[F_{n,1}(\widehat{q}_{1,t_0}(\tau_1)) - F_{n,1}(q_{1,t_0}(\tau_1))] - [F_1(\widehat{q}_{1,t_0}(\tau_1)) - F_1(q_{1,t_0}(\tau_1))]| \geq 8\varepsilon\} \\ \leq \;& 192 n^2 \varepsilon^{-4} \exp\{-\varepsilon^2/(128\delta^2)\} + 256\delta^{-8} \exp(-n\delta^2), \end{aligned} \tag{4.8}$$

Proof. This lemma can be derived as a refinement version of Theorem 37 in Chapter II of Pollard (1984), following the similar idea as in the proof of Theorem 3.11 in Zhu (1993). □

4.5.1.3 Lemma A4.5 and the proof

Lemma A4.5. *Given that Conditions 4.1 and 4.2 holds. Assume the density function f_1 and its first derivation with respect to τ are bounded away from zero and infinity*

on $\Delta = [\tau_L, \tau_U]$. *For any* $c > 0$ *and* $0 < \zeta \leq 1/2$, *there exists positive constant* c_1 *such that for sufficiently large* n,

$$pr\{\sup_{\tau\in\Delta} |\widehat{q}_{1,t_0}(\tau) - q_{1,t_0}(\tau)| > cn^{\zeta-1/2}\} \leq O\{\exp[-c_1 n^{2\zeta} - \log(n^{\zeta-1/2})]\}.$$

Proof. From the condition, we learn that the first derivative of f_1 with respect to τ are bounded away from zero and infinity on Δ. Then there exist $w_0 \geq 0$, such that $|q_{1,t_0}(\tau_a) - q_{1,t_0}(\tau_b)| \leq w_0|\tau_a - \tau_b|$. Given $0 < \delta < \min\{1, \frac{3w_0}{(\tau_U-\tau_L)}\}$, we can define a grid partition for $[\tau_L, \tau_U]$ as $\tau_L = \tau_0 < \tau_1 < \cdots < \tau_{N_\delta} = \tau_U$ with $N_\delta = \lfloor \frac{6w_0(\tau_U-\tau_L)}{\delta} \rfloor$, where $\lfloor x \rfloor$ denotes the largest integer smaller than x. With this partition, the grid would be $|\tau_k - \tau_{k-1}| \leq \frac{\delta}{3w_0}$. Then we have $|q_{1,t_0}(\tau_k) - q_{1,t_0}(\tau_{k-1})| \leq \frac{\delta}{3}$.

We learn from the definition for $\widehat{q}_{1,t_0}(\tau_k)$ and $q_{1,t_0}(\tau_k)$ that they are non-decreasing. Suppose $|\widehat{q}_{1,t_0}(\tau_k) - q_{1,t_0}(\tau_k)| \leq \frac{\delta}{3}$ and $|\widehat{q}_{1,t_0}(\tau_{k-1}) - q_{1,t_0}(\tau_{k-1})| \leq \frac{\delta}{3}$, then for $\tau_{k-1} \leq x \leq \tau_k$, we have

$$\widehat{q}_{1,t_0}(x) - q_{1,t_0}(x) \leq \widehat{q}_{1,t_0}(\tau_k) - q_{1,t_0}(\tau_k) + q_{1,t_0}(\tau_k) - q_{1,t_0}(x) \leq \frac{2\delta}{3} < \delta.$$

The other direction can be shown by the same arguments. Then we have $|\widehat{q}_{1,t_0}(x) - q_{1,t_0}(x)| < \delta$ for any $\tau_{k-1} \leq x \leq \tau_k$. Thus if $\sup_{\tau\in[\tau_L,\tau_U]} |\widehat{q}_{1,t_0}(\tau) - q_{1,t_0}(\tau)| > \delta$, there exist some $0 \leq k \leq N_\delta$ such that $|\widehat{q}_{1,t_0}(\tau_k) - q_{1,t_0}(\tau_k)| > \frac{\delta}{3}$. Let $\delta = cn^{\zeta-1/2}$. Then we have

$$\text{pr}\left(\sup_{\tau\in[\tau_L,\tau_U]} |\widehat{q}_{1,t_0}(\tau) - q_{1,t_0}(\tau)| > cn^{\zeta-1/2}\right) \leq N_\delta \text{pr}\left(|\widehat{q}_{1,t_0}(\tau_k) - q_{1,t_0}(\tau_k)| > \frac{c}{3}n^{\zeta-1/2}\right)$$

From Breslow and Crowley (1974), we have the consistency and weak convergency of $\widehat{G}^r_{C,t_0}(s)$ to $G^r_{C,t_0}(s)$. By WLLN and CLT, we have $\widehat{\pi}_{t_0} \to_p \pi_{t_0}$ and $\sqrt{n}(\widehat{\pi}_{t_0} - \pi_{t_0}) =$

$O_p(1)$. By CMT, we have $\sqrt{n}(\widehat{G}^r_{C,t_0}(s)\widehat{\pi}_{t_0} - G^r_{C,t_0}(s)\pi_{t_0}) = O_p(1)$. Write

$$\begin{aligned}
& L_n(\tau) = \frac{1}{n}\sum_{i=1}^{n}\left\{\tau - \left\{1 - \frac{I[X_i^r(t_0) > q_{1,t_0}(\tau), M_i(t_0)=1]}{\widehat{G}^r_{C,t_0}(q_{1,t_0}(\tau))\widehat{\pi}_{t_0}}\right\}\right\} \\
\approx\ & \frac{1}{n}\sum_{i=1}^{n}\left\{\tau - \left\{1 - \frac{I[X_i^r(t_0) > q_{1,t_0}(\tau), M_i(t_0)=1]}{G^r_{C,t_0}(q_{1,t_0}(\tau))\pi_{t_0}}\right\}\right\} \\
& -\frac{1-\tau}{G^r_{C,t_0}(q_{1,t_0}(\tau))}\left\{\widehat{G}^r_{C,t_0}(q_{1,t_0}(\tau)) - G^r_{C,t_0}(q_{1,t_0}(\tau))\right\} + \frac{1-\tau}{\pi_{t_0}}\left\{\widehat{\pi}_{t_0} - \pi_{t_0}\right\} \\
& +\frac{1}{n}\sum_{i=1}^{n}\left\{\frac{I[X_i^r(t_0) > q_{1,t_0}(\tau), M_i(t_0)=1]}{G^r_{C,t_0}(q_{1,t_0}(\tau))\pi_{t_0}} - (1-\tau)\right\}\left\{\frac{G^r_{C,t_0}(q_{1,t_0}(\tau))\pi_{t_0}}{\widehat{G}^r_{C,t_0}(q_{1,t_0}(\tau))\widehat{\pi}_{t_0}} - 1\right\} \\
\equiv\ & L_{n,1}(\tau) + L_{n,2}(\tau) + L_{n,3}(\tau) + o_p(n^{-1/2})
\end{aligned}$$

From Lemma A4.1,

$$\widehat{q}_{1,t_0}(\tau_k) - q_{1,t_0}(\tau_k) = \{f_1(q_{1,t_0}(\tau_k))\}^{-1}L_n(\tau_k) + R_n(\tau_k)$$

where $R_n(\tau_k) = O_p[n^{-3/4}(\log n)^{1/2}(\log\log n)^{1/4}]$. Then we have for sufficient large n,

$$\begin{aligned}
& \text{pr}\left(|\widehat{q}_{1,t_0}(\tau_k) - q_{1,t_0}(\tau_k)| > \frac{c}{3}n^{\zeta-1/2}\right) \\
\leq\ & \text{pr}\left(|L_n(\tau_k)| > \frac{\inf_{\tau\in[\tau_L,\tau_U]} f_1(q_{1,t_0}(\tau))c}{6}n^{\zeta-1/2}\right) \ (\text{Let } c^* = \frac{\inf_{\tau\in[\tau_L,\tau_U]} f_1(q_{1,t_0}(\tau))c}{6}) \\
\leq\ & \text{pr}\left(|L_{n,1}(\tau_k)| > \frac{c^*}{6}n^{\zeta-1/2}\right) + \text{pr}\left(|L_{n,2}(\tau_k)| > \frac{c^*}{6}n^{\zeta-1/2}\right) + \text{pr}\left(|L_{n,3}(\tau_k)| > \frac{c^*}{6}n^{\zeta-1/2}\right)
\end{aligned}$$

Since $M(t_0) \perp X^r(t_0)$, by Hoeffding's inequality, we have

$$\begin{aligned}
& \text{pr}\left(|L_{n,1}(\tau_k)| > \frac{c^*}{6}n^{\zeta-1/2}\right) \\
=\ & \text{pr}\left(\left|\left\{\frac{1}{n}\sum_{i=1}^{n}\frac{I[X_i^r(t_0) > q_{1,t_0}(\tau), M_i(t_0)=1]}{G^r_{C,t_0}(q_{1,t_0}(\tau))\pi_{t_0}}\right\} - (1-\tau_k)\right| > \frac{c^*}{6}n^{\zeta-1/2}\right) \\
=\ & O[\exp(-c_{1,1}n^{2\zeta})]
\end{aligned}$$

and

$$\begin{aligned}\mathrm{pr}\left(|L_{n,3}(\tau_k)| > \frac{c^*}{6}n^{\zeta-1/2}\right) &= \mathrm{pr}\left(\left|\left\{\frac{1}{n}\sum_{i=1}^{n} I(M_i(t_0)=1)\right\} - \pi_{t_0}\right| > \frac{c^*\pi_{t_0}}{6(1-\tau_k)}n^{\zeta-1/2}\right) \\ &= O[\exp(-c_{1,3}n^{2\zeta})]\end{aligned}$$

for some positive constant $c_{1,1}$ and $c_{1,3}$ as n is sufficiently large.

From Pepe (1991), we have

$$n^{1/2}|\widehat{G}^r_{C,t_0}(q_{1,t_0}(\tau))-G^r_{C,t_0}(q_{1,t_0}(\tau))| \approx n^{-1/2}\sum_{i=1}^{n} G^r_{C,t_0}(q_{1,t_0}(\tau))\int_0^{q_{1,t_0}(\tau)} y(s)^{-1}dM_i^{G^r_{C,t_0}}(s),$$

where $y(t) = pr(X^r \geq t)$ and $M_i^{G^r_{C,t_0}}(\cdot)$ is the martingale process. Then we have from Azuma–Hoeffding inequality

$$\begin{aligned}&\mathrm{pr}\left(|L_{n,2}(\tau_k)| > \frac{c^*}{6}n^{\zeta-1/2}\right) \\ =\ &\mathrm{pr}\left(\left|\widehat{G}^r_{C,t_0}(q_{1,t_0}(\tau)) - G^r_{C,t_0}(q_{1,t_0}(\tau))\right| > \frac{c^*G^r_{C,t_0}(q_{1,t_0}(\tau))}{6(1-\tau_k)}n^{\zeta-1/2}\right) \\ =\ &\mathrm{pr}\left(\left|\sum_{i=1}^{n} G^r_{C,t_0}(q_{1,t_0}(\tau))\int_0^{q_{1,t_0}(\tau)} y(s)^{-1}dM_i^{G^r_{C,t_0}}(s) - 0\right| > \frac{c^*G^r_{C,t_0}(q_{1,t_0}(\tau))}{12(1-\tau_k)}n^{\zeta-1/2}\right) \\ =\ &O[\exp(-c_{1,2}n^{2\zeta})]\end{aligned}$$

for some positive constant $c_{1,2}$ as n is sufficiently large.

These lead to

$$\begin{aligned}&\mathrm{pr}\left(\sup_{\tau\in[\tau_L,\tau_U]}|\widehat{q}_{1,t_0}(\tau) - q_{1,t_0}(\tau)| > cn^{\zeta-1/2}\right) \\ \leq\ &N_\delta O(\exp(-c_1n^{2\zeta})) = O\{\exp[-c_1n^{2\zeta} - \log(n^{\zeta-1/2})]\}.\end{aligned}$$

for a positive constant $c_1 = \min\{c_{1,1}, c_{1,2}, c_{1,3}\}$.

□

4.5.2 Proof of Theorem 4.1

We first establish the limit null distribution for $\frac{n^{1/2}\widehat{c}(\tau_1,v;t_0)}{\sqrt{\tau_1(1-\tau_1)F_{n,2}(v)(1-F_{n,2}(v))}}$.

Denote

$$W_{n,12}(\tau_1,v;t_0)=\frac{n^{1/2}\widehat{c}(\tau_1,v;t_0)}{\sqrt{\tau_1(1-\tau_1)F_{n,2}(v)(1-F_{n,2}(v))}}=n^{1/2}\frac{F_{n,12}(\widehat{q}_{1,t_0}(\tau_1),v)-F_{n,1}(\widehat{q}_{1,t_0}(\tau_1))\cdot F_{n,2}(v)}{\sqrt{\tau_1(1-\tau_1)F_{n,2}(v)(1-F_{n,2}(v))}}$$

and

$$W^F_{n,12}(\tau_1,v;t_0)=n^{1/2}\frac{F_{n,12}(\widehat{q}_{1,t_0}(\tau_1),v)-F_{n,1}(\widehat{q}_{1,t_0}(\tau_1))\cdot F_{n,2}(v)}{\sqrt{\tau_1(1-\tau_1)F_2(v)(1-F_2(v))}}.$$

We can write

$$\begin{aligned}
& W_{n,12}(\tau_1,v;t_0)-n^{1/2}\frac{F_{12}(q_{1,t_0}(\tau_1),v)-\tau_1F_2(v)}{\sqrt{\tau_1(1-\tau_1)F_2(v)(1-F_2(v))}}\\
= \;& W^F_{n,12}(\tau_1,v;t_0)-n^{1/2}\frac{F_{12}(q_{1,t_0}(\tau_1),v)-\tau_1F_2(v)}{\sqrt{\tau_1(1-\tau_1)F_2(v)(1-F_2(v))}}\\
& +\left[1-\frac{\sqrt{F_2(v)(1-F_2(v))}}{\sqrt{F_{n,2}(v)(1-F_{n,2}(v))}}\right]\cdot\left\{W^F_{n,12}(\tau_1,v;t_0)-\frac{F_{12}(q_{1,t_0}(\tau_1),v)-\tau_1F_2(v)}{\sqrt{\tau_1(1-\tau_1)F_2(v)(1-F_2(v))}}\right\}.
\end{aligned}$$

For $W^F_{n,12}(\tau_1,v;t_0)-n^{1/2}\frac{F_{12}(q_{1,t_0}(\tau_1),v)-\tau_1F_2(v)}{\sqrt{\tau_1(1-\tau_1)F_2(v)(1-F_2(v))}}$, we can express it as

$$W^F_{n,12}(\tau_1,v;t_0)-n^{1/2}\frac{F_{12}(q_{1,t_0}(\tau_1),v)-\tau_1F_2(v)}{\sqrt{\tau_1(1-\tau_1)F_2(v)(1-F_2(v))}}=\frac{I_1+I_2+I_3}{\sqrt{\tau_1(1-\tau_1)F_2(v)(1-F_2(v))}},$$

where

$$\begin{aligned}
I_1 &= n^{1/2}\{[F_{n,12}(\widehat{q}_{1,t_0}(\tau_1),v)-F_{n,12}(q_{1,t_0}(\tau_1),v)]-F_2(v)[F_{n,1}(\widehat{q}_{1,t_0}(\tau_1))-F_{n,1}(q_{1,t_0}(\tau_1))]\},\\
I_2 &= n^{1/2}\{[F_{n,1}(\widehat{q}_{1,t_0}(\tau_1))-F_{n,1}(q_{1,t_0}(\tau_1))][F_{n,2}(v)-F_2(v)]+[F_{n,2}(v)-F_2(v)][F_{n,1}(q_{1,t_0}(\tau_1))-\tau_1]\}\\
I_3 &= n^{1/2}\{[F_{n,12}(q_{1,t_0}(\tau_1),v)-F_2(v)F_{n,1}(q_{1,t_0}(\tau_1))-\tau_1F_{n,2}(v)+\tau_1F_2(v)]\\
&\quad -[F_{12}(q_{1,t_0}(\tau_1),v)-\tau_1F_2(v)]\}.
\end{aligned}$$

From the condition, we learn that the derivative of $f_1(q_{1,t_0}(\tau_1))$ is finite and bounded below by 0. By Taylor expansion, we have

$$\begin{aligned} I_1 &= n^{1/2}\{[F_{n,12}(\widehat{q}_{1,t_0}(\tau_1), v) - F_{n,12}(q_{1,t_0}(\tau_1), v)] - F_2(v)[F_{n,1}(\widehat{q}_{1,t_0}(\tau_1)) - F_{n,1}(q_{1,t_0}(\tau_1))]\} \\ &= n^{1/2}\{\{pr[T^r \leq q_{1,t_0}(\tau_1) \mid V(t_0) = v] - \tau_1\} f_1(q_{1,t_0}(\tau_1))(\widehat{q}_{1,t_0}(\tau_1) - q_{1,t_0}(\tau_1))\} \\ &\quad + o(n^{1/2}|\widehat{q}_{1,t_0}(\tau_1) - q_{1,t_0}(\tau_1)|). \end{aligned}$$

From Lemma A4.5, we have for sufficient large n,

$$pr\{\sup_{\tau \in \Delta} |\tau - F_{n,1}(q_{1,t_0}(\tau))| > cn^{\zeta - 1/2}\} \leq O\{\exp[-c_1 n^{2\zeta} - \log(n^{\zeta - 1/2})]\}.$$

By WLLN, we have

$$F_{n,2}(v) \to_p F_2(v).$$

Under the null hypothesis, we have $F_{12}(q_{1,t_0}(\tau_1), v) = \tau_1 F_2(v)$ for any $\tau_1 \in [\tau_L, \tau_U]$, which implies $pr[T^r \leq \widehat{q}_{1,t_0}(\tau_1) \mid V(t_0) = v] = \tau_1$. From these results and Lemma A4.1, Lemma A4.4 and Lemma A4.5, we have $I_1 = o_p(n^{-1/2})$ and $I_2 = o_p(n^{-1/2})$,

which leads to

$$
\begin{aligned}
W^F_{n,12}(\tau_1, v; t_0) &\approx n^{1/2}\{\frac{[F_{n,12}(q_{1,t_0}(\tau_1), v) - F_2(v)F_{n,1}(q_{1,t_0}(\tau_1)) - \tau_1 F_{n,2}(v) + \tau_1 F_2(v)]}{\sqrt{\tau_1(1-\tau_1)F_2(v)(1-F_2(v))}}\} \\
&= \frac{n^{1/2}}{\sqrt{\tau_1(1-\tau_1)F_2(v)(1-F_2(v))}} \cdot \\
&\quad \Bigg\{ -\frac{1}{n}\sum_{i=1}^{n} \frac{I(X_i^r(t_0) > q_{1,t_0}(\tau_1), V_i(t_0) \le v, M_i(t_0) = 1)}{\widehat{G}^r_{C,t_0}(q_{1,t_0}(\tau_1))\widehat{\pi}_{t_0}} \\
&\quad + F_2(v)[\frac{1}{n}\sum_{i=1}^{n} \frac{I(X_i^r(t_0) > q_{1,t_0}(\tau_1), M_i(t_0) = 1)}{\widehat{G}^r_{C,t_0}(q_{1,t_0}(\tau_1))\widehat{\pi}_{t_0}}] \\
&\quad + (1-\tau_1)\frac{1}{n}\sum_{i=1}^{n} \frac{I(V_i(t_0) \le v, M_i(t_0) = 1)}{\widehat{\pi}_{t_0}} - (1-\tau_1)F_2(v) \Bigg\} \\
&= \frac{n^{1/2} I(M_i(t_0) = 1)}{\widehat{\pi}_{t_0}\sqrt{\tau_1(1-\tau_1)F_2(v)(1-F_2(v))}} \cdot \\
&\quad \Bigg\{ -\frac{1}{\widehat{G}^r_{C,t_0}(q_{1,t_0}(\tau_1))}\frac{1}{n}\sum_{i=1}^{n} I[X_i^r(t_0) > q_{1,t_0}(\tau_1][I(V_i(t_0) \le v) - F_2(v)] \\
&\quad + (1-\tau_1)\frac{1}{n}\sum_{i=1}^{n}[I(V_i(t_0) \le v) - F_2(v)] \Bigg\}.
\end{aligned}
$$

Under the null hypothesis, we can write

$$
\begin{aligned}
W_{n,12}(\tau_1, v; t_0) &\approx W^F_{n,12}(\tau_1, v; t_0) \\
&\approx \frac{n^{1/2} I(M_i(t_0) = 1)}{\pi_{t_0}\sqrt{\tau_1(1-\tau_1)F_2(v)(1-F_2(v))}} \cdot \\
&\quad \Bigg\{ -\frac{1}{G^r_{C,t_0}(q_{1,t_0}(\tau_1))}\frac{1}{n}\sum_{i=1}^{n} I[X_i^r(t_0) > q_{1,t_0}(\tau_1][I(V_i(t_0) \le v) - F_2(v)] \\
&\quad + (1-\tau_1)\frac{1}{n}\sum_{i=1}^{n}[I(V_i(t_0) \le v) - F_2(v)] \Bigg\}. \\
&\equiv -n^{-1/2}\sum_{i=1}^{n} \frac{\xi_i(\tau_1, v; t_0)}{\sqrt{\tau_1(1-\tau_1)}}
\end{aligned}
$$

holds uniformly for any (τ_1, v), where

$$\begin{aligned}\xi_i(\tau_1, v; t_0) &= \frac{I(M_i(t_0)=1)}{\pi_{t_0}\sqrt{F_2(v)(1-F_2(v))}} \times \\ &\{I[X_i^r(t_0) > q_{1,t_0}(\tau_1)][I(V_i(t_0) \le v) - F_2(v)]/G^r_{C,t_0}(q_{1,t_0}(\tau_1)) \\ &-(1-\tau_1)[I(V_i(t_0) \le v) - F_2(v)]\}.\end{aligned}$$

Define $\mathcal{F} = \{\frac{\xi_i(\tau_1,v;t_0)}{\sqrt{\tau_1(1-\tau_1)}}, \tau_1 \in \Delta, v \in \Omega_{V(t_0)}, t_0 \in \mathcal{T}\}$. The function class $\mathcal{F}$ is Donsker and thus Glivenko-Cantelli (van der Vaart et al., 1996) since the class of indicator functions is Donsker and τ_1, $F_2(v)$, $1/\pi_{t_0}$ and $1/G^r_{C,t_0}(q_{1,t_0}(\tau_1))$ are uniformly bounded. As a result of Donsker theorem,

$$W_{n,12}(\tau_1, v; t_0) \to \chi(\tau_1, v; t_0),$$

where $\chi(\tau_1, v; t_0)$ for any $t_0 \in \mathcal{T}$ is a separable Gaussian process depending on (τ_1, v) for $(\tau_1, v) \in \Delta \otimes \Omega_{V(t_0)}$ with $E\{\chi(\tau_1, v; t_0)\} = 0$ and covariance matrix

$$E\{\chi(\tau_1, v; t_0)\chi(\tau_1', v'; t_0)\} = \frac{\{\min(\tau_1, \tau_1') - \tau_1\tau_1'\}\{\min(F_2(v), F_2(v')) - F_2(v)F_2(v')\}}{\tau_1(1-\tau_1)F_2(v)[1-F_2(v)]\tau_1'(1-\tau_1')F_2(v')[1-F_2(v')]}.$$

Then, by the extended continuous mapping theorem (Theorem 1.11.1 in van der Vaart et al. (1996)), under the null hypothesis, we have

$$\int_\Delta \int_{\Omega_{V(t_0)}} \frac{nc^2(\tau_1, v; t_0)}{\tau_1(1-\tau_1)F_{n,2}(v; t_0)(1-F_{n,2}(v; t_0))} d\mu_1(\tau_1)d\mu_2(v) \to_d \int_\Delta \int_{\Omega_{V(t_0)}} \chi^2(\tau_1, v; t_0)d\mu_1(\tau_1)d\mu_2(v).$$

Since $\mathcal{T}$ is a finite set, by continuous mapping theorem, we establish the limiting null distribution as

$$\begin{aligned}n\widehat{q}_{max}(T, \bar{Z}; \Delta, \mathcal{T}) &\to_d \max_{t_0 \in \mathcal{T}} \int_\Delta \int_{\Omega_{V(t_0)}} \chi^2(\tau_1, v; t_0)d\mu_1(\tau_1)d\mu_2(v); \\ n\widehat{q}_{sum}(T, \bar{Z}; \Delta, \mathcal{T}) &\to_d \sum_{t_0 \in \mathcal{T}} \int_\Delta \int_{\Omega_{V(t_0)}} \chi^2(\tau_1, v; t_0)d\mu_1(\tau_1)d\mu_2(v).\end{aligned}$$

This completes the proof of Theorem 4.1.

4.5.3 Proof of Theorem 4.2

We first investigate the asymptotic limit of $nq_{max}(T, \bar{Z}; \Delta, \mathcal{T})$ under the alternative hypothesis $H_{a,max}$. Let $\gamma_0 = q_{max}(T, \bar{Z}; \Delta, \mathcal{T})$. Under the alternative hypothesis $H_{a,max}$, we have $\gamma_0 > 0$. Since $an^{-1} = o_p(1)$, we have $P(an^{-1} > \gamma_0/2) \to 0$ as $n \to \infty$. Under the alternative hypothesis $H_{a,max}$, we have $\gamma_0 > 0$ and thus

$$P(\widehat{q}_{max}(T, \bar{Z}; \Delta, \mathcal{T}) > \gamma_0/2) \to P(\gamma_0 > \gamma_0/2) = 1.$$

This suggests that for any a,

$$\begin{aligned} & P(n\widehat{q}_{max}(T, \bar{Z}; \Delta, \mathcal{T}) > a) = P(\widehat{q}_{max}(T, \bar{Z}; \Delta, \mathcal{T}) > an^{-1}) \\ \geq \; & P(\widehat{q}_{max}(T, \bar{Z}; \Delta, \mathcal{T}) > \gamma_0/2) - P(an^{-1} > \gamma_0/2) \end{aligned}$$

It then follows that $P(n\widehat{q}_{max}(T, \bar{Z}; \Delta, \mathcal{T}) > a) \to 1$ as $n \to \infty$ under $H_{a,max}$. Denote $C_{\max,\alpha}$ as the α-level critical value determined upon the limit null distribution of $n\widehat{q}_{max}(T, \bar{Z}; \Delta, \mathcal{T})$, which is greater than 0. Then we have

$$P(n\widehat{q}_{max}(T, \bar{Z}; \Delta, \mathcal{T}) > C_{\max,\alpha}) \to 1$$

as $n \to \infty$ given $H_{a,max}$ holds. This implies that $n\widehat{q}_{max}(T, \bar{Z}; \Delta, \mathcal{T})$ is a consistent test against $H_{a,max}$.

Follow similar lines, we can show that $n\widehat{q}_{sum}(T, \bar{Z}; \Delta, \mathcal{T})$ is a consistent test against $H_{a,sum}$. This completes the proof of Theorem 4.2.

4.5.4 Justification of the Resampling Procedure

Firstly, as justified in the Proof of Theorem 4.1, there is weak convergence of $W_{n,12}$ to a mean zero Gaussian process with covariance process

$$E\{\chi(\tau_1, v; t_0)\chi(\tau_1', v'; t_0)\} = \frac{\{\min(\tau_1, \tau_1') - \tau_1\tau_1'\}\{\min(F_2(v), F_2(v')) - F_2(v)F_2(v')\}}{\tau_1(1-\tau_1)F_2(v)[1-F_2(v)]\tau_1'(1-\tau_1')F_2(v')[1-F_2(v')]}.$$

Next, given that $\{\iota_i^b\}_{i=1}^n$ are i.i.d. random variables following a standard normal distribution. Conditional on the observed data $\{(X_i^r(t_0), V_i(t_0), \delta_i, M_i(t_0))\}_{i=1}^n$, we have the asymptotic covariance matrix of $\frac{n^{-1/2}\sum_{l=1}^n \widehat{\xi}_i(\tau_1, v; t_0)\iota_i^b}{\sqrt{\tau_1(1-\tau_1)}}$ as

$$\begin{aligned} & E\left\{ \frac{n^{-1/2}\sum_{l=1}^n \widehat{\xi}_i(\tau_1', v'; t_0)\iota_i^b}{\sqrt{\tau_1'(1-\tau_1')}} \cdot \frac{n^{-1/2}\sum_{l=1}^n \widehat{\xi}_i(\tau_1, v; t_0)\iota_i^b}{\sqrt{\tau_1(1-\tau_1)}} \middle| \{(X_i^r(t_0), V_i(t_0), \delta_i, M_i(t_0))\}_{i=1}^n \right\} \\ = & \frac{n^{-1}\sum_{i=1}^n \widehat{\boldsymbol{\xi}}_i^{(j)}(\tau')\widehat{\boldsymbol{\xi}}_i^{(j)}(\tau)}{\tau_1'(1-\tau_1')\tau_1(1-\tau_1)}, \end{aligned}$$

which converges in probability to $E\{\chi(\tau_1, v; t_0)\chi(\tau_1', v'; t_0)\}$.

Following the arguments in Lin et al. (1993), conditional on $\{(X_i^r(t_0), V_i(t_0), \delta_i, M_i(t_0))\}_{i=1}^n$, $\frac{n^{-1/2}\sum_{l=1}^n \widehat{\xi}_i(\tau_1, v; t_0)\iota_i^b}{\sqrt{\tau_1(1-\tau_1)}}$ is zero-mean Gaussian with covariance function converge to the same limit as $W_{n,12}$. Applying the extended continuous mapping theorem as in the proof of Theorem 4.1, we have that under the null hypothesis, the conditional distribution of $n\widehat{q}_{max,b}(T, \bar{Z}; \Delta, \mathcal{T}$ (or $n\widehat{q}_{sum,b}(T, \bar{Z}; \Delta, \mathcal{T})$ given the observed data is asymptotically equivalent to the unconditional distributions of $nq_{max}(T, \bar{Z}; \Delta, \mathcal{T}$ (or $nq_{sum}(T, \bar{Z}; \Delta, \mathcal{T})$. This completes the justification for using the resampling procedure in Section 4.1.3.

Chapter 5

Summary and Future Work

5.1 Summary

In this dissertation, we propose to assess the covariate effects by adopting a global view pertaining to the concept of interval quantile independence. In the first topic, we introduce a flexible and robust method for evaluating the effects of a univariate covariate on a time-to-event outcome. In the second topic, we develop a group testing framework to identify variables that are dynamically associated with the outcome, and introduces a screening framework for excluding irrelevant variables in ultra-high dimensional settings with high confidence. In the third topic, we suggest an index for measuring and testing the departure from interval quantile independence for time-dependent covariates and time-to-event outcomes across a set of landmark times.

In the first project, we develop a new testing framework for evaluating a survival prognostic factor. The main thrust of the new framework lies in its flexibility of accommodating a dynamic covariate effect, which is achieved through adapting the spirit of globally concerned quantile regression. Our testing procedures are conveniently developed based on existing results on fitting a working quantile regression model with randomly censored data. It is important to note that the validity of the testing procedures does not require that the working model is the true model. Moreover, the proposed methods can be readily extended to handle more complex survival outcomes, such as time to event subject to competing risks. Another commendable extension of this work is to generalize the current null hypothesis and testing procedures to permit evaluating multiple prognostic factors simultaneously. This work also lays a key foundation for developing a nonparametric screening method for helping identify useful prognostic factors among a large number of candidates.

In the second project, we develop a new testing and screening framework that can help determine outcome-relevant covariates in classic univariate and multivariate settings and ultra-high dimensional settings. The proposed methods sensibly adopt a global perspective that examines covariate effects over a continuum of outcome

quantiles. Without assuming the correct specification of the working linear quantile regression model, we establish the theoretical properties the proposed testing and screening procedures. We demonstrate the favorable performance of the proposed methods with extensive simulation studies and an application to a microarray data set.

In the third project, we introduce a new test statistics for evaluating the covariate effect of time dependent covariate on the time-to-event outcome. We develop a new testing framework, and suggest a resampling procedure accordingly. The proposed method is fully nonparametric and sensibly adopt a global perspective. We provide rigorous justifications for the limiting null distribution, the consistency of the proposed testing procedure under a general class of alternative hypotheses and the proposed resampling procedure. The proposed methods are illustrated with extensive simulation studies and an application to the data set from the FIRST study.

5.2 Future work

In the third project, we proposed a new global-concerned test statistic that is powerful to assess the dynamic effects of time dependent covariates on the time-to-event outcome. However, this is a univariate approach that does not allow for adjusting for potential confounders. When the confounding effect can be captured by a small number of strata, we can consider evaluating the covariate effect within each stratum. This is not appealing because the sample size decreases with stratification. Also, it is preferable to provide an omnibus evaluation for the entire population. In our future work, we plan to develop a robust strategy to evaluate the effect of a time-dependent covariate while properly adjusting for the potential confounders. A possible direction is to adapt the idea of propensity score. We will explore and investigate possible solutions that can broaden the scope of the work in the third project.

Bibliography

Andersen, P. K. and Gill, R. D. (1982), 'Cox's regression model for counting processes: a large sample study', The Annals of Statistics **10**(4), 1100–1120.

Andersen, P. K. and Perme, M. P. (2010), 'Pseudo-observations in survival analysis', Statistical methods in medical research **19**(1), 71–99.

Angrist, J., Chernozhukov, V. and Fernández-Val, I. (2006), 'Quantile regression under misspecification, with an application to the us wage structure', Econometrica **74**(2), 539–563.

Austin, P. C. (2011), 'An introduction to propensity score methods for reducing the effects of confounding in observational studies', Multivariate behavioral research **46**(3), 399–424.

Bahadur, R. R. (1966), 'A note on quantiles in large samples', The Annals of Mathematical Statistics **37**(3), 577–580.

Barut, E., Fan, J. and Verhasselt, A. (2016), 'Conditional sure independence screening', Journal of the American Statistical Association **111**(515), 1266–1277.

Bellera, C. A., MacGrogan, G., Debled, M., de Lara, C. T., Brouste, V. and Mathoulin-Pélissier, S. (2010), 'Variables with time-varying effects and the cox model: some statistical concepts illustrated with a prognostic factor study in breast cancer', BMC medical research methodology **10**(1), 1–12.

Belloni, A. and Chernozhukov, V. (2011), 'L1-penalized quantile regression in high-dimensional sparse models', The Annals of Statistics **39**(1), 82–130.

Bilias, Y., Chen, S. and Ying, Z. (2000), 'Simple resampling methods for censored regression quantiles', Journal of Econometrics **99**(2), 373–386.

Boos, D. D., Stefanski, L. A. et al. (2013), Essential statistical inference, Springer.

Bottai, M., Cai, B. and McKeown, R. E. (2010), 'Logistic quantile regression for bounded outcomes', Statistics in medicine **29**(2), 309–317.

Boucheron, S., Lugosi, G. and Massart, P. (2013), Concentration inequalities: A nonasymptotic theory of independence, Oxford university press.

Breslow, N. and Crowley, J. (1974), 'A large sample study of the life table and product limit estimates under random censorship', The Annals of statistics pp. 437–453.

Cárcamo, J., Rodríguez, L.-A. and Cuevas, A. (2019), 'Directional differentiability for supremum-type functionals: statistical applications', arXiv preprint arXiv:1902.01136 .

Cheng, M.-Y., Honda, T., Li, J., Peng, H. et al. (2014), 'Nonparametric independence screening and structure identification for ultra-high dimensional longitudinal data', Annals of Statistics **42**(5), 1819–1849.

Cox, D. R. and Oakes, D. (2018), Analysis of survival data, Chapman and Hall/CRC.

Cui, Y. and Peng, L. (2022), 'Assessing dynamic covariate effects with survival data', Lifetime Data Analysis pp. 1–25.

Darling, D. A. (1957), 'The kolmogorov-smirnov, cramer-von mises tests', The Annals of Mathematical Statistics **28**(4), 823–838.

Delgado, M. A. (1993), 'Testing the equality of nonparametric regression curves', Statistics & probability letters **17**(3), 199–204.

Dette, H., Neumeyer, N. et al. (2001), 'Nonparametric analysis of covariance', the Annals of Statistics **29**(5), 1361–1400.

Dickson, E. R., Grambsch, P. M., Fleming, T. R., Fisher, L. D. and Langworthy, A. (1989), 'Prognosis in primary biliary cirrhosis: model for decision making', Hepatology **10**(1), 1–7.

Efron, B. and Tibshirani, R. (2007), 'On testing the significance of sets of genes', The annals of applied statistics **1**(1), 107–129.

Fan, J., Feng, Y. and Song, R. (2011), 'Nonparametric independence screening in sparse ultra-high-dimensional additive models', Journal of the American Statistical Association **106**(494), 544–557.

Fan, J. and Lv, J. (2008), 'Sure independence screening for ultrahigh dimensional feature space', Journal of the Royal Statistical Society: Series B (Statistical Methodology) **70**(5), 849–911.

Fan, J., Ma, Y. and Dai, W. (2014), 'Nonparametric independence screening in sparse ultra-high-dimensional varying coefficient models', Journal of the American Statistical Association **109**(507), 1270–1284.

Fan, J., Samworth, R. and Wu, Y. (2009), 'Ultrahigh dimensional feature selection: beyond the linear model', The Journal of Machine Learning Research **10**, 2013–2038.

Fan, J., Song, R. et al. (2010), 'Sure independence screening in generalized linear models with np-dimensionality', The Annals of Statistics **38**(6), 3567–3604.

Feng, X., He, X. and Hu, J. (2011), 'Wild bootstrap for quantile regression', Biometrika **98**(4), 995–999.

Grambsch, P. M. and Therneau, T. M. (1994), 'Proportional hazards tests and diagnostics based on weighted residuals', Biometrika **81**(3), 515–526.

Gutenbrunner, C., Jurečková, J., Koenker, R. and Portnoy, S. (1993a), 'Tests of linear hypotheses based on regression rank scores', Journaltitle of Nonparametric Statistics **2**(4), 307–331.

Gutenbrunner, C., Jurečková, J., Koenker, R. and Portnoy, S. (1993b), 'Tests of linear hypotheses based on regression rank scores', Journaltitle of Nonparametric Statistics **2**(4), 307–331.

Hall, P. and Hart, J. D. (1990), 'Bootstrap test for difference between means in nonparametric regression', Journal of the American Statistical Association **85**(412), 1039–1049.

Hall, P., Huber, C. and Speckman, P. L. (1997), 'Covariate-matched one-sided tests for the difference between functional means', Journal of the American Statistical Association **92**(439), 1074–1083.

Hall, P. and Miller, H. (2009), 'Using generalized correlation to effect variable selection in very high dimensional problems', Journal of Computational and Graphical Statistics **18**(3), 533–550.

Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. and Stahel, W. A. (1986), Robust statistics: the approach based on influence functions, Wiley-Interscience; New York.

Hardle, W. and Marron, J. S. (1990), 'Semiparametric comparison of regression curves', The Annals of Statistics pp. 63–89.

He, X., Wang, L., Hong, H. G. et al. (2013), 'Quantile-adaptive model-free variable screening for high-dimensional heterogeneous data', The Annals of Statistics **41**(1), 342–369.

Hjort, N. L. and Pollard, D. (2011), 'Asymptotics for minimisers of convex processes', arXiv preprint arXiv:1107.3806 .

Huang, J., Ma, S. and Zhang, C.-H. (2008), 'Adaptive lasso for sparse high-dimensional regression models', Statistica Sinica pp. 1603–1618.

Huang, Y. (2010), 'Quantile calculus and censored regression', The Annals of Statistics **38**(3), 1607–1637.

Huber, P. J. (2004), Robust statistics, Vol. 523, John Wiley & Sons.

Kasy, M. (2018), 'Uniformity and the delta method', Journal of Econometric Methods **8**(1).

Kiefer, J. (1967), 'On bahadur's representation of sample quantiles', The Annals of Mathematical Statistics **38**(5), 1323–1342.

Kim, Y., Choi, H. and Oh, H.-S. (2008), 'Smoothly clipped absolute deviation on high dimensions', Journal of the American Statistical Association **103**(484), 1665–1673.

King, E., Hart, J. D. and Wehrly, T. E. (1991), 'Testing the equality of two regression curves using linear smoothers', Statistics & Probability Letters **12**(3), 239–247.

Kleinbaum, D. G. and Klein, M. (2010), Survival analysis, Vol. 3, Springer.

Koenker, R. (2005), Quantile regression, Vol. 38, Cambridge Univ Pr.

Koenker, R. (2022), 'quantreg: Quantile regression'. R package version 5.87.
URL: *https://CRAN.R-project.org/package=quantreg*

Koenker, R. and Bassett Jr, G. (1978), 'Regression quantiles', Econometrica: journal of the Econometric Society pp. 33–50.

Koenker, R. and Bassett Jr, G. (1982), 'Robust tests for heteroscedasticity based on regression quantiles', Econometrica: Journal of the Econometric Society pp. 43–61.

Koenker, R., Chernozhukov, V., He, X. and Peng, L. (2017), Handbook of quantile regression, CRC press.

Kosorok, M. R. (2006), 'Introduction to empirical processes and semiparametric inference'.

Koul, H. L. and Schick, A. (1997), 'Testing for the equality of two nonparametric regression curves', Journal of Statistical Planning and Inference **65**(2), 293–314.

Koul, H. L. and Schick, A. (2003), 'Testing for superiority among two regression curves', Journal of statistical planning and inference **117**(1), 15–33.

Kuitunen, I., Ponkilainen, V. T., Uimonen, M. M., Eskelinen, A. and Reito, A. (2021), 'Testing the proportional hazards assumption in cox regression and dealing with possible non-proportionality in total joint arthroplasty research: methodological perspectives and review', BMC musculoskeletal disorders **22**(1), 1–7.

Kulasekera, K. (1995), 'Comparison of regression curves using quasi-residuals', Journal of the American Statistical Association **90**(431), 1085–1093.

Kutner, N. G., Clow, P. W., Zhang, R. and Aviles, X. (2002), 'Association of fish intake and survival in a cohort of incident dialysis patients', American journal of kidney diseases **39**(5), 1018–1024.

Ledoux, M. and Talagrand, M. (2013), Probability in Banach Spaces: isoperimetry and processes, Springer Science & Business Media.

Li, G., Peng, H., Zhang, J., Zhu, L. et al. (2012), 'Robust rank correlation based screening', The Annals of Statistics **40**(3), 1846–1877.

Li, J., Zheng, Q., Peng, L. and Huang, Z. (2016), 'Survival impact index and ultrahigh-dimensional model-free screening with survival outcomes', Biometrics **72**(4), 1145–1154.

Li, R. and Peng, L. (2014), 'Varying coefficient subdistribution regression for left-truncated semi-competing risks data', Journal of multivariate analysis **131**, 65–78.

Li, R. and Peng, L. (2017), 'Assessing quantile prediction with censored quantile regression models', Biometrics **73**(2), 517–528.

Li, R., Zhong, W. and Zhu, L. (2012), 'Feature screening via distance correlation learning', Journal of the American Statistical Association **107**(499), 1129–1139.

Lin, D. Y., Wei, L.-J. and Ying, Z. (1993), 'Checking the cox model with cumulative sums of martingale-based residuals', Biometrika **80**(3), 557–572.

Liu, J., Li, R. and Wu, R. (2014), 'Feature selection for varying coefficient models with ultrahigh-dimensional covariates', Journal of the American Statistical Association **109**(505), 266–274.

Mai, Q. and Zou, H. (2015), 'The fused kolmogorov filter: A nonparametric model-free screening method', The Annals of Statistics **43**(4), 1471–1497.

Mas-Colell, A. (1979), 'Homeomorphisms of compact, convex sets and the jacobian matrix', SIAM Journal on Mathematical Analysis **10**(6), 1105–1109.

Munk, A., Dette, H. et al. (1998), 'Nonparametric comparison of several regression functions: exact and asymptotic theory', The Annals of Statistics **26**(6), 2339–2368.

Newton, M. A., Quintana, F. A., Den Boon, J. A., Sengupta, S. and Ahlquist, P. (2007), 'Random-set methods identify distinct aspects of the enrichment signal in gene-set analysis', The Annals of Applied Statistics **1**(1), 85–106.

Oakes, D. (2000), 'Survival analysis', Journal of the American Statistical Association **95**(449), 282–285.

Park, T., Shao, X., Yao, S. et al. (2015), 'Partial martingale difference correlation', Electronic Journal of Statistics **9**(1), 1492–1517.

Peng, L. (2021), 'Quantile regression for survival data', Annual review of statistics and its application **8**, 413–437.

Peng, L. and Fine, J. (2008), 'Nonparametric tests for continuous covariate effects with multistate survival data', Biometrics **64**(4), 1080–1089.

Peng, L. and Fine, J. P. (2009), 'Competing risks quantile regression', Journal of the American Statistical Association **104**(488), 1440–1453.

Peng, L. and Huang, Y. (2008), 'Survival analysis with quantile regression models', Journal of the American Statistical Association **103**(482), 637–649.

Pepe, M. S. (1991), 'Inference for events with dependent risks in multiple endpoint studies', Journal of the American Statistical Association **86**(415), 770–778.

Pollard, D. (1984), Convergence of Stochastic Processes, Springer-Verlag.

Portnoy, S. (2003), 'Censored quantile regression', Journal of American Statistical Association **98**, 1001–1012.

Powell, J. L. (1984), 'Least absolute deviations estimation for the censored regression model', Journal of Econometrics **25**(3), 303–325.

Powell, J. L. (1986), 'Censored regression quantiles', Journal of econometrics **32**(1), 143–155.

Putter, H. (2013), Landmarking, in 'Handbook of Survival Analysis', Chapman & Hall/CRC, Boca Raton, chapter 21, pp. 441–456.

Rizopoulos, D. (2012), Joint Models for Longitudinal and Time-to-Event Data: With Applications in R, CRC Press.

Robins, J. M. and Rotnitzky, A. (1992), Recovery of information and adjustment for dependent censoring using surrogate markers, in 'AIDS epidemiology', Springer, pp. 297–331.

Robins, J. and Tsiatis, A. A. (1992), 'Semiparametric estimation of an accelerated failure time model with time-dependent covariates', Biometrika **79**(2), 311–319.

Scheetz, T. E., Kim, K.-Y. A., Swiderski, R. E., Philp, A. R., Braun, T. A., Knudtson, K. L., Dorrance, A. M., DiBona, G. F., Huang, J., Casavant, T. L. et al. (2006), 'Regulation of gene expression in the mammalian eye and its relevance to eye disease', Proceedings of the National Academy of Sciences **103**(39), 14429–14434.

Shao, X. and Zhang, J. (2014), 'Martingale difference correlation and its use in high-dimensional variable screening', Journal of the American Statistical Association **109**(507), 1302–1318.

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S. et al. (2005), 'Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles', Proceedings of the National Academy of Sciences **102**(43), 15545–15550.

Székely, G. J., Rizzo, M. L., Bakirov, N. K. et al. (2007), 'Measuring and testing dependence by correlation of distances', The annals of statistics **35**(6), 2769–2794.

Therneau, T. M. and Grambsch, P. M. (2000), Modeling Survival Data: Extending the Cox Model, Springer.

Thorogood, J., Persijn, G., Schreuder, G. M., D'amaro, J., Zantvoort, F., Van Houwelingen, J. and Van Rood, J. (1990), 'The effect of hla matching on kidney graft survival in separate posttransplantation intervals.', Transplantation **50**(1), 146–150.

van der Vaart, A., van der Vaart, A., van der Vaart, A. and Wellner, J. (1996), Weak Convergence and Empirical Processes: With Applications to Statistics, Springer Series in Statistics, Springer.
URL: *https://books.google.com/books?id=OCenCW9qmp4C*

Van Houwelingen, H. C. (2007), 'Dynamic prediction by landmarking in event history analysis', Scandinavian Journal of Statistics **34**(1), 70–85.

Venables, W. N. and Ripley, B. D. (2013), Modern applied statistics with S-PLUS, Springer Science & Business Media.

Verweij, P. J. and van Houwelingen, H. C. (1995), 'Time-dependent effects of fixed covariates in cox regression', Biometrics pp. 1550–1556.

Wald, A. (1943), 'Tests of statistical hypotheses concerning several parameters when the number of observations is large', Transactions of the American Mathematical society **54**(3), 426–482.

Wang, H. J. and Fygenson, M. (2009), 'Inference for censored quantile regression models in longitudinal studies', The Annals of Statistics **37**(2), 756–781.

Wang, H. J., McKeague, I. W. and Qian, M. (2018), 'Testing for marginal linear effects in quantile regression', Journal of the Royal Statistical Society. Series B, Statistical methodology **80**(2), 433.

Wang, H. and Wang, L. (2009), 'Locally weighted censored quantile regression', Journal of the American Statistical Association **104**(487), 1117–1128.

Wang, L. and Akritas, M. G. (2006), 'Testing for covariate effects in the fully nonparametric analysis of covariance model', Journal of the American Statistical Association **101**(474), 722–736.

Wang, L., Wu, Y. and Li, R. (2012), 'Quantile regression for analyzing heterogeneity in ultra-high dimension', Journal of the American Statistical Association **107**(497), 214–222.

Wellner, J. A. (2007), 'On an exponential bound for the kaplan–meier estimator', Lifetime data analysis **13**(4), 481–496.

Wu, Y. and Yin, G. (2015), 'Conditional quantile screening in ultrahigh-dimensional heterogeneous data', Biometrika **102**(1), 65–76.

Ying, Z., Jung, S. and Wei, L. (1995), 'Survival analysis with median regression models', Journal of the American Statistical Association **90**, 178–184.

Zhang, X., Yao, S., Shao, X. et al. (2018), 'Conditional mean and quantile dependence testing in high dimension', The Annals of Statistics **46**(1), 219–246.

Zhao, L. (2004), 'Linear hypothesis testing in censored regression models', Statistica Sinica pp. 333–347.

Zheng, Q., Peng, L. and He, X. (2015), 'Globally adaptive quantile regression with ultra-high dimensional data', Annals of statistics **43**(5), 2225.

Zhou, L. (2006), 'A simple censored median regression estimator', Statistica Sinica **16**(3), 1043–1058.

Zhu, L.-P., Li, L., Li, R. and Zhu, L.-X. (2011), 'Model-free feature screening for ultrahigh-dimensional data', Journal of the American Statistical Association **106**(496), 1464–1475.

Zhu, L.-X. (1993), 'Convergence rates of the empirical processes indexed by the classes of functions and applications', Journal of Systems Science and Mathematical Sciences **13**(1), 33.

Zhu, L., Zhang, Y., Xu, K. et al. (2018), 'Measuring and testing for interval quantile dependence', The Annals of Statistics **46**(6A), 2683–2710.

Zou, H. (2006), 'The adaptive lasso and its oracle properties', Journal of the American statistical association **101**(476), 1418–1429.