Composite Quantile Regression for Varying Coefficient Models with Response Data Missing at Random

Shuanghua Luo 1, Cheng-yi Zhang 2,* and Meihua Wang 3

1 School of Science, Xi’an Polytechnic University, Xi’an 710048, China
2 School of Economics and Finance, Xi’an Jiaotong University, Xi’an 710061, China
3 School of Economics and Management, Xidian University, Xi’an 710071, China

* Correspondence: cyzhang08@xjtu.edu.cn; Tel.: +86-180-9219-2208

Received: 24 July 2019; Accepted: 19 August 2019; Published: 21 August 2019

Abstract: Composite quantile regression (CQR) estimation and inference are studied for varying coefficient models with response data missing at random. Three estimators including the weighted local linear CQR (WLLCQR) estimator, the nonparametric WLLCQR (NWLLCQR) estimator, and the imputed WLLCQR (IWLLCQR) estimator are proposed for unknown coefficient functions. Under some mild conditions, the proposed estimators are asymptotic normal. Simulation studies demonstrate that the unknown coefficient estimators with IWLLCQR are superior to the other two with WLLCQR and NWLLCQR. Moreover, bootstrap test procedures based on the IWLLCQR fittings is developed to test whether the coefficient functions are actually varying. Finally, a type of investigated real-life data is analyzed to illustrated the applications of the proposed method.

Keywords: varying coefficient model; composite quantile regression; missing at random; inverse probability weighting; imputed method

1. Introduction

The varying coefficient model, proposed originally by Hastie and Tibshirani [1], is flexible and powerful to examine the dynamic changes of regression coefficients over some factors such as time and age and has gained much popularity during the past few decades (see [2–6]).

A classical varying coefficient model has the following structure:

\[ Y = X^T \beta(U) + \epsilon, \]  

where \( Y \in \mathbb{R} \) is a response variable, \( X = (X_1, \cdots, X_p)^T \in \mathbb{R}^p \) is a covariate vector, \( \beta(\cdot) = (\beta_1(\cdot), \cdots, \beta_p(\cdot))^T \in \mathbb{R}^p \) is an unknown coefficient vector function with a smoothing variable \( U \), and \( \epsilon \) is a random error independent of \( (X, U) \).

Recently, some estimates of \( \beta(\cdot) \) for Model (1) with the least squares regression have attracted many researchers’ attention. Above all, Hastie and Tibshirani [1] considered \( L_2 \) penalized least squares estimation and attained some good results. Following, Fan et al. [4] and Fan et al. [7] applied the least squares regression to propose a two-step local polynomial estimation procedure and a profile estimator for Model (1), respectively, and designed some suitable statistical inference procedures. However, there arises a dilemma that these estimation procedures of the least squares could be very sensitive to outliers [8]. In order to overcome this problem, quantile regression proposed by Koenker [9] can be thought of as an alternative because as a mean model, traditional least squares regression only gives the effects of the covariates at the center of the distribution, while quantile regression can not only directly estimate the ones at different quantiles, but also characterize the entire conditional distribution.
of a dependent variable of the regression [8]. Thus, this regression has a much better robust property when processing outlier observations.

Due to its significant theoretical advances, some scholars had integrated the quantile regression into the varying coefficient model. Kim [10] attained the quantile regression model with the varying coefficient. For the processing of time series data, Cai and Xu [11] developed nonparametric quantile estimations with dynamic smooth coefficient models. Later, Cai and Xiao [12] applied dynamic models with partially-varying coefficients to investigate semiparametric quantile regression and obtained some useful results. Tang [13] derived a robust quantile regression estimation using the spatial semiparametric partially-linear regression model with a varying coefficient. Unfortunately, a relative small efficiency may result by a single quantile regression procedure compared with the least squares regression. In order to overcome this drawback, it is very necessary to get a desirable efficient and stable estimator. In recent years, an oracle procedure of composite quantile regression (CQR) was proposed by Zou and Yuan [14] to select the significant variables, and some important theoretical and applied results were derived. So far, the CQR method is widely used in many situations. For example, some efficient estimators based on the CQR method were proposed by Kai et al. [15] and Guo et al. [16] for semi-parametric partially-linear models with a varying coefficient. In addition, a data-driven weighted CQR (WCQR) estimation was studied by Sun et al. [17] and Yang et al. [18] for linear models with a varying coefficient, respectively.

Although the QR has significant theoretical properties such that its literature on the complete data has been rapidly growing, people have paid scant attention to the incomplete data, i.e., the data samples containing missing values since this class of data may lead easily to substantially-distorted results. In fact, as is commonplace, missing data often appear in real life. There are various reasons such as failure on the part of investigators when gathering correct information, the unwillingness of some sampled units when supplying the desired information, loss of information caused by uncontrollable factors, and so forth, resulting in the data missing. In the early 1970s, the advances in computer technology such that many laborious numerical calculations were possible to perform spurred the literature on statistical analysis of real data containing missing values in applied work; see [19–25]. Despite a long history on missing data analysis, little work on QR has taken missing data into account. Recently, an iterative imputation procedure was developed by Wei et al. [26] in a linear QR model with non-i.i.d. error terms for the covariates with missing values. A smoothed empirical likelihood analysis was discussed by Lv and Li [27] for partially-linear quantile regression with missing response. An inverse probability weighting QR approach was proposed by Sherwood et al. [28] in the last few years for analyzing healthcare cost data with missing covariates at random. The QR for competing risk data was studied by Sun et al. [29] when the failure type was missing. An efficient QR analysis was discussed by Chen [30] with missing observations. Some imputation methods were proposed by Shu [31] for quantile estimation under data missing at random.

In this paper, a coherent inference framework based on CQR estimation and inference is explored for varying coefficient models with response data missing at random. The main contribution of this paper can be summarized as follows:

- A composite quantile regression estimation (CQRE) method is proposed for the analysis of varying coefficient models with response data missing at random. This method has the following two advantages: (1) the CQRE method can effectively overcome not only the drawback of a relative small efficiency that may result from a single quantile regression procedure compared with the least-squares regression, but also the interference of non-normal error; hence, it improves its estimation efficiency significantly; (2) since different quantiles are used in the imputation instead of actually observed responses or means and the robustness of quantile regression is inherited, the CQRE method is less sensitive to outliers; thus, the CQRE method is more effective and robust than the single quantile regression method and the classical least squares method.

- Three estimators including the weighted local linear CQR (WLLCQR) estimator, the nonparametric WLLCQR (NWLLCQR) estimator, and the imputed WLLCQR (IWLLCQR) estimator are proposed.
for an unknown coefficient function in the varying coefficient model to establish the asymptotic
normality of these estimators under some mild conditions.

The rest of this paper is organized as follows. The CQR varying coefficient model will be
introduced with missing response data in Section 2 to construct a class of estimators for an unknown
coefficient function. Then, some theoretical results on the asymptotic property of the proposed
estimators are proposed in Section 3. In Section 4, a bootstrap-based test procedure is developed
to perform a simulation study in Section 5 that demonstrates the finite-sample performance of the
proposed method. Following, an application to a real dataset illustrates the effectiveness of our
approach in Section 6. In addition, some discussions and conclusion remarks are presented in Section 7
and Section 8, respectively. Finally, the proof of the main results is given in Appendix A.

2. Estimation Based on the CQR Varying Coefficient Model With Missing Response

In this section, the CQR varying coefficient model will be introduced with missing response data
to construct a class of estimators for an unknown coefficient function. In particular, as the main estimate
methods in this paper, three estimators including the weighted local linear CQR (WLLCQR) estimator,
the nonparametric WLLCQR (NWLLCQR) estimator, and the imputed WLLCQR (IWLLCQR) estimator
are constructed and emphasized.

Let \(\{X_i, U_i, Y_i, \delta_i\} : i = 1, 2, \ldots, n\) be a random sample coming from Model (1), such that:

\[
Y_i = X_i^T \beta(U_i) + \varepsilon_i, \quad i = 1, 2, \ldots, n,
\]

where all the \(X_i \in \mathbb{R}^p\) and \(U_i \in \mathbb{R}\) are always observed, and \(\beta(\cdot) = (\beta_1(\cdot), \ldots, \beta_p(\cdot))^T \in \mathbb{R}^p\)
is the coefficient vector function. Further, \(\delta_i = 0\) if \(Y_i\) is missing and \(\delta_i = 1\) otherwise. We assume that
throughout this paper, \(Y_i\) is missing at random (MAR) for some \(i\). This assumption indicates that \(\delta_i\) and
\(Y_i\) are conditionally independent given \(X_i\) and \(U_i\), that is,

\[
P(\delta_i = 1 | X_i, U_i, Y_i) = P(\delta = 1 | X_i, U_i) = p(X_i, U_i) = \pi(Z_i),
\]

where \(Z_i = (X_i, U_i)^T\). Moreover, we also assume that across different quantile regression models,
there is the same coefficient vector function \(\beta(\cdot)\). Thus, we can express the conditional \(\tau\)-quantile function
of \(Y\) as:

\[
Q_\tau(X, U) = X^T \beta(U) + c_\tau,
\]

where \(c_\tau\) is the \(\tau\)-quantile of \(\varepsilon\). If \(\beta_j(\cdot)\) is differentiable, Taylor’s expansion yields that:

\[
\beta_j(U) \approx \beta_j(u) + \beta_j'(u)(U - u) = a_j + b_j(U - u),
\]

for \(j = 1, 2, \ldots, p\), where \(u\) is a fixed value of a random variable and \(U\) lies in a neighborhood of \(u\).
For the case of no missing response data, minimizing the following criterion:

\[
\sum_{i=1}^{n} \rho_\tau\left( Y_i - c_\tau - X_i^T [a + b(U_i - u)] \right) K_h(U_i - u), \quad (4)
\]

we can attain the local linear quantile regression (LLQR) estimator of \(\beta(u)\), where \(\rho_\tau(u) = u(\tau - I_{u<0})\)
is called the quantile loss function of \(\tau\)-quantile regression, \(a = (a_1, \ldots, a_p)^T\), \(b = (b_1, \ldots, b_p)^T\),
and \(K_h(\cdot) = K_h(\cdot/h)/h\) is a Gaussian kernel function with bandwidth \(h\). In order to improve the
quantile regression estimation efficiently, the local linear composite quantile regression (LLCQR) estimation
is adopted from Guo et al. [16] for the varying coefficient models. Let \(q\) be the number of quantiles
and \(\tau_k = k/(1 + q)\) for \(k = 1, \ldots, q\). The loss function of the LLCQR estimation is defined as:

\[
\sum_{k=1}^{q} \sum_{i=1}^{n} \rho_{\tau_k}\left( Y_i - c_k - X_i^T [a + b(U_i - u)] \right) K_h(U_i - u), \quad (5)
\]
where \(c_k\) is the \(\tau_k\)-quantile of \(\varepsilon\).

In what follows, this technique of LLCQR will be extended to handle the case of response data missing at random.

### 2.1. WLLCQR Estimation

The inverse probability weighting (IPW) version of local linear CQR estimation will be considered to handle missing responses data at random, that is the CC (complete-case) analysis will be adjusted by using the inverse of the selection probability as the weight. However, the nonparametric smoothing estimation of \(\pi(\cdot)\) will encounter the curse of dimensionality when the dimension of \(Z\) is high enough. Motivated by Wang [24], we use the inverse marginal probability weighted approach.

Let \(P(\delta = 1 | X = x, U_i = u) = P(\delta = 1 | U_i = u) = \Delta(u)\), i.e., the propensity score just depends on \(U_i\). When the inverse marginal probability function \(\Delta(u)\) is known, the WLLCQR estimator \(\hat{\beta}(u)\) of \(\beta(u)\) is defined as:

\[
(\hat{c}, \hat{a}, \hat{b}) = \arg \min_{c,a,b} \sum_{k=1}^{q} \sum_{i=1}^{n} \delta_i \frac{\delta_i}{\hat{\Lambda}(U_i)} \rho_{\hat{\tau}} \left( Y_i - c_k - X_i^T \left[ a + b(U_i - u) \right] \right) K_h(U_i - u),
\]

where \(c = (c_1, \cdots, c_q)^T\) and \(\Delta(u) = P(\delta = 1 | U_i = u).\) Here, \(\hat{\beta}(u) = \hat{\alpha}\) is called the WLLCQR estimator of \(\beta(u)\) with \(\Delta(u)\).

### 2.2. Nonparametric WLLCQR Estimation

However, the inverse marginal probability function in practical situations is usually unknown, and thus, it needs to be estimated. We often employ nonparametric smoothing estimation approaches to estimate the unknown selection probability \(\Delta(\cdot)\). The Nadaraya–Watson estimation [32] is one of these nonparametric smoothing estimation approaches. We can define the Nadaraya–Watson estimator of \(\Delta(u)\) as:

\[
\hat{\Delta}(u) = \frac{\sum_{i=1}^{n} \delta_i L_{h_0}(U_i - u)}{\sum_{i=0}^{n} L_{h_0}(U_i - u)},
\]

where \(L_{h_0}(\cdot) = L_{h_0}(\cdot | h_0)\) is a density kernel function and \(h_0\) is a bandwidth. Therefore, the NWLLCQR estimation procedure with \(\hat{\Delta}(u)\) is formally defined as:

\[
(\hat{c}_N, \hat{a}_N, \hat{b}_N) = \arg \min_{c,a,b} \sum_{k=1}^{q} \sum_{i=1}^{n} \delta_i \frac{\delta_i}{\hat{\Lambda}(U_i)} \rho_{\hat{\tau}} \left( Y_i - c_k - X_i^T \left[ a + b(U_i - u) \right] \right) K_h(U_i - u),
\]

where \(\hat{\beta}_N(u) = \hat{\alpha}_N\) is called the NWLLCQR estimator of \(\beta(u)\) with \(\hat{\Delta}(u)\).

### 2.3. Imputed WLLCQR Estimation

Although both the WLLCQR estimator and NWLLCQR estimator can well estimate the inverse marginal probability function, the information contained in the data is not explored fully. Now, we use quantile regression imputation to resolve the issue by imputing \(Y_i\) by \(X_i^T \hat{\beta}_C(U)\) if \(Y_i\) is missing, where \(\hat{\beta}_C(U) = \hat{\alpha}\) and \(\hat{\alpha}\) is defined in:

\[
(\hat{c}, \hat{a}, \hat{b}) = \arg \min_{c,a,b} \sum_{k=1}^{q} \sum_{i=1}^{n} \delta_i \rho_{\hat{\tau}} \left( Y_i - c_k - X_i^T \left[ a + b(U_i - u) \right] \right) K_h(U_i - u).
\]

Therefore, the imputed WLLCQR estimation procedure can be defined as:

\[
(\hat{c}_I, \hat{a}_I, \hat{b}_I) = \arg \min_{c,a,b} \sum_{k=1}^{q} \sum_{i=1}^{n} \rho_{\hat{\tau}} \left( Y_i^* - c_k - X_i^T \left[ a + b(U_i - u) \right] \right) K_h(U_i - u),
\]

where \(Y_i^* = \frac{\delta_i}{\hat{\Lambda}(U_i)} Y_i + (1 - \frac{\delta_i}{\hat{\Lambda}(U_i)}) X_i^T \hat{\beta}_C(U_i)\) and \(\hat{\beta}_I(u) = \hat{\alpha}_I\) is called the IWLLCQR estimator of \(\beta(u)\).
Remark 1. Since local results in interpolation vary greatly and are unstable, as a smoothing method, the kernel function is used in Equations (4)–(10) such that the interpolation results of these equations are much smoother and stabler.

3. Asymptotic Properties

In this section, the asymptotic distribution will be considered for the estimators proposed in Section 2 to establish some theoretical results of these estimators.

Let \( f(\cdot) \) and \( f_U(\cdot) \) be the density functions of \( \varepsilon \) and \( U \), respectively. For simplicity, the following notations: \( \mu_j = \int u^j K(u) du \) for \( j = 0, 1, 2 \), \( \eta_i = \sum_{k=1}^{q} I(\varepsilon_i \leq c_k) - \tau_k \) for \( i = 1, 2, \ldots, n \), and \( D_u = E(X_i X_i^T | U = u) \sum_{k=1}^{q} f(c_k) \) will be used in this section.

Now, the following results are established.

Theorem 1. Suppose that Conditions C1–C7 in the Appendix hold. If \( \Delta(u) \) is known, then:

\[
\sqrt{n} \left( \hat{\beta}(u) - \beta(z) - \frac{1}{2} h^2 \mu_2 \beta''(u) \right) \xrightarrow{d} N(0, \frac{\nu_0}{f_U(u)} D_u^{-1} \Omega_u D_u^{-1}),
\]

where \( \xrightarrow{d} \) represents the convergence in the distribution, \( \Omega_u = E\{ \eta^2 \frac{\pi(U)}{f(U)} X_i X_i^T | U = u \} - E\{ \frac{1-\Delta(U)}{\pi(U)} E\{ X_i \eta_i | U \} \} \).”

Theorem 2. Suppose \( \Delta(u) > 0 \) is a smoothing function of \( u \), based on Conditions C1–C7 in the Appendix holding. Then:

\[
\sqrt{n} \left( \hat{\beta}_E(u) - \beta(z) - \frac{1}{2} h^2 \mu_2 \beta''(u) \right) \xrightarrow{d} N(0, \frac{\nu_0}{f_U(u)} D_u^{-1} \Omega_u D_u^{-1}),
\]

where \( \Omega_u^* = E\{ \eta^2 \frac{\pi(U)}{f(U)} X_i X_i^T | U = u \} - E\{ (1 - \pi(U)) (2\pi(U) + 1) E\{ X_i \eta_i | U \} \} \).

Theorem 3. Assuming \( \Delta(u) > 0 \) is a smoothing function of \( u \), based on the Conditions C1–C7 in the Appendix hold. Then:

\[
\sqrt{n} \left( \hat{\beta}_I(u) - \beta(z) - \frac{1}{2} h^2 \mu_2 \beta''(u) \right) \xrightarrow{d} N(0, \frac{\nu_0}{f_U(u)} D_u^{-1} \Omega_u^* D_u^{-1}),
\]

where:

\[
\Omega_u^* = E\{ \eta^2 \frac{\pi(U)}{f(U)} X_i X_i^T | U = u \} - E\{ (1 - \pi(U)) (2\pi(U) + 1) E\{ X_i \eta_i | U \} \} \).
\]

4. A Bootstrap-Based Goodness-of-Fit Test

In investigating the varying coefficient model, how to test whether unknown coefficient functions are actually varying is of importance. In this section, the testing problem is considered for Model (1) under response missing, and then, a goodness-of-fit test is proposed based on the difference between the weighted residual sums of the quantile (WRSQ) and the LLCQR fittings under both the null and alternative hypotheses.

The following testing problem:

\[
H_0 : \beta(u) = \beta \quad \text{versus} \quad H_1 : \beta(u) \neq \beta,
\]

for simplicity, is considered, where \( \beta \) is a constant vector. The model (1) becomes a classical linear model with missing responses under the null hypothesis. The WRSQ under \( H_0 \) is defined as:

\[
WRSQ_0 = \sum_{k=1}^{q} \sum_{i=1}^{n} \rho_{\Delta i} \left( Y_i - \hat{\epsilon}_i - X_i^T \hat{\beta}_I \right),
\]
where \( \hat{c}_1, \ldots, \hat{c}_q \) and \( \hat{\beta}_I \) are given by the following IWCQR estimation procedure:

\[
(\hat{c}_1, \ldots, \hat{c}_q, \hat{\beta}_I) = \arg \min_{\xi} \sum_{k=1}^{q} \sum_{i=1}^{n} \rho_{\xi}(y_i - c_k - X_i^T \hat{\beta}).
\]

Similarly, the WRSQ under \( H_1 \) can be defined as:

\[
\text{WRSQ}_1 = \sum_{k=1}^{q} \sum_{i=1}^{n} \rho_{\xi}(y_i^* - \hat{c}_i - X_i^T \hat{\beta}_1(U_i)),
\]

where \( c_{1i}, \ldots, c_{qi} \) and \( \hat{\beta}_1(u) \) are given in (10). Then, the following test statistic is given as:

\[
T_\nu = \frac{\text{WRSQ}_0 - \text{WRSQ}_1}{\text{WRSQ}_0} = \frac{\text{WRSQ}_0}{\text{WRSQ}_1} - 1.
\]

For a large value of \( T_\nu \), the null hypothesis (11) is rejected. In what follows, based on the bootstrap method, we evaluate the \( p \) values of the test along the lines of Wong et al. [33] and Guo et al. [16]:

**Step 1.** Assume the number of complete data is \( m \). We get the IWLLCQR estimator \( \hat{\beta}_I(U_i) \).

**Step 2.** The bootstrap residuals \( \hat{c}_i^* \) are generated from series \( \{\hat{e}_i - \bar{\hat{e}}\}_{i=1}^{n} \), where:

\[
\hat{e}_i = Y_i^* - X_i^T \tilde{\hat{\beta}}(U_i), \bar{\hat{e}} = \frac{1}{m} \sum_{i=1}^{m} \hat{e}_i.
\]

**Step 3.** Step 2 is repeated for \( M \) times, and then, series sets \( E_j = \{Y_{ij}, X_{ij}, U_{ij}, \delta_{ij}\}_{j=1}^{n} \) are obtained for \( j = 1, \ldots, M \). The bootstrap test statistic is calculated for each bootstrap sample \( E_j \), denoted by \( T_\nu^{*} \).

**Step 4.** The \( p \) value is approximately estimated by \( \hat{p} = \frac{S}{M} \), where \( S \) is the cardinality of the set \( \mathcal{S} = \{j \mid T_\nu^{*} \geq T_\nu, j = 1, \ldots, M\} \).

### 5. Simulation Study

A simulation study was carried out to investigate the finite-sample properties of our proposed method by a comparison among the WLLCQR estimation method, the NWLLCQR estimation method, the IWLLCQR estimation method, the INWLLCQR (imputed not weighted LLCQR) estimation method, and the WLLCQR estimation method without data missing, defined in (5).

In numerical studies, generally, the kernel function \( K(x) \) is taken to be \( K(x) = 0.75(1 - x^2)_I(|x| \leq 1) \). Here, this function is still adopted. It follows that the cross-validation method is used to select the optimal bandwidths \( h_{opt} \). In the subsequent examples, let the composite level \( q = 8 \).

**Example 1.** Consider the following model:

\[
Y_i = \sin(\pi U_i)X_i + \epsilon_i,
\]

where \( X_i = Z_{1i} + Z_{2i} + Z_{3i}, U_i = Z_{1i} + Z_{2i}, Z_{1i}, Z_{2i}, Z_{3i} \) are independent, \( Z_{1i} \) and \( Z_{2i} \) follow a uniform distribution on \([-1, 1] \), \( Z_{3i} \sim N(0, 1) \), and three error distributions of \( \epsilon_i \) are considered including \( \epsilon_i \sim N(0, 1) \), \( \epsilon_i \sim \text{t}(3) \), and \( \epsilon_i \sim 0.8N(0, 1) + 0.2N(0, 3^2) \).

An analysis of the fitting of five different estimators including WLLCQR\(_A\), NWLLCQR\(_A\), IWLLCQR\(_A\), INWLLCQR, and WLLCQR is done by using the following three selection probability functions:

**Case 1:** \( \Delta_1(u) = 0.9 + 0.2u \) if \( |u| \leq 0.5 \), and 0.95 otherwise.

**Case 2:** \( \Delta_2(u) = 0.7 + 0.2u \) if \( |u| \leq 0.5 \), and 0.75 otherwise.
Case 3: $\Delta_3(u) = 0.5 + 0.2u$ if $|u| \leq 0.5$, and 0.55 otherwise.

The average missing rates of $Y$ corresponding to these three selection probability functions are approximately 0.15, 0.36, and 0.45, respectively. For each of the three cases, we generated 500 Monte Carlo random samples of size 200. The performance of the estimators is illustrated via the MSE. The simulation results are given in Table 1.

| Model Error | $\Delta(u)$ | $n$ | WLLCQR | NWLLCQR | IWLLCQR | INWLLCQR | WLLCQR |
|-------------|-------------|-----|--------|---------|---------|----------|--------|
| Error(1)    | $\Delta_1(u)$ | 200 | 0.01219 | 0.01213 | 0.01192 | 0.01201 | 0.01182 |
|             | 100         |     | 0.01017 | 0.01012 | 0.00997 | 0.01003 | 0.00973 |
|             | $\Delta_2(u)$ | 200 | 0.1845 | 0.1792 | 0.1701 | 0.1715 | 0.1645 |
|             | 100         |     | 0.1701 | 0.1698 | 0.1641 | 0.1654 | 0.1583 |
|             | $\Delta_3(u)$ | 200 | 0.2685 | 0.2518 | 0.2346 | 0.2371 | 0.2207 |
|             | 100         |     | 0.2685 | 0.2518 | 0.2346 | 0.2371 | 0.2207 |

From Table 1, we can make the following observations:

Under the same selection probability function $\Delta(u)$ and the same sample size $n$, the MSE of IWLLCQR is only slightly smaller than the ones of WLLCQR and NWLLCQR, respectively; the MSE of INWLLCQR is also slightly smaller than the ones of WLLCQR and NWLLCQR, because much more information on missing data is considered in IWLLCQR and INWLLCQR, while the MSE of INWLLCQR is slightly greater than the one of IWLLCQR. Further the MSE of IWLLCQR is only slightly greater than the one of WLLCQR; this further confirms that the IWLLCQR method is a safe alternative to WLLCQR and NWLLCQR.

Now, the simulated curves are plotted with the case of $\varepsilon_i \sim N(0, 1)$ under different levels of missing rates. Here, the results are presented only when $n = 200$, while the ones for the case of $n = 100$ are not given since these results were similar. Figures 1–3 summarize the finite sample performance of the NWLLCQR, IWLLCQR, INWLLCQR, and WLLCQR methods for $\beta(u)$ under different levels of missing rates. The red dashed curve, the blue dashed curve, the blue dotted curve, and the green dashed-dotted curve represent the results obtained by the NWLLCQR method, the IWLLCQR method, the INWLLCQR method, and the WLLCQR method, respectively. In addition, the red solid curve denotes the real curve of $\beta(u)$. From Figures 1–3, we can see that:  

1. The simulation results based on the IWLLCQR method were similar to those based on the NWLLCQR method, the INWLLCQR method, and the WLLCQR method under a lower level of missing rate. However, the IWLLCQR method outperformed the INWLLCQR method, and the INWLLCQR method outperformed the WLLCQR method, under a higher level of missing rate.

2. It can be easily found that the simulated curve obtained by the IWLLCQR method was very close to the true curve. Thus, the imputed estimation was reasonable. However, the bias of the INWLLCQR method was slightly greater than those for the IWLLCQR and the WLLCQR method.
Example 2. To examine the performance of the proposed test method, we consider the following model:

$$Y = \beta(U)X + \varepsilon,$$

where $X \sim N(0, 1)$, $\varepsilon \sim t(5)$, $U$ follows a uniform distribution on $[0, 1]$, and $X$, $U$, and $\varepsilon$ are independent.
In order to illustrate our methods by using the dataset, artificial missing data were created by deleting some of the response values in the dataset at random. Assume that 40% of the response values in this data are missed in this example. Consider the testing problem:

\[ H_0 : \beta(u) = 1 \quad v.s \quad H_1 : \beta(u) = 1 + \lambda(u^2 - 0.5) \quad (0 \leq \lambda \leq 1). \]

In what follows, the proposed test procedure is applied in a simulation with 500 replications. For each replication, 500 samples were generated, and the bootstrap sampling was repeated 300 times. Suppose the significance level \( \alpha = 0.05 \). Figure 4 shows that the simulated powers increased quickly as \( \lambda \) increased. In particular, the simulated size of the test \( T_n \) was 0.043, which is close to the true significant level of \( \alpha = 0.05 \), when the null hypothesis holds. This demonstrates that the bootstrap estimate of the null distribution was considerably effective, which shows that our test was very powerful.

![Figure 4.](image)

Figure 4. The comparison between the true curve and the NWLLCQR\(\hat{\Delta}\), IWLLCQR\(\hat{\Delta}\), INWLLCQR, and WLLCQR simulation curve when \( n = 200 \) and the selection probability function is \( \Delta_{3}(u) \).

6. A Real Data Example

In this section, we apply the methods proposed in this paper to the dataset on air pollution that the Norwegian Public Roads Administration collected. The dataset, which can be found in StatLib, consists of 500 observations. The varying coefficient model based on the CQR method was used by Guo and Tian [16] to fit the relation among the hourly values of the logarithm of the number of cars per hour \( (X_1) \), wind speed \( (X_2) \), the logarithm of the concentration of NO\(_2\) \( (Y) \), and the hour of the day \( (T) \). We deleted about 35% of the completely observed \( Y \) randomly to illustrate our proposed methods. Now, we investigate the varying coefficient model with the response data missing:

\[ Y = \beta_1(T)X_1 + \beta_2(T)X_2 + \varepsilon. \]

(15)

Since the coefficient functions of the model (15) are really time varying, we need to consider the following testing problem:

\[ H_0 : \beta(\cdot) = \beta \quad v.s \quad H_1 : \beta(\cdot) \neq \beta \]

(16)

where \( \beta = (\beta_1, \beta_2)^T \) is a constant vector. The model (15) is just a classical linear model if the null hypothesis in (16) is true. For the testing problem (16), we should reject the null hypothesis \( H_0 \) at a significance level of 0.05 because the \( p \) value of test \( T_n \) was 0.00 based on 500 resampling bootstraps.

In addition, the estimated functions of \( \hat{\beta}_1(\cdot) \) and \( \hat{\beta}_2(\cdot) \) are given, and the computational results came from 500 simulation runs. The estimated coefficients and the standard deviations are summarized in Table 2 for WLLCQR\(\hat{\Delta}\), NWLLCQR\(\hat{\Delta}\), IWLLCQR\(\hat{\Delta}\), INWLLCQR, and WLLCQR.
Table 2. The coefficient estimates and sample standard deviations (in parentheses) for the air pollution data.

|                | WLLCQR | NWLLCQR | IWLLCQR | INWLLCQR | WLLCQR |
|----------------|--------|---------|---------|----------|--------|
| $\beta_1(t)$  | $-0.312 (0.046)$ | $-0.307 (0.045)$ | $-0.315 (0.042)$ | $-0.309 (0.044)$ | $-0.316 (0.041)$ |
| $\beta_2(t)$  | $-0.379 (0.104)$ | $-0.378 (0.102)$ | $-0.375 (0.099)$ | $-0.376 (0.101)$ | $-0.374 (0.098)$ |

From Table 2, we can find that IWLLCQR̂∆ and WLLCQR had much smaller standard deviations than WLLCQR∆, WLLCQR̂∆, and INWLLCQR, respectively. In what follows, we give estimated functions of $\beta_1(\cdot)$ and $\beta_2(\cdot)$ along with the 95% bootstrap confidence bands. The results in Figures 5 and 6 show that $\beta_1(u)$ and $\beta_2(u)$ were time varying. Furthermore, we can also see that the IWLLCQR̂∆ method had almost equal confidence intervals as the WLLCQR method. Hence, the IWLLCQR̂∆ method was reasonable.

![Figure 5](image1.png)  
**Figure 5.** The comparison between the true curve and the NWLLCQR̂Δ, IWLLCQR̂Δ, INWLLCQR, and WLLCQR simulation curve when $n = 200$ and the selection probability function is $\Delta_3(u)$.

![Figure 6](image2.png)  
**Figure 6.** The comparison between the true curve and the NWLLCQR̂Δ, IWLLCQR̂Δ, INWLLCQR, and WLLCQR simulation curve when $n = 200$ and the selection probability function is $\Delta_3(u)$. 
7. Discussions

In the simulation study and the practical applications, we mainly found that the CQRE method was much more stable, efficient, and effective for varying coefficient models with response data missing at random than the single QRE method and least squares method when the sample size was large enough and the error faced a different distribution, which means the bias was at a relatively low level and the correct selection rate relatively higher.

In the face of high-dimensional data, these three method were not ideal, but the CQRE method was relatively better. This case arises widely in many research fields such as reliability life testing, genetic data research, medical tracking trials, population census, economics and finance, environment monitoring and biomedical research, etc. How to modify our method to improve the performance of the CQRE method for high-dimensional varying coefficient models with response data missing at random is an important topic that we will study further.

On the other hand, this paper with only response data missing at random studied the CQR estimation and inference for varying coefficient models. However, it did not consider the case of covariant data missing, nor even the more general case of both response and covariant data missing. These problems are more challenging topics that we will explore and study further in the coming year.

8. Concluding Remarks

In this paper, a CQRE method was proposed for varying coefficient models with response data missing at random to develop three estimators including the WLLCQR estimator, the NWLLCQR estimator, and the IWLLCQ estimator for unknown coefficient functions and establish some results on the asymptotic normality of these proposed estimators under some mild conditions. Following, a bootstrap-based test procedure was designed to perform a simulation study, which demonstrated that the unknown coefficient estimators with IWLLCQR were superior to the other two ones with WLLCQR and NWLLCQR. Meanwhile, based on the IWLLCQR fittings, a bootstrap test procedure was also designed to test whether the coefficient functions were actually varying. Finally, a type of investigated real-life dataset was analyzed to illustrate that the CQRE method was much more stable, efficient, and effective for varying coefficient models with response data missing at random than the single QRE method and least squares method.

Author Contributions: All the authors inferred the main conclusions and approved of the current version of this manuscript.

Acknowledgments: The authors would like to thank the anonymous referees for their valuable comments and suggestions, which actually stimulated this work. The work was supported by the National Natural Science Foundations of China (11601409, 71301155 and 11201362), the Natural Science Foundations of Shaanxi Province of China (2016JM1009), and the Natural Science Foundations of the Department of Shaanxi Province of China (2017J0344).

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

| QR | quantile regression |
| CQR | composite QR |
| WCQR | weighted CQR |
| CQRE | CQR estimation |
| LQQR | local linear QR |
| LLCQR | local linear CQR |
| WLLCQR | weighted LLCQR |
| NWLLCQR | nonparametric WLLCQR |
| IWLLCQR | imputed WLLCQR |
Appendix A

The following conditions are needed for the results in Section 3.

(C1) \{ (Y_i, X_i) : i = 1, 2, \ldots, n \} are independent and identically distributed random vectors.
(C2) The density function \( f(\cdot) \) of \( \epsilon \) has a continuous and uniformly-bounded derivative, namely \( 0 < \sup_s f'(s) < B_0 \).
(C3) Matrix \( E(X_i^T X_i | T = t) \) is a positive definite matrix, and \( E(X_i | T) = 0 \).
(C4) Random variable \( U \) has a second-order differentiable density function \( f_U(u) > 0 \) in some neighborhood of \( u \).
(C5) The coefficient function \( \beta(u) \) is second-order differentiable in a neighborhood of a given \( u \), and \( \beta'' \neq 0 \) is continuous.
(C6) The kernel function \( K(\cdot) \) is a symmetric density function with a compact support, whose bandwidth \( h \to 0, nh \to \infty \) as \( n \to \infty \).
(C7) The bandwidth \( h_0 \to 0, h_0/h \to 0 \), and \( nh_0 \to \infty \) as \( n \to \infty \).
(C8) The selection probability function \( \Delta(u) > 0 \) has a bounded and continuous second derivative on the support of \( U \).

The following lemma is useful for proving some theorems given in Section 3.

Lemma A1 (See Lemma 2 in [15]). Let \( (Y_1, X_1), (Y_2, X_2), \ldots, (Y_n, X_n) \) be independent and identically distributed (i.i.d.) random vectors, where the \( Y_i \)'s are scalar random variables. Suppose that \( E|Y_i|^3 < \infty \) and \( \sup_x \int |y|^3 \rho(x,y)dy < \infty \), where \( f(\cdot, \cdot) \) represents the density of \( (X, Y) \). Let \( K(\cdot) \) be a bounded positive function with \( h \) bounded support, satisfying the Lipschitz condition. Then:

\[
\sup_x \left| \frac{1}{n} \sum_{i=1}^{n} \left[ K_h(X_i - x) Y_i - E[K_h(X_i - x) Y_i] \right] \right| = O_p \left( \frac{h n^{1/2} (1/h)}{\sqrt{n h}} \right). \tag{A1}
\]

In what follows, the main theorems in Section 3 will be proven.

Proof of Theorem 1. Let \( K_i(u) = K_h \left( \frac{(U_i - u)}{h} \right), s_i(u) = (U_i - u)/h, \eta_{ik}(u) = I(\xi_i \leq \xi_k - r_i(u) - \tau_k) \) with \( r_i(u) = X_i^T \left( \beta(U_i) - \beta(U_t) - \beta'(U_t - U_{i-1}) \right), \theta = \sqrt{n h} \left( \xi_1 - c_1, \ldots, \xi_q - c_q, [\hat{a} - \beta(U_t)]', [\hat{b} - \beta'(U_t)]' \right), X_{ik}(u) = \{ e_k', X_i', X_i^T(U_i - u)/h \}' \) with \( e_k \) a \( q \)-dimensional vector with one at the \( k \)th position and zero, elsewhere. Since:

\[
(\hat{c}, \hat{a}, \hat{b}) = \arg \min_{c,a,b} \sum_{k=1}^{q} \sum_{i=1}^{n} \delta_i \Delta(U_i) \rho_{\hat{\eta}} \left( Y_i - c_k - X_i^T [a + b(U_i - u)] \right) K_h(U_i - u).
\]

\( \theta \) is the minimizer of the criterion:

\[
T_n(\pi(U_i), \theta) = \sum_{k=1}^{q} \sum_{i=1}^{n} \delta_i K_h(u) \Delta(U_i) \left( \rho_{\hat{\eta}} (\epsilon_i - c_k + r_i(u) - \Delta_{ijk}) - \rho_{\hat{\eta}} (\epsilon_i - c_k + r_i(u)) \right),
\]

where \( \Delta_{ijk} = X_{ijk}^T \theta / \sqrt{n h} \). Applying the following identity (see Knight [34]):

\[
\rho_{\pi} (x - y) - \rho_{\pi} (x) = y[I(x < s) - \tau] + \int_0^y [I(x \leq s) - I(x < s)] ds,
\]

we can rewrite \( T_n(\Delta(U_i), \theta) \) as:

\[
T_n(\Delta(U_i), \theta) = \sum_{k=1}^{q} \sum_{i=1}^{n} \delta_i K_h(u) \Delta(U_i) \left( \Delta_{ijk} I(\xi_i \leq \xi_k - r_i(u)) \right) - \tau_k
+ \int_0^{\Delta_{ijk} (I(\xi_i \leq \xi_k - r_i(u)) - I(\xi_k - \xi_i - r_i(u)) - I(\xi_i - \xi_k - r_i(u))]} dz
= \left[ W_n(u) \right]^T \theta + \sum_{k=1}^{q} B_{nk}(\theta),
\]

where \( W_n(u) \) is the weight matrix.
where
\[ W_n(u) = \frac{1}{\sqrt{n}h} \sum_{k=1}^n \sum_{i=1}^n \frac{\delta_i K_i(u)}{A(U_i)} \eta_{i,k}(u) X_{i,k}(u) \]
and,
\[ B_{n,k}(\theta) = \sum_{i=1}^n \frac{\delta_i K_i(u)}{A(U_i)} \int_0^{\Delta_{i,k}} [I(\epsilon_i - c_k + r_i(u) \leq z) - I(\epsilon_i - c_k + r_i(u) \leq 0)] \, dz. \]

Then, it follows from Lemma A1 that \( B_{n,k}(\theta) = E(B_{n,k}(\theta)) + o_p(1) \). Denote:
\[ S_n(u) = \text{Diag}\left\{ \frac{S_{n,11}(u)}{S_{n,21}(u)}, \frac{S_{n,12}(u)}{S_{n,22}(u)} \right\}, \]
where:
\[ S_{n,11}(u) = \frac{1}{nh} \text{Diag} \left\{ f(c_1) \sum_{i=1}^n K_i(u), \ldots, f(c_q) \sum_{i=1}^n K_i(u) \right\}, \]
\[ S_{n,12}(u) = \frac{1}{nh} \left[ f(c_1) \sum_{i=1}^n K_i(u) X_i, \ldots, f(c_q) \sum_{i=1}^n K_i(u) X_i \right]^T, \]
\[ S_{n,22}(u) = \frac{1}{nh} \sum_{k=1}^q f(c_k) \sum_{i=1}^n K_i(u) X_i X_i^T, \]
\[ S_{n,33}(u) = \frac{1}{nh} \sum_{k=1}^q f(c_k) \sum_{i=1}^n K_i(u) X_i^T X_i^T. \]

We observe that by the iterative expectation:
\[
\sum_{k=1}^q E[B_{n,k}(\theta)|X,U] = \sum_{k=1}^q \sum_{i=1}^n K_i(u) f_{\Delta_{i,k}} E[I(\epsilon_i - c_k + r_i(u) \leq z) - I(\epsilon_i - c_k + r_i(u) \leq 0)] \, dz
\]
\[
= \sum_{k=1}^q \sum_{i=1}^n K_i(u) f_{\Delta_{i,k}} [F(c_k - r_i(u) + z) - F(c_k - r_i(u))] \, dz
\]
\[
= \frac{1}{2} \theta^T \left( \frac{1}{nh} \sum_{k=1}^q \sum_{i=1}^n K_i(u) f(c_k) (1 + o(1)) X_{i,k}(u) X_{i,k}^T(u) \right) \theta + o_p(1)
\]
\[
= \frac{1}{2} \theta^T S_n(u) \theta + o_p(1).
\]

As in Parzen [35], we have:
\[ S_{n,11}(u) \overset{p}{\rightarrow} S_{11}(u) = f_U(u) \text{Diag}\{f(c_1), \ldots, f(c_q)\}, \]
\[ S_{n,12}(u) \overset{p}{\rightarrow} S_{12}(u) = f_U(u) E(X_i|U = u) [f(c_1), \ldots, f(c_q)]]\],
\[ S_{n,22}(u) \overset{p}{\rightarrow} S_{22}(u) = f_U(u) \sum_{k=1}^q f(c_k) E(X_i X_i^T|U = u), \]
\[ S_{n,33}(u) \overset{p}{\rightarrow} S_{33}(u) = f_U(u) \sum_{k=1}^q f(c_k) E(X_i X_i^T|U = u). \]

Based on the above results, we can prove that:
\[ T_n(\Delta(U_i), \theta) = \frac{1}{2} \theta^T S_n(u) \theta + |W_n(u)|^T \theta + o_p(1), \]
where:
\[ S(u) = \text{Diag}\left\{ \frac{S_{11}(u)}{S_{21}(u)}, \frac{S_{12}(u)}{S_{22}(u)} \right\}, \]
\[ S_{33}(u) \right\}. \]
By Corollary 2 of Knight [34], we have \( \theta \xrightarrow{p} - \left|S(u)\right|^{-1}W_n(u) \). Assume Condition (C3) is satisfied. Then, \( S(u) = \text{Diag}\{S_{11}(u), S_{22}(u), S_{33}(u)\} \). Simple calculation of the block matrix yields:

\[
\sqrt{n h} \left( \hat{\beta}(u) - \beta(u) \right) \xrightarrow{p} - \left|S_{22}(u)\right|^{-1} W_{n,2}(u),
\]

(A2)

where \( W_{n,2}(u) = \frac{1}{\sqrt{n h}} \sum_{i=1}^{n} \frac{\delta_i K_i(u)}{\Delta(U_i)} \eta_i(u) X_i \). Let \( \hat{W}_{n,2}(u) = \frac{1}{\sqrt{n h}} \sum_{i=1}^{n} \frac{\delta_i K_i(u)}{\Delta(U_i)} \eta_i X_i \). It is easy to verify that \( E(\hat{W}_{n,2}(u)) = 0 \). As in Parzen [35], we obtain:

\[
\text{Var}(\hat{W}_{n,2}(u)) = \frac{1}{n h} \sum_{i=1}^{n} \frac{K_i^2(u)}{\Delta(U_i)} \eta_i^2 X_i X_i^T = f_U(u) v_0 E \left[ \frac{\eta_i^2}{\Delta(U_i)} X_i X_i^T | U = u \right].
\]

By the central limit theorem, we get \( \hat{W}_{n,2}(u) \xrightarrow{L} N(0, f_U(u) v_0 \Omega_u) \). Similar to Kai et al. (2011) [15], we can show that:

\[
\text{Var}(\hat{W}_{n,2}(u) - W_{n,2}(u)|X, U) \leq \frac{\eta^2}{n h} \sum_{i=1}^{n} \frac{\delta_i K_i(u)}{\Delta^2(U_i)} X_i X_i^T \times \max_k \{ F(c_k + |r_i(u)| - F(c_k) \} = o_p(1).
\]

By Slutsky’s theorem, we obtain:

\[
W_{n,2}(u) - E[W_{n,2}(u)] \xrightarrow{L} N(0, f_U(u) v_0 \Omega_u).
\]

(A3)

Since:

\[
\frac{1}{\sqrt{n h}} E[W_{2n}(u)|X, U] = \frac{1}{\sqrt{n h}} \sum_{k=1}^{q} \sum_{i=1}^{n} K_i(u) [F(c_k - r_i(u) - F(c_k)] X_i
\]

\[
= - \frac{1}{\sqrt{n h}} \sum_{k=1}^{q} \sum_{i=1}^{n} K_i(u) f(c_k) [1 + o(1)] |r_i(u)| X_i,
\]

it is easy to obtain:

\[
\frac{1}{\sqrt{n h}} E[W_{2n}(u)] = - \frac{\mu_2 h^2}{2} f_U(u) S_{22} \beta''(u) + o(h^2).
\]

(A4)

Together with (A2) and (A4), we have:

\[
\sqrt{n h} \left( \hat{\beta}(u) - \beta(u) - \frac{1}{2} \mu_2 h^2 \beta''(u) \right) \xrightarrow{L} N \left( 0, \frac{v_0}{f_U(u)} D_u^{-1} \Omega_u D_u^{-1} \right).
\]

This completes the proof of Theorem 1. \( \Box \)

**Proof of Theorem 2.** Let \( \theta^* = \sqrt{n h} \{ \hat{c}_1 - c_1, \cdots, \hat{c}_q - c_q, [\hat{\mathbf{a}}^* - \beta(u)]^T, h[\hat{\mathbf{b}}^* - \beta'(u)]^T \}^T \). Similar to the proof of Theorem 1, we have:

\[
T_n^* (\hat{\Delta}(U), \theta^*) = \sum_{k=1}^{q} \sum_{i=1}^{n} \frac{\delta_i K_i(u)}{\Delta^2(U_i)} [\Delta_{ik} I_{(i \leq q)} - \tau_k] + \int_0^{\Delta_{ik}} [I_{(i < q + r_i(u) \leq z)} - I_{(i < q + r_i(u) \leq 0)}] \, dz
\]

\[
= \left[ W_{n,2}(u) \right]^T \theta^* + \sum_{k=1}^{q} B_{n,k}^*(\theta^*),
\]

where:

\[
W_{n,2}(u) = \frac{1}{\sqrt{n h}} \sum_{k=1}^{q} \sum_{i=1}^{n} \frac{\delta_i K_i(u)}{\Delta(U_i)} \eta_{i,k}(u) X_{i,k}(u),
\]

\[
B_{n,k}^*(\theta^*) = \sum_{i=1}^{n} \frac{\delta_i K_i(u)}{\Delta(U_i)} \int_0^{\Delta_{ik}} [I_{(i < q + r_i(u) \leq z)} - I_{(i < q + r_i(u) \leq 0)}] \, dz.
\]

Let:

\[
H_{n,k}^*(\theta^*) = \sum_{i=1}^{n} \frac{\delta_i (\Delta(U_i) - \Delta_{ik}(U_i))}{\Delta(U_i) \Delta_{ik}(U_i)} \times \int_0^{\Delta_{ik}} K_i(u) [I_{(i < q + r_i(u) \leq z)} - I_{(i < q + r_i(u) \leq 0)}] \, dz.
\]
Then, \( B_{n,k}^*(\theta^*) = B_{n,k}(\theta^*) + H_{n,k}^*(\theta^*) \). It is easy to verify that:

\[
\sum_{i=1}^{n} \frac{\delta_i}{\Delta(U_i)} \int_{0}^{\Delta(U_i)} K_i(u)\{I_{(\epsilon_i - c_k + r_i(u) \leq z)} - I_{(\epsilon_i - c_k + r_i(u) \leq 0)}\} \, dz = O_p(1). \tag{A5}
\]

Considering the fact that \( \sup_u |\hat{\Delta}(u) - \Delta(u)| = o(1) \), it follows from (A5) that \( H_{n,k}^*(\theta^*) = o_p(1) \), and then:

\[
\sum_{k=1}^{d} B_{n,k}^*(\theta^*) = \sum_{k=1}^{d} B_{n,k}(\theta^*) + o_p(1). \tag{A6}
\]

Similar to the proof of Theorem 1, we can prove that:

\[
\sqrt{n}h(\hat{\beta}_N(u) - \beta(u)) \xrightarrow{p} - [S_{22}(u)]^{-1} W_{n,2}^*(u), \tag{A7}
\]

where \( W_{n,2}^*(u) = \frac{1}{\sqrt{nh}} \sum_{i=1}^{n} \delta_i K_i(u) \eta_i X_i \). Let \( \hat{W}_{n,2}(u) = \frac{1}{\sqrt{nh}} \sum_{i=1}^{n} \delta_i K_i(u) \eta_i X_i \). By the proof of Theorem 3 in Wong [33], we can obtain:

\[
W_{n,2}(u) = \frac{1}{\sqrt{nh}} \sum_{i=1}^{n} \delta_i K_i(u) \eta_i X_i + \frac{1}{\sqrt{nh}} \sum_{i=1}^{n} \delta_i \frac{\Delta(U_i)}{\Delta(U_i)} E[K_i(u)\eta_i X_i|Z_i] + o(h^2)
\]

\[
= \hat{W}_{n,21}(u) + W_{n,22}(u) + o(h^2),
\]

where \( E(\hat{W}_{n,21}(u)) = 0 \) and \( E(\hat{W}_{n,22}(u)) = 0 \). Furthermore,

\[
Var(\hat{W}_{n,21}(u)) = f_{U}(u)v_0 E\left[ \frac{\eta^2_i}{\Delta(U_i)} X_i X_i^T | U_i \right],
\]

\[
Var(\hat{W}_{n,22}(u)) = f_{U}(u)v_0 E\left[ \frac{1 - \Delta(U_i)}{\Delta(U_i)} E\{X_i \eta_i | U_i \} \otimes 2 \right],
\]

\[
Cov(\hat{W}_{n,21}(u), \hat{W}_{n,22}(u)) = f_{U}(u)v_0 E\left[ \frac{1 - \Delta(U_i)}{\Delta(U_i)} E\{X_i \eta_i | U_i \} \otimes 2 \right].
\]

Completing the calculation, we obtain:

\[
Var(\hat{W}_{n,2}(u)) = f_{U}(u)v_0 \left\{ E\left[ \frac{\eta^2_i}{\Delta(U_i)} X_i X_i^T | U_i \right] - E\left[ \frac{1 - \Delta(U_i)}{\Delta(U_i)} E\{X_i \eta_i | U_i \} \otimes 2 \right] \right\} + o(1).
\]

Based on the above results, it follows that \( \hat{W}_{n,2}(u) \xrightarrow{L} N(0, f_{U}(u)v_0 \Omega_u) \). Similar to the proof of Theorem 1, we have: \( Var(\hat{W}_{n,2}(u) - W_{n,2}(u)|X, U) = o_p(1) \). Thus:

\[
W_{n,2}(u) - E[W_{n,2}(u)] \xrightarrow{L} N(0, f_{U}(u)v_0 \Omega_u) \tag{A8}
\]

By Lemma A1, we get:

\[
\frac{1}{nh} \sum_{i=1}^{n} \delta_i K_i(u) \eta_i X_i \xrightarrow{p} E\left[ \frac{1}{nh} \sum_{i=1}^{n} \delta_i K_i(u) \eta_i X_i \right] = O(h^2).
\]

Since \( \frac{1}{\Delta(U_i)} - \frac{1}{\Delta(U_i)} = o_p(1) \), then

\[
\frac{1}{\sqrt{nh}} W_{n,2}(u) = \frac{1}{\sqrt{nh}} \sum_{i=1}^{n} \delta_i K_i(u) \eta_i X_i + \frac{1}{\sqrt{nh}} \sum_{i=1}^{n} \delta_i \left( \frac{1}{\Delta(U_i)} - \frac{1}{\Delta(U_i)} \right) K_i(u) \eta_i X_i
\]

\[
= \frac{1}{\sqrt{nh}} \sum_{n=1}^{n} W_{n,2}(u) + o_p(h^2).
\]
Thus, we can show that:
\[
\frac{1}{\sqrt{n\Delta}} E[W_{n,2}^*(u)] = \frac{1}{\sqrt{n\Delta}} E[W_{n,2}(u)] + o(n^2). \tag{A9}
\]
Following (A7), (A9), and Theorem 1, we complete the proof of Theorem 2. □

**Proof of Theorem 3.** Write \( \theta^{**} = \sqrt{n\Delta}\{\varepsilon_1 - c_1, \cdots, \varepsilon_q - c_q, \{\hat{\theta}^{**} - \beta(u)\}^T, h\{\hat{\beta}^{**} - \beta'(u)\}^T. \)

\( \Delta_{i,k}^* = X_i^T\theta^{**}/\sqrt{n\Delta}, \eta_{i,k}(u) = I((\delta_{1\Delta}^*)_{i,k} - r_i(u))' \) then we have:

\[
Y_i^* = \frac{\delta_i}{\hat{\pi}(U_i)} Y_i + (1 - \frac{\delta_i}{\hat{\pi}(U_i)}) X_i^T \hat{\beta}(U_i) = \frac{\delta_i}{\hat{\pi}(U_i)} \varepsilon_i + X_i^T \hat{\beta}(U_i) + o_p(1),
\]

Similar to the proof of Theorem 2, we have:

\[
T_n^{**} = \sum_{k=1}^n \sum_{j=1}^n \delta_j K_i(u) \left\{ \Delta_{i,k}^* \left[ I_{(\delta_{1\Delta}^*)_{i,k} \leq r_i(u) - \tau_k} \right] + \int_0^{\Delta_{i,k}^*} \left[ I_{(\delta_{1\Delta}^*)_{i,k} \epsilon - \tau_k} \right] \right\} dz - \sum_{k=1}^n \sum_{j=1}^n \left( 1 - \frac{\delta_j}{\hat{\pi}(U_j)} \right) K_j(u) \left\{ \int_{(\delta_{1\Delta}^*)_{i,k} \leq r_i(u) - \tau_k} \right\} dz + \int_{(\delta_{1\Delta}^*)_{i,k} \leq r_i(u) - \tau_k} \right\} dz.
\]

We can prove:

\[
\sum_{i=1}^n \left( 1 - \frac{\delta_i}{\hat{\pi}(U)} \right) K_i(u) \int_{0}^{\Delta_{i,k}^*} \left[ I_{(\delta_{1\Delta}^*)_{i,k} \leq r_i(u) - \tau_k} \right] dz = o_p(1).
\]

Similar to the proof of Theorem 1, we can complete the proof. □

**References**

1. Hastie, T.J.; Tibshirani, R.J. Varying-coefficient models. *J. R. Stat. Soc. Ser. A* 1993, 55, 757–796. [CrossRef]
2. Chiapp, C.T.; Rice, J.A.; Wu, C.O. Smoothing spline estimation for varying coefficient models with repeatedly measured dependent variables. *J. Am. Stat. Assoc.* 2001, 96, 605–619. [CrossRef]
3. Eubank, R.L.; Huang, C.; Maldonado, Y.M.; Wang, N.; Wang, S.; Buchanan, R.J. Smoothing spline estimation in varying coefficient models. *J. R. Stat. Soc. Ser. A* 2004, 66, 653–667. [CrossRef]
4. Fan, J.; Zhang, J.T. Statistical estimation in varying coefficient models. *Ann. Stat.* 1999, 27, 1491–1518.
5. Huang, J.; Wu, C.O.; Zhou, L. Varying coefficient models and basis function approximations for the analysis of repeated measurements. *Biometrika* 2002, 89, 111–128. [CrossRef]
6. Wu, C.O.; Yu, K.F.; Chiang, C.T. A two-step smoothing method for varying coefficient models with repeated measurements. *Ann. Inst. Stat. Math.* 2000, 52, 519–543. [CrossRef]

7. Fan, J.; Huang, T. Profile likelihood inferences on semiparametric varying-coefficient partially linear models. *Bernoulli* 2001, 11, 1031–1057. [CrossRef]

8. Whang, Y.J. Smoothed empirical likelihood methods for quantile regression models. *Econ. Theory* 2006, 22, 173–205. [CrossRef]

9. Koenker, R. *Quantile Regression*; Cambridge University Press: Cambridge, UK, 2005.

10. Kim, M.O. Quantile regression with varying coefficients. *Ann. Stat.* 2007, 35, 92–108. [CrossRef]

11. Cai, Z.; Xu, X. Nonparametric quantile estimations for dynamic smooth coefficient models. *J. Am. Stat. Assoc.* 2008, 103, 1595–1608. [CrossRef]  

12. Cai, Z.; Xiao, Z. Semiparametric quantile regression estimation in dynamic models with partially varying coefficients. *J. Econom.* 2012, 167, 413–425. [CrossRef]

13. Tang, Q.G. Robust estimation for spatial semiparametric varying coefficient partially linear regression. *Stat. Pap.* 2015, 56, 1137–1161.

14. Zou, H.; Yuan, M. Composite quantile regression and the oracle model selection theory. *Ann. Stat.* 2008, 36, 1108–1126. [CrossRef]

15. Kai, B.; Li, R.; Zou, H. New efficient estimation and variable selection methods for semiparametric varying coefficient partially linear models. *Ann. Stat.* 2011, 39, 305–332. [CrossRef] [PubMed]

16. Guo, J.; Tian, M.Z. New efficient and robust estimation in varying coefficient models with heteroscedasticity. *Stat. Sin.* 2012, 22, 1075–1101.

17. Sun, J.; Gai, Y.; Lin, L. Weighted local linear composite quantile estimation for the case of general error distributions. *J. Stat. Plan. Inference* 2013, 143, 1049–1063. [CrossRef]

18. Yang, H.; Lv, J.; Guo, C.H. Weighted composite quantile regression estimation and variable selection for varying coefficient models with heteroscedasticity. *J. Korean Stat. Soc.* 2015, 44, 77–94. [CrossRef]

19. Luo, S.; Zhang, C.-Y. Nonparametric M-type regression estimation under missing response data. *Stat. Pap.* 2016, 57, 641–664. [CrossRef]

20. Rubin, D.B. Inference and missing data. *Biometrika* 1976, 63, 581–592. [CrossRef]

21. Sterne, J.; White, I.; Carlin, J.; Spratt, M.; Kenward, M.; Carpenter, J. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* 2009, 338, b2393. [CrossRef] [PubMed]

22. Wang, Q.; Linton, O.; Härdle, W. Semiparametric regression analysis with missing response at random. *J. Am. Stat. Assoc.* 2004, 99, 334–345. [CrossRef]

23. Wang, Q.; Sun, Z. Estimation in partially linear models with missing responses at random. *J. Multivar. Anal.* 2007, 98, 1470–1493. [CrossRef]

24. Wang, Q.; Rao, N.K. Empirical Likelihood-based inference under imputation for missing response data. *Ann. Stat.* 2002, 30, 896–924. [CrossRef]

25. Xue, L.G. Empirical likelihood confidence intervals for response mean with data missing at random. *Scand. J. Stat.* 2009, 36, 671–685. [CrossRef]

26. Wei, Y.; Ma, Y.; Carroll, R. Multiple imputation in quantile regression. *Biometrika* 2012, 99, 423–438. [CrossRef] [PubMed]

27. Lv, X.; Li, R. Smoothed empirical likelihood analysis of partially linear quantile regression models with missing response variables. *Adv. Stat. Anal.* 2013, 97, 317–347. [CrossRef]

28. Sherwood, B.; Wang, L.; Zhou, X. Weighted quantile regression for analyzing health care cost data with missing covariates. *Stat. Med.* 2013, 32, 4967–4979. [CrossRef]

29. Sun, Y.; Wang, Q.; Gilbert, P. Quantile regression for competing risks data with missing cause of failure. *Ann. Stat.* 2012, 22, 703–728. [CrossRef]

30. Chen, X.; Wan, T.K.; Zhou, Y. Efficient quantile regression analysis with missing observations. *J. Am. Stat. Assoc.* 2015, 110, 723–741. [CrossRef]

31. Kim, S.Y. Imputation methods for quantile estimation under missing at random. *Stat. Its Interface* 2013, 6, 369–377.

32. Nageswara, S.; Rao, V. Nadaraya-Watson estimator for sensor fusion. *Opt. Eng.* 1997, 36, 642–647.

33. Wong, H.; Guo, S.J.; Chen, M.; Wai-Cheung, I.P. On locally weighted estimation and hypothesis testing on varying coefficient models with missing covariates. *J. Stat. Plan. Inference* 2009, 139, 2933–2951. [CrossRef]
34. Knight, K. Limiting distributions for L1 regression estimators under general conditions. *Ann. Stat.* 1998, 26, 755–770.

35. Parzen, E. On estimation of a probability density function and model. *Ann. Math. Stat.* 1962, 33, 1065–1076. [CrossRef]

© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).