

# Robust Inference for Time-Varying Coefficient Models with Longitudinal Data

Zhaofeng Wang<sup>1</sup>, Jiancheng Jiang<sup>2</sup>, Qunyi Qiu<sup>3</sup>

<sup>1</sup>Jishou University, Jishou, China

<sup>2</sup>University of North Carolina, Charlotte, NC, USA

<sup>3</sup>Peking University, Beijing, China

Email: jdwzf@126.com, jjiang1@uncc.edu, qyqiu@math.pku.edu.cn

Received 28 September 2015; accepted 14 December 2015; published 17 December 2015

Copyright © 2015 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

---

## Abstract

Time-varying coefficient models are useful in longitudinal data analysis. Various efforts have been invested for the estimation of the coefficient functions, based on the least squares principle. Related work includes smoothing spline and kernel methods among others, but these methods suffer from the shortcoming of non-robustness. In this paper, we introduce a local M-estimation method for estimating the coefficient functions and develop a robustified generalized likelihood ratio (GLR) statistic to test if some of the coefficient functions are constants or of certain parametric forms. The robustified GLR test is robust against outliers and the error distribution. This provides a useful robust inference tool for the models with longitudinal data. The bandwidth selection issue is also addressed to facilitate the implementation in practice. Simulations show that the proposed testing method is more powerful in some situations than its counterpart based on the least squares principle. A real example is also given for illustration.

## Keywords

Local Polynomial Smoothing, Longitudinal Data, Local M-Estimators, Generalized Likelihood Ratio Tests

---

## 1. Introduction

The defining characteristic of a longitudinal data study is that individuals are measured repeatedly over a given time period, longitudinal studies are in contrast to cross-sectional studies, in which a single outcome is measured for each individual. The repeated measurements within each subject are generally correlated with each other, but different individuals can be assumed to be independent. The primary advantage of a longitudinal study is its effectiveness for studying changes over time. Statistical research in this field has been very active, and many pa-

rametric models have been developed. See Diggle *et al.* [1], Davidian and Giltinan [2], Vonesh and Chinchilli [3], and the references therein.

As overwhelming longitudinal data exist in biomedical studies, there are increasing demands for a generally applicable inference tool for analysing these datasets. The parametric models are efficient for analysing longitudinal data but may suffer from mis-specification. To reduce possible modeling bias, different nonparametric and semiparametric methods have been studied in the literature, for example Zeger and Diggle [4], Müller [5], Eubank *et al.* [6], He *et al.* [7] [8], among others. In particular, a useful nonparametric model for analysing time-varying effects of covariates receives much attention. Examples include Hoover *et al.* [9], Fan and Zhang [10], Huang *et al.* [11] and others. These works focus on the least-squares based estimation. While they are useful in some applications, the shortcoming for lack of robustness is naturally raised. This motivates us to consider a robust inference tool for the time-varying coefficient models.

In this paper, we consider a local M-estimation approach based on local polynomial smoother and a robustified “generalized likelihood ratio (*GLR*)” statistic to test if parts of the coefficients are constants or of certain parametric forms. This in particular allows one to nonparametrically check the goodness-of-fit of the usual linear models widely used in practice (see for example Diggle *et al.* [1]). We conduct extensive simulations to demonstrate that the proposed estimation method is robust against outliers and error distributions, and that the robustified *GLR* tests are powerful than its counterpart when the error deviates away from normality.

This paper is organized as follows. In section 2, the local M-estimation approach is introduced; a data-driven bandwidth selection rule is also given. In Section 3, we focus on the robustified *GLR* tests. Simulations are conducted in Section 4, where the performances of the different tests are compared, and a real example is also used to illustrate the proposed method. Finally, the paper is concluded by a discussion.

## 2. Model and Estimation

### 2.1. Local M-Estimation

Consider the following time-varying coefficient model,

$$Y(t) = X(t)^T \beta(t) + \varepsilon(t), \quad (1)$$

where  $\beta(t) = (\beta_0(t), \beta_1(t), \dots, \beta_k(t))^T$  for  $k \geq 0$  and  $\varepsilon(t)$  is a zero-mean correlated stochastic process that cannot be explained by the covariates  $X(t)$ . As in Hoover *et al.* [9] we regard the repeated observations  $(Y_{ij}, X_{ij}, t_{ij})$ , for  $i = 1, 2, \dots, n$  and  $j = 1, 2, \dots, n_i$ , as a random sample from model (1), that is,

$$Y_{ij} = X_{ij}^T \beta(t_{ij}) + \varepsilon_i(t_{ij}), \quad (2)$$

where  $Y_{ij} = Y_i(t_{ij})$  is the observed response,  $X_{ij} = X_i(t_{ij}) = (X_{ij0}, X_{ij1}, \dots, X_{ijk})^T$  is the observed covariates for the  $i$ th subject at time  $t_{ij}$ , and  $\varepsilon_i(t)$  is a zero-mean stochastic process with covariance function  $r(s, t) = \text{Cov}(\varepsilon_i(s), \varepsilon_i(t))$ . For this model, we assume that the measurements on the responses for different subjects are independent, but  $X_{ij}$  may be correlated at different time points within each subject,  $\varepsilon_i(t)$ 's are also independent for different subjects. Model (1) or (2) is a useful extension to the usual linear model for longitudinal/panel data analysis in Lindsey [12], Jones [13], Diggle, *et al.* [1], Hand and Crowder [14], among others.

There are several methods for estimating the coefficient  $\beta(t)$ , for example, the smoothing spline method in Brumback and Rice [15], Hoover *et al.* [9], and Chiang *et al.* [16], the kernel smoother in Hoover *et al.* [9], Wu *et al.* [17], Wu and Chiang [18], and Chiang *et al.* [16], and other methods such as the two-step estimation in Fan and Zhang [10] and the global smoothing procedure using basis function approximations in Huang *et al.* [11]. These methods are all based on the least squares principle and suffer from the shortcoming of non-robustness. It is worthy of developing a robust estimation and testing method for the model (1). Generally speaking, one can develop different testing methods for different estimating approaches. To introduce our method, we begin with the local polynomial estimation (see Hoover *et al.* [9]), that is, to find  $b_{rl}$ 's to minimize

$$\sum_{i=1}^n \sum_{j=1}^{n_i} \left[ Y_{ij} - \sum_{l=0}^k X_{ijl} \sum_{r=0}^d (t_{ij} - t)^r b_{rl} \right]^2 K_h(t - t_{ij}), \quad (3)$$

where  $K_h(\cdot) = h^{-1}K(\cdot/h)$ , and  $d$  denotes the order of the polynomial used in smoothing.

The above local least squares based estimation is not robust against outliers and heavy tailed errors. To fix this problem, we propose to estimate the coefficient function by minimizing

$$\sum_{i=1}^n \sum_{j=1}^{n_i} \rho \left( Y_{ij} - \sum_{l=0}^k X_{ijl} \sum_{r=0}^d (t_{ij} - t)^r b_{rl} \right) K_h(t - t_{ij}), \quad (4)$$

where  $\rho(\cdot)$  is an outlier resistant function. The coefficient functions  $\beta_l(t)$ 's is estimated by  $\hat{b}_{0l}$ 's ( $l = 0, \dots, k$ ), the solutions of the above optimization problem. The resulting estimators are so-called the local M-type estimator of  $\beta(t)$ . If  $\rho(\cdot)$  has a derivative  $\psi(\cdot)$ , then the solutions of (4) solve the following equations:

$$\sum_{i=1}^n \sum_{j=1}^{n_i} \psi \left( Y_{ij} - \sum_{l=0}^k X_{ijl} \sum_{r=0}^d (t_{ij} - t)^r b_{rl} \right) (t_{ij} - t)^r X_{ijl} K_h(t - t_{ij}) = 0, \quad (5)$$

for  $l = 0, \dots, k$  and  $r = 0, \dots, d$ .

The above method is in the same spirit of the local M-estimation studied for cross-sectional data in Fan and Jiang [19] and Jiang and Mack [20]. It can be shown that the estimator is  $\sqrt{n}$ -consistent under certain conditions. The estimator involves a selection of the outlier resistant function. Much work in the literature has demonstrated that the Huber's  $\psi_k$ -class, *i.e.*,  $\psi_k(x) = \max\{-k, \min(k, x)\}$ , is satisfactory for robust estimation in the location model and nonparametric regression (Huber [21]; Jiang and Mack [20]), where a method for determining the parameter  $k$  was studied in Jiang and Mack [20]. However, our experience shows that a rule of thumb for the choice of  $k$ , such as  $k = 1.35s$  where  $s$  is the robust standard deviation of the residuals (Serfling [22]), is simple and satisfactory in the present situation. Other choices of the outlier resistant function are also possible, such as  $\rho(x) = |x|$  which leads to the least absolute deviation estimation (Jiang *et al.*, [23]).

The Newton algorithm can be used to find the solutions to the Equations (5). If the initial values of the parameters for iteration are good enough, for example, from the least squares estimation, then the iterative solutions can be found in a few steps, which is theoretically verified in Jiang and Mack [20] for nonparametrically modeling dependent data. In our simulations, the least squares based estimators in (2) will be employed as the initial values.

## 2.2. Bandwidth Selection

The performance of the estimator in (4) depends on the smoothing parameter  $h$ . There are several approaches to the selection of the bandwidth, such as the cross-validation (CV) and generalized cross-validation (GCV) criteria in Hastie and Tibshirani [24]. We here extend the CV method to the present situation for determining the bandwidth  $h$ . Specifically, denote by  $\hat{\beta}^{(-i)}(t)$  the solution of (4) or (5) based on all the observations without the measurements for  $i$ th subject. Then the bandwidth  $h$  can be estimated by the minimizer  $\hat{h}_{cv}$  of

$$CV(h) = \frac{1}{N} \sum_{i=1}^n \sum_{j=1}^{n_i} \rho \left( Y_{ij} - X_{ij}^T \hat{\beta}^{(-i)}(t_{ij}) \right), \quad (6)$$

where  $N = \sum_{i=1}^n n_i$ .

The minimization of (6) is time-consuming for simulations, even though it is not for real data analysis. We here suggest a pre-determination of  $\hat{h}_{cv}$  before simulations. Specifically, the procedure is detailed as follows:

1) Generate several samples (20 for instance) from (2), then minimize (6) to get the  $\hat{h}_{cv}$  for each sample and compute the average  $\bar{h}_{cv}$  of the  $\hat{h}_{cv}$ 's;

2) Fix the bandwidth  $\bar{h}_{cv}$  in each simulation to find the estimator in (4).

In step 1), the computational burden can be further reduced if one uses an easy-evaluated bandwidth as initial value, such as the one from the GCV criterion for the local least squares based estimator. Since the estimated bandwidths for the local least squares estimation and the local M-estimation are highly correlated in general, the minimization of  $CV(h)$  will be achieved in a few iterations.

## 3. Robustified GLR Tests

For simplicity, consider the following testing problem

$$H_0 : \boldsymbol{\beta}(t) = \boldsymbol{\beta} \leftrightarrow H_1 : \boldsymbol{\beta}(t) \neq \boldsymbol{\beta}, \quad (7)$$

where  $\boldsymbol{\beta}(t) = (\beta_0(t), \beta_1(t), \dots, \beta_k(t))^T$ , and  $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_k)^T$  is an unspecified constant vector. Let

$$RSS_0 = \sum_{i=1}^n \sum_{j=1}^{n_i} \rho \left( Y_{ij} - \sum_{l=0}^k X_{ijl} \hat{\beta}_l \right),$$

$$RSS_1 = \sum_{i=1}^n \sum_{j=1}^{n_i} \rho \left( Y_{ij} - \sum_{l=0}^k X_{ijl} \hat{\beta}_l(t) \right),$$

where  $\hat{\beta}_l$ 's are the usual M-estimators of the coefficients under  $H_0$ , and  $\hat{\beta}_l(\cdot)$ 's are the local M-estimators of the coefficient functions under  $H_1$ . Define the following testing statistic for the testing problem (7):

$$T_N = \frac{N}{2} \log \frac{RSS_0}{RSS_1} \approx \frac{N}{2} \frac{RSS_0 - RSS_1}{RSS_1}. \quad (8)$$

Intuitively, large values of  $T_N$  provide evidences against  $H_0$ . The proposed statistic is basically based on the comparison of residuals between the null and the alternative. It can be regarded as a robustification for the GLR testing statistic studied in Fan *et al.* [25], Fan and Huang [26], and Fan and Jiang ([27] [28]). In particular, if  $\rho(x) = x^2$  is employed, then  $T_N$  is in the same formulas as those in the afore-mentioned papers. For the errors departing away from normality, one can expect the  $T_N$  to be robust and more powerful than its local least-squares counterpart with  $\rho(x) = x^2$ . For more general testing problems such as  $H'_0 : \boldsymbol{\beta}(\cdot) = \boldsymbol{\beta}_\theta(\cdot)$  for  $\theta \in \Theta$ , which tests whether  $\boldsymbol{\beta}(\cdot)$  admits a certain parametric form, the above testing statistic can be similarly constructed if one replaces the  $\hat{\beta}_l$  with the local M-estimator under  $H'_0$ .

### Bootstrap Estimate of Null Distribution

To implement the robustified GLR tests, one needs to obtain the null distribution of  $T_N$  under  $H_0$ . In the following, we use simulations to compute the null distribution of the test statistic for a finite sample. It can be applied to both the local least squares and the local M-estimation methods. The computational algorithm is given as follows. For an easy illustration, we first consider the situation with  $n_1 = n_2 = \dots = n_n$ .

- 1) Obtain the nonparametric estimate  $\hat{\boldsymbol{\beta}}(\cdot)$  along with its associated bandwidth  $\hat{h}$ .
- 2) Compute the testing statistic  $T_N$  and the residual vectors  $\hat{\varepsilon}_i(\cdot)$  under the alternative, for  $i = 1, 2, \dots, n$ .
- 3) For each given  $\mathbf{X}_i$ , draw a bootstrap residual vectors  $\varepsilon_i^*(\cdot)$  from the centered empirical distribution of  $\hat{\varepsilon}_i(\cdot)$  and compute  $Y_i^* = \mathbf{X}_i^T \boldsymbol{\beta} + \varepsilon_i^*$ .
- 4) Use the generated random sample  $(\mathbf{X}_{ij}, Y_{ij}^*)$  ( $i = 1, \dots, n; j = 1, \dots, n_i$ ) and the bandwidth  $\hat{h}$  to calculate the value  $T_N^*$  of the robustified GLR testing statistic.
- 5) Repeat steps 3 and 4  $B$  times to obtain the bootstrap statistics  $T_{N1}^*, \dots, T_{NB}^*$ .
- 6) The  $P$ -value of the testing statistic  $T_N$  is the percentage of the bootstrap statistics  $\{T_{N1}^*, \dots, T_{NB}^*\}$  that exceed  $T_N$ .

Similar simulation approaches to determining the  $p$ -value of a testing statistic were given in Fan and Jiang [27] for additive models and Hui and Jiang [29] for DTARCH models. For the case that  $n_i$ 's are not equal, the conditional bootstrap method is infeasible, but one can use a resampling subject bootstrap method (see for example Huang, Wu and Zhou, [11]) to replace the bootstrap method above. Specifically, let  $\{Y_{ij}^p; i = 1, \dots, n, j = 1, \dots, n_i\}$  be generated as follows:

$$Y_{ij}^p = \mathbf{X}_{ij}^T \hat{\boldsymbol{\beta}} + \hat{\varepsilon}_{ij},$$

where  $\hat{\boldsymbol{\beta}}$  is the estimate under  $H_0$  and  $\hat{\varepsilon}_{ij}$ 's are the residuals from the null model. Resample  $n$  subjects with replacement from  $\{(Y_{ij}^p, \mathbf{X}_{ij}, t_{ij}); i = 1, \dots, n, j = 1, \dots, n_i\}$  to obtain a bootstrap sample

$\{Y_{ij}^*, \mathbf{X}_{ij}^*, t_{ij}^*; 1 \leq i \leq n, 1 \leq j \leq n_i^*\}$ . Let  $T_N^*$  be the value of the testing statistic for the bootstrap sample. With  $B$

independent replications, we obtain  $B$  bootstrap estimators  $\{T_{N1}^*, \dots, T_{NB}^*\}$ . Then p-value of the testing statistic  $T_N$  can be obtained as in step 6 above.

## 4. Numerical Studies

### 4.1. Robustness of the Estimation

In this section, we compare the performance of the local M-estimation with the local least squares estimation. For the outlier-resistant function, we opt for the Huber  $\psi_k$ -functions in our numerical study:

$$\psi(u) = \min(k, \max(u, -k)).$$

The rule of thumb in Section 2.1 for the choice of  $k$  will be employed.

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

$$Y_i(t_{ij}) = X_{ij0}\beta_0(t_{ij}) + X_{ij1}\beta_1(t_{ij}) + \varepsilon_i(t_{ij}), \quad (9)$$

where  $\beta_0(t) = 0.5 + 3t$ ,  $\beta_1(t) = t + 2\exp(-16t^2)$ ,  $\{X_{ij0}\}$  and  $\{X_{ij1}\}$  are respectively from  $AR(1)$  models in such a way that  $X_{ij0} = 0.2X_{i(j-1)0} + \epsilon_{ij0}$ ,  $\epsilon_{ij0} \stackrel{iid}{\sim} N(0, 0.4^2)$ ,  $X_{ij1} = 0.2X_{i(j-1)1} + \epsilon_{ij1}$ , and  $\epsilon_{ij1} \stackrel{iid}{\sim} N(0, 1)$ . For simplicity, we used the time points  $t_{ij} = 0.1j$  which are equally spaced for each  $i$ . The errors  $\{\varepsilon_i(t_{ij})\}$  are iid for all  $i$  and  $j$ . We considered the following four distributions for the errors:

(A1)  $\varepsilon_i(t_{ij}) \sim N(0, 1)$ ;

(A2)  $\varepsilon_i(t_{ij}) \sim t_{(3)}$ ;

(A3)  $\varepsilon_i(t_{ij}) \sim \text{centered } \chi_{(5)}^2$ ;

(A4)  $\varepsilon_i(t_{ij}) \sim 0.9N(0, 1) + 0.1N(0, 5^2)$ .

We simulated 200 samples of size  $n = 100$  and  $n_i = 20$  from the model (9). The averages of  $\hat{\beta}(t)$  and the 2.5% and 97.5% sample quantiles of  $\hat{\beta}(t)$  among simulations were calculated at each time points. In addition, we also computed the mean absolute deviation error (MADE) of the estimators at the time points (for  $k = 1, 2$ ),

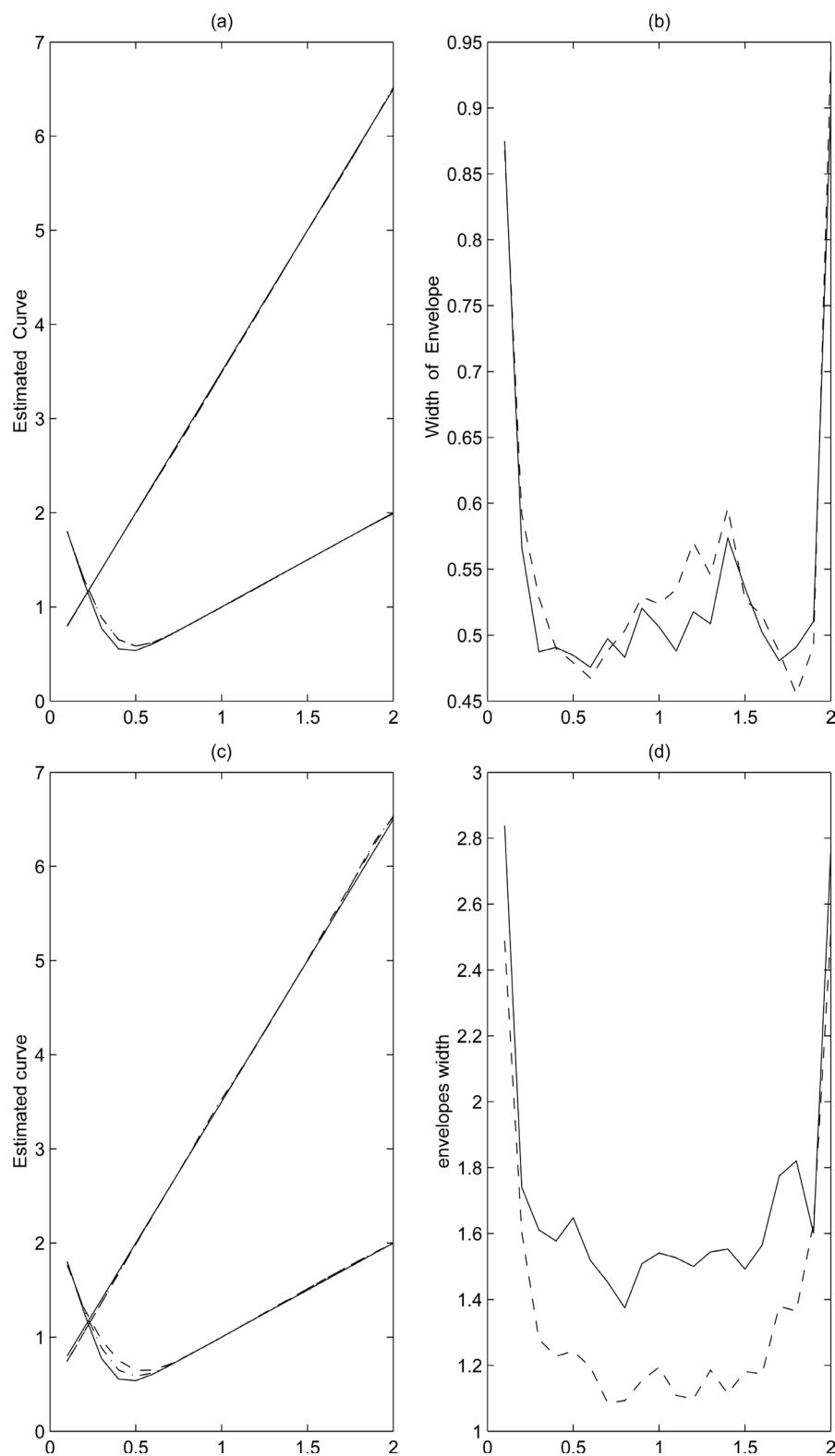
$$\text{MADE}(\hat{\beta}_k(t)) = \frac{1}{\sum_{i=1}^n n_i} \sum_{i,j} |\hat{\beta}_k(t_{ij}) - \beta_k(t_{ij})|.$$

The average of values of the above MADE over simulations was calculated and reported in [Table 1](#) for the error models (A1) and (A3). The results show that the local LS estimator and the local M-estimators have similar accuracy if the error is normal, but the latter performs better than the former if the error deviates away from normality. We display in [Figure 1](#) the estimators of the coefficient functions along with the widths of the envelopes formed by pointwise 2.5% and 97.5% quantiles of the estimators among simulations, where  $\beta_0(t)$  is easier to estimate than  $\beta_1(t)$  and the estimated envelopes' widths for  $\hat{\beta}_0(t)$  are not reported for saving space. It is evident that both the local LS estimator and the local M-estimator have little biases, and the latter is much better than the former in terms of the envelopes' widths when the error deviates away from the normal distribution.

We did simulations with the error distributions in (A2) and (A4). The results are similar and omitted for saving space.

**Table 1.** MADEs under the errors (A1) and (A3) for Example 1.

MADE	(A1)		(A3)	
	$\hat{\beta}_0(t)$	$\hat{\beta}_1(t)$	$\hat{\beta}_0(t)$	$\hat{\beta}_1(t)$
Local LS-estimator	0.1124	0.0415	0.3541	0.1060
Local M-estimator	0.1150	0.0420	0.2810	0.0978



**Figure 1.** Estimated curves and widths of envelopes. Upper panel: results for the error in (A1); lower panel: results for the error in (A2). (a) and (c): the average of estimated curves; solid—true curves, dash-dotted—local LS estimator, dashed—local M-estimator. (b) and (d): pointwise widths of envelopes for  $\beta_1(t)$ , which were formed by the pointwise 2.5% & 97.5% quantiles of the estimators among simulations; solid—local LS estimator, dashed—local M-estimator.

**Example 2.** In this example, we consider the model (9) with much more complex structures for the covariates, the errors, and the coefficient functions than those in **Example 1**. Different from **Example 1**, the two covariates are now correlated, and the error  $\varepsilon_i(t_{ij})$  are correlated within subjects. Specifically,

$\beta_0(t) = 0.3 \exp(-7(t-0.3)^2) + 0.7 \exp(-7(t-1.3)^2)$ ,  $\beta_1(t) = 0.5 + 0.3t + 4t^2$ ,  $X_{ij0}$  and  $X_{ij1}$  are from the following two-dimensional AR model,

$$\begin{pmatrix} X_{ij0} \\ X_{ij1} \end{pmatrix} = \begin{pmatrix} 0.3 & 0.4 \\ 0.1 & 0.6 \end{pmatrix} \begin{pmatrix} X_{i(j-1)0} \\ X_{i(j-1)1} \end{pmatrix} + \begin{pmatrix} \epsilon_{ij0} \\ \epsilon_{ij1} \end{pmatrix},$$

where  $\epsilon_{ij0} \sim N(0, 0.4^2)$  and  $\epsilon_{ij1} \sim N(0, 1)$ . We used the same  $t_{ij}$ 's as in **Example 1**. The errors  $\varepsilon_i(t_{ij})$  are iid for all different subjects but may be correlated within subjects, which were generated from the following three distributions:

(B1)  $\varepsilon_i(t_{ij}) \sim N(0, 1)$ ;

(B2)  $\varepsilon_i(t_{ij}) = z_i(t_{ij}) + 0.5z_i(t_{i(j-1)})$ , where  $z_i(t_{ij}) \sim N(0, 1)$ ;

(B3)  $\varepsilon_i(t_{ij}) = z_i(t_{ij}) + 0.5z_i(t_{i(j-1)})$ , where  $z_i(t_{ij}) \sim t(3)$ .

We conducted 400 simulations. In each simulation, a samples of size  $n = 100$  and  $n_i = 20$  was drawn. We calculated the estimators along with the related MADEs and envelopes. **Table 2** reports the MADEs for the estimated coefficient functions. It is evidenced that the local M-estimator is better than its counterpart under the heavy tailed error in (B3). Both estimators have similar MADEs under the other two errors in (B1) and (B2).

Since  $\beta_1(t)$  is much simpler than  $\beta_0(t)$ , we now consider only the the estimated curves for  $\beta_0(t)$ . **Figure 2** displays the estimated curves for  $\beta_0(t)$  along with the 95% confidence intervals under the error (B2), and the widths of the envelopes for  $\beta_0(t)$  under the errors in (B2) and (B3). The average of the estimator in **Figure 2(a)** and the corresponding 95% confidence intervals in **Figure 2(b)** show that the local M-estimator well captures the structure of the true curve  $\beta_0(t)$ . It is seen from **Figure 2(c)** and **Figure 2(d)** that the local M-estimator is better than the local LS estimator when the error deviates away from the norm distribution.

## 4.2. Null Distribution and Power of Test

In this section, we conduct simulations to show that our conditional bootstrap method gives a good estimate of the null distribution, and to compare the powers of different nonparametric tests.

**Example 3.** To compare the powers of different tests, we consider the testing problem (7) for the model (9). The null model is the constant-coefficient model with  $\beta_0(t) = 0.5$ ,  $\beta_1(t) = 0.05$ ,  $X_{ij0}$  and  $X_{ij1}$  are the same as in **Example 1**. The error  $\varepsilon_i(t_{ij})$  is generated from one of the following distributions:

(C1)  $\varepsilon_i(t_{ij}) \sim N(0, 1)$ ;

(C2)  $\varepsilon_i(t_{ij}) \sim t_{(3)}$ ;

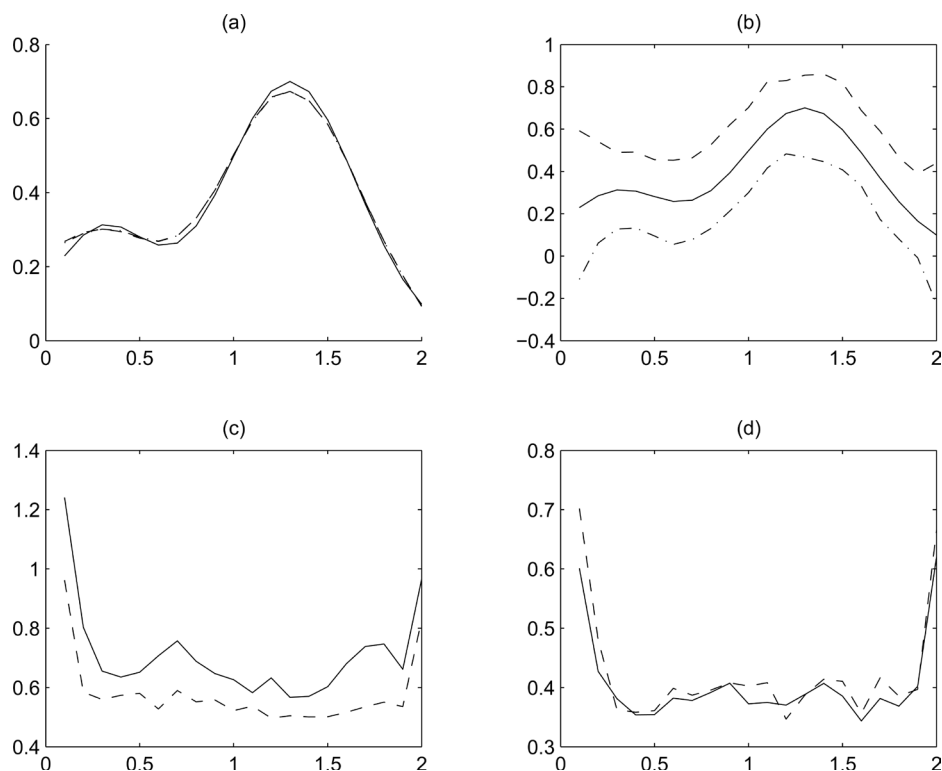
(C3)  $\varepsilon_i(t_{ij}) \sim 0.9N(0, 1) + 0.1N(0, 5^2)$ .

We use a sequence of alternative models,  $\beta_0(t) = 0.5 + \theta t$ ,  $\theta \in [0, 2]$ , to calculate the powers of the tests.

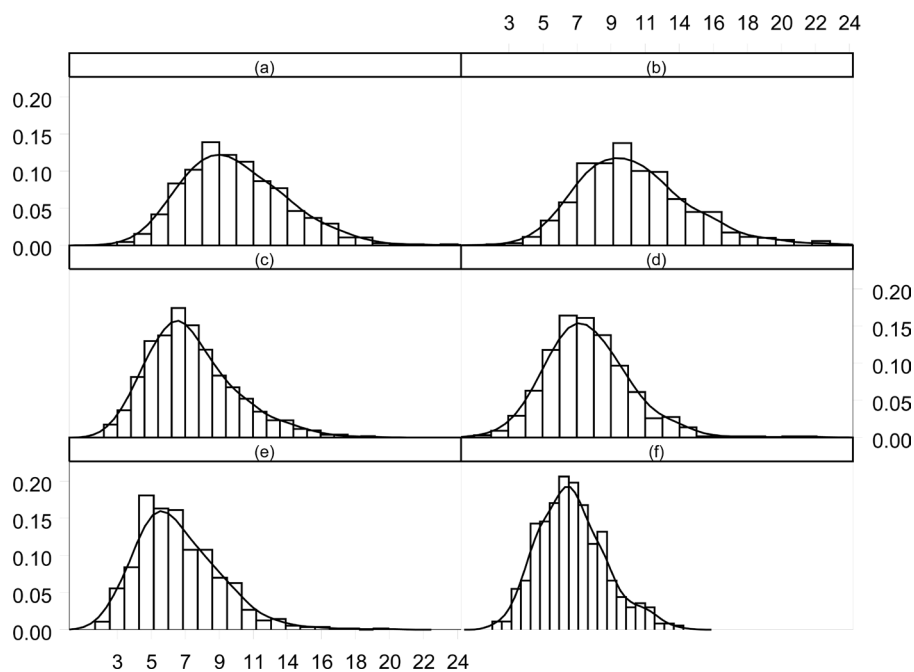
We conducted 400 simulations. The sample size is the same as in **Example 1**. The bootstrap replicates number is  $B = 600$ . **Figure 3** displays the histograms and the estimated probability densities of the testing statistic based on the common kernel method under the null model. It is seen that the null distributions of the testing statistics are very close, which shows that in finite samples the two tests hold close levels and hence have approx-

**Table 2.** MADEs for **Example 2**.

MADE	(B1)		(B2)		(B3)	
	$\hat{\beta}_0(\cdot)$	$\hat{\beta}_1(\cdot)$	$\hat{\beta}_0(\cdot)$	$\hat{\beta}_1(\cdot)$	$\hat{\beta}_0(\cdot)$	$\hat{\beta}_1(\cdot)$
Local LS-estimator	0.0713	0.0529	0.0828	0.0589	0.1408	0.0884
Local M-estimator	0.0733	0.0534	0.0860	0.0534	0.1164	0.0740



**Figure 2.** Estimated curves and envelopes. Upper panel: estimated curves for  $\beta_0(t)$  along with its 95% confidence intervals based on the local M-estimation; solid—true curve, dashed—estimated curves. Lower panel: widths of envelopes formed by 2.5% and 97.5% sample quantiles among simulations, solid—for local LS estimator, dashed—for local M-estimator. (a) (b) (d): the results under the error in (B2); (c): the results under the error in (B3).



**Figure 3.** Histograms and estimated probability densities of  $T_N$ 's. Left panel: based on the local LS-estimator; right panel: based on the local M-estimator with Huber  $\psi_k$ -function. (a) (b): results under the error in (C1); (c) (d): results under the error in (C2); (e) (f): results under the error in (C3).

imately equal type I errors. **Figure 4** shows the powers of the tests based on the local LS and local M-estimators. It demonstrates that the proposed test is more powerful than that based on local LS-estimator when the error deviates away from the normal distribution. Both tests are approximately the same powerful under the normal error.

### 4.3. A Real Example

We here illustrate how to use the proposed method in practice. Consider the body-weight of male Wistar rats dataset in Brunner *et al.* ([30], Table A11). For this dataset, the objective of the experience is to assess the toxicity of a drug on the body-weight evolution of Wistar rats. A group of ten rats was given a placebo, while a second group of ten was given a high dose of the drug. For each rat in the study, its body-weight was observed once a week over a period of 22 weeks. **Figure 5(a)** and **Figure 5(b)** displays the data curves for the two groups.

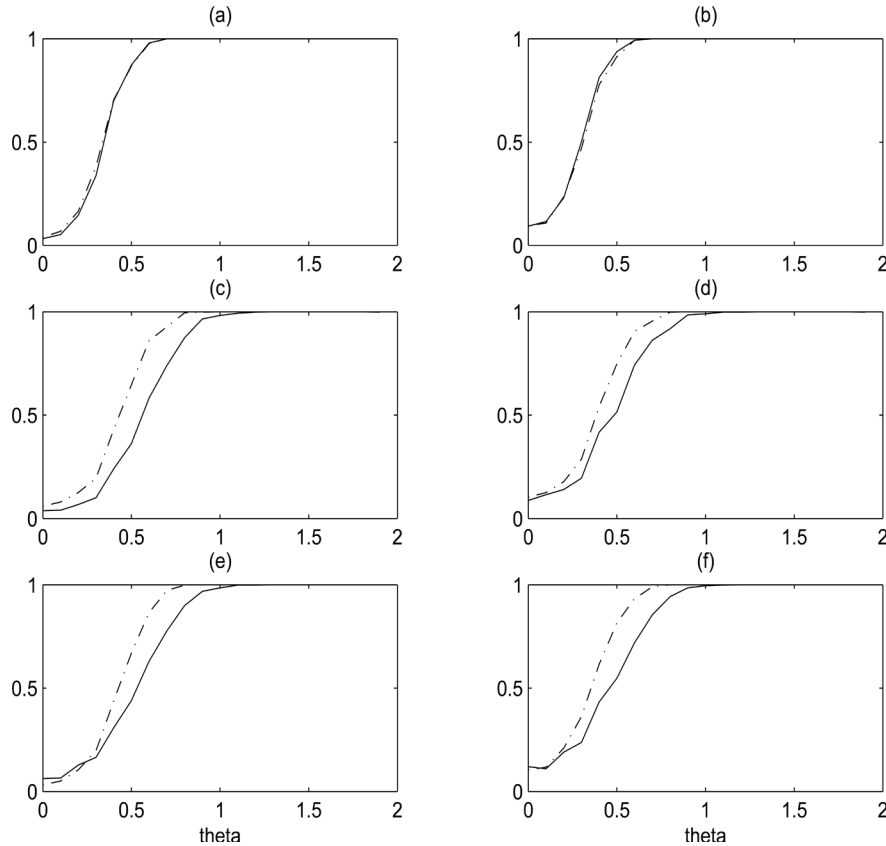
To check if the body-weights of the two test groups differ in their evolution over time, Brunner *et al.* [2] compared the time curves of the mean body-weights for both groups, and evaluated the ANOVA-type statistics. Their results do not support the conjecture of different body-weight evolutions in the two groups.

We are interested in investigating the conjecture. Since **Figure 5(a)** and **Figure 5(b)** exhibit the time-varying feature of the body-weights of rats, it seems reasonable to model the dataset via the following time-varying coefficient model:

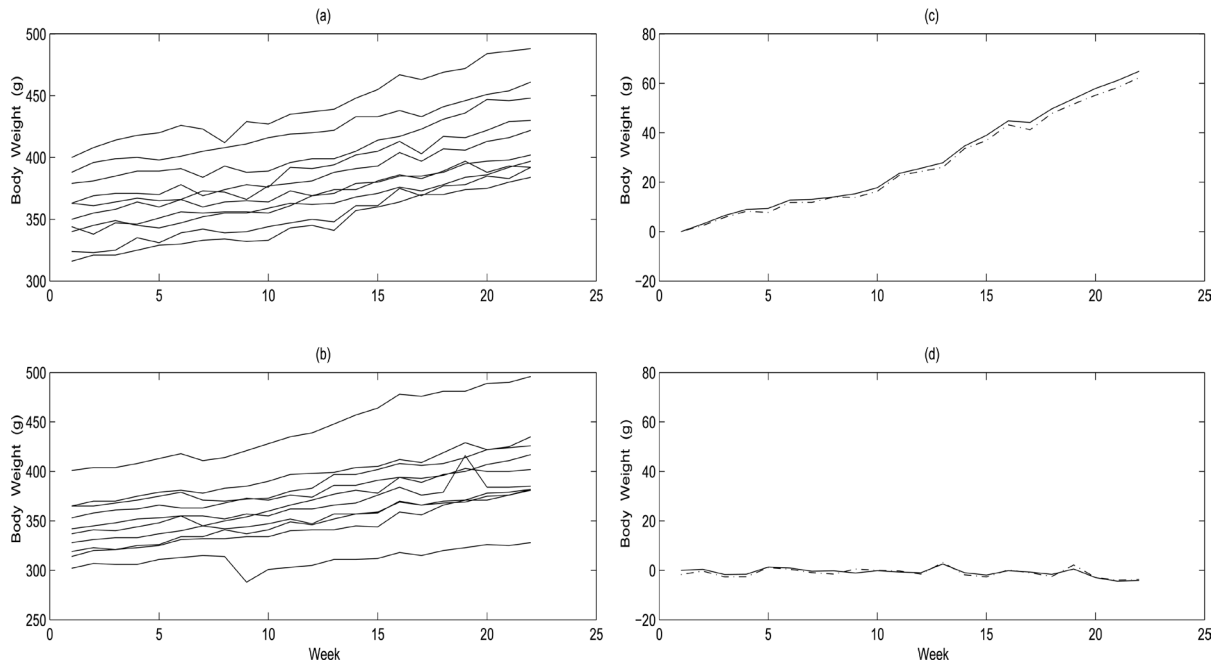
$$Y(t) = \beta_0(t) + X_1(t)\beta_1(t) + \varepsilon(t), \quad (4.10)$$

where  $X_1(t)$  equals 1 if in the treatment group and equals 0 if in the control group. Then  $\beta_0(t)$  reflects the evolution of average body-weight for the control group, and  $\beta_1(t)$  reflects that for the treatment group.

Since in the beginning the average body-weight of the rats in the placebo group is bigger than that in the



**Figure 4.** Powers of  $T_N$ 's. Left panel: powers under significance level 95%; right panel: powers under significance level 90%. (a) (b): results under the error in (C1); (c) (d): results under the error in (C2); (e) (f): results under the error in (C3). Solid - power based on the local LS-estimator; dash-dotted—power based on the local M-estimator with Huber's  $\psi_k$ -function.



**Figure 5.** Data curves and the estimated coefficient functions. Left panel: (a) body-weights for the placebo group; (b) body-weights for treatment group. Right panel: (c) estimated curve for  $\beta_0(t)$ ; (d) estimated curve for  $\beta_1(t)$ ; solid—the local LS estimator, dash-dotted—the local M-estimate.

treatment group, we subtract the average body-weight in each group from the body-weight of each subject. This will make the estimators of the coefficient functions be approximately zeros at the beginning. **Figure 5(c)** and **Figure 5(d)** report the estimators of  $\beta_0(t)$  and  $\beta_1(t)$ . It seems that the  $\beta_0(t)$  is time-varying but the  $\beta_1(t)$  is not, which plausibly reflects that the drug did not affect the average body-weights of rats.

Now we consider the following two hypothesis testing problems:

1)  $H_{01} : \beta(t) = \beta \leftrightarrow H_1 : \beta(t) \neq \beta$ , where  $\beta = (\beta_1, \beta_2)'$  is an unknown parameter vector. This is used to test if the average body-weights of rats evolve over time.

2)  $H_{02} : \beta_1(t) = 0 \leftrightarrow H_1 : \beta_1(t) \neq 0$ . This checks if the drug affects the average body-weights of rats in study.

We used  $B = 600$  bootstrap replicates for computing the null distributions of the testing statistics. For  $H_{01}$ , the  $P$ -values are 0.0117 and 0.005 respectively for the local LS-estimation based test and the local M-estimation based test. It seems that the null hypothesis does not hold, which means that the average body-weight of rats for each group changes over time. For  $H_{02}$ , the  $P$ -values are 0.9633 and 0.8833 for the tests respectively based on the local LS and local M-estimation methods. This evidences that for the dataset the drug has no effect on the average body-weights of rats. This is consistent to the result of Brunner *et al.* [30].

## 5. Discussion

We have introduced a robust inference method based on the local M-estimation method and the robustified *GLR* test. It is demonstrated that the local M-estimators are robust against outliers and error distributions, and the proposed robustified *GLR* test is more powerful than its counterpart (the *GLR* test) under certain situations with heavy tailed errors, while both of them perform well under the normal error. The proposed inference approach seems appealing in robustly modeling longitudinal data.

Our method is also applicable to other estimating methods, such as the global smoothing one in Huang *et al.* [11], but will not be discussed further.

## Acknowledgements

We thank the Editor and the referee for their comments. Supported in part by the NSFC grant 71361010 and by

funds provided by the University of North Carolina at Charlotte. Correspondence should be addressed to Jian-cheng Jiang at University of North Carolina at Charlotte, USA (E-mail: [jjjiang1@unc.edu](mailto:jjjiang1@unc.edu)).

## References

- [1] Diggle, P.J., Liang, K.-Y. and Zeger, S.L. (1994) Analysis of Longitudinal Data. Clarendon Press, Oxford.
- [2] Davidian, M. and Giltinan, D.M. (1995) Nonlinear Models for Repeated Measurement Data. Chapman and Hall, London.
- [3] Vonesh, E.F. and Chinchilli, V.M. (1997) Linear and Nonlinear Models for Analysis of Repeated Measurements. Marcel Dekker, New York.
- [4] Zeger, S.L. and Diggle, P.L. (1994) Semiparametric Models for Longitudinal Data with Application to CD4 Cell Numbers in HIV Seroconverters. *Biometrics*, **50**, 689-699. <http://dx.doi.org/10.2307/2532783>
- [5] Müller, H.-G. (1998) Nonparametric Regression Analysis of Longitudinal Data. Lecture Notes in Statistics, Vol. 46, Springer-Verlag, Berlin.
- [6] Eubank, R.L., Huang, C., Maldonado, Y.M., Wang, N., Wang, S. and Buchanan, R.J. (2004) Smoothing Spline Estimation in Varying-Coefficient Models. *Journal of the Royal Statistical Society*, **66**, 653-667. <http://dx.doi.org/10.1111/j.1467-9868.2004.B5595.x>
- [7] He, X., Fung, W.K. and Zhu, Z.Y. (2005) Robust Estimation in Generalized Partial Linear Models for Clustered Data. *Journal of the American Statistical Association*, **100**, 1176-1184. <http://dx.doi.org/10.1198/016214505000000277>
- [8] He, X., Zhu, Z.Y. and Fung, W.K. (2002) Estimation in a Semiparametric Model for Longitudinal Data with Unspecified Dependence Structure. *Biometrika*, **89**, 579-590. <http://dx.doi.org/10.1093/biomet/89.3.579>
- [9] Hoover, D.R., Rice, J.A., Wu, C.O. and Yang, L.-P. (1998) Nonparametric Smoothing Estimates of Time-Varying Coefficient Models with Longitudinal Data. *Biometrika*, **85**, 809-822. <http://dx.doi.org/10.1093/biomet/85.4.809>
- [10] Fan, J. and Zhang, J.T. (2000) Two-Step Estimation of Functional Linear Models with Applications to Longitudinal Data. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **62**, 303-322. <http://dx.doi.org/10.1111/1467-9868.00233>
- [11] Huang, J., Wu, C. and Zhou, L. (2002) Vary-Coefficient Models and Basis Function Approximations for the Analysis of Repeated Measurements. *Biometrika*, **89**, 111-128. <http://dx.doi.org/10.1093/biomet/89.1.111>
- [12] Lindsey, J.K. (1993) Models for Repeated Measurements. Oxford University Press, Oxford.
- [13] Jones, R.M. (1993) Longitudinal Data with Serial Correlation: A State-Space Approach. Chapman and Hall, London. <http://dx.doi.org/10.1007/978-1-4899-4489-4>
- [14] Hand, D. and Crowder, M. (1996) Practical Longitudinal Data Analysis. Chapman and Hall, London. <http://dx.doi.org/10.1007/978-1-4899-3033-0>
- [15] Brumback, B. and Rice, J.A. (1988) Smoothing Spline Models for the Analysis of Nested and Crossed Samples of Curves (with Discussion). *Journal of the American Statistical Association*, **93**, 961-994. <http://dx.doi.org/10.1080/01621459.1998.10473755>
- [16] Chiang, C.T., Rice, J.A. and Wu, C.O. (2001) Smoothing Spline Estimation for Varying Coefficient Models with Repeatedly Measured Dependent Variables. *Journal of the American Statistical Association*, **96**, 605-619. <http://dx.doi.org/10.1198/016214501753168280>
- [17] Wu, C.O., Chiang, C.T. and Hoover, D.R. (1998) Asymptotic Confidence Regions for Kernel Smoothing of a Varying Coefficient Model with Longitudinal Data. *Journal of the American Statistical Association*, **93**, 1388-1402. <http://dx.doi.org/10.1080/01621459.1998.10473800>
- [18] Wu, C.O. and Chiang, C.T. (2000) Kernel Smoothing on Varying Coefficient Models with Longitudinal Dependent Variable. *Statistica Sinica*, **10**, 433-456.
- [19] Fan, J. and Jiang, J. (2000) Variable Bandwidth and One-Step Local M-Estimator. *Science in China Series A*, **43**, 65-81. <http://dx.doi.org/10.1007/BF02903849>
- [20] Jiang, J. and Mack, Y.P. (2001) Robust Local Polynomial Regression for Dependent Data. *Statistica Sinica*, **11**, 705-722.
- [21] Huber, P.J. (1964) Robust Estimation of a Location Parameter. *The Annals of Mathematical Statistics*, **35**, 73-101. <http://dx.doi.org/10.1214/aoms/1177703732>
- [22] Serfling, R.J. (1980) Approximation Theorems of Mathematical Statistics. Wiley, New York. <http://dx.doi.org/10.1002/9780470316481>
- [23] Jiang, J., Jiang, X. and Song, X. (2014) Weighted Composite Quantile Regression Estimation of DTARCH Models.

---

*Econometrics Journal*, **17**, 1-23. <http://dx.doi.org/10.1111/ectj.12023>

- [24] Hastie, T.J. and Tibshirani, R.J. (1990) Generalized Additive Models. Chapman and Hall, New York.
- [25] Fan, J., Zhang, C.-M. and Zhang, J. (2001) Generalized Likelihood Ratio Statistics and Wilks Phenomenon. *The Annals of Statistics*, **29**, 153-193. <http://dx.doi.org/10.1214/aos/996986505>
- [26] Fan, J. and Huang, T. (2005) Profile Likelihood Inferences on Semiparametric Varying-Coefficient Partially Linear Models. *Bernoulli*, to Appear. <http://dx.doi.org/10.3150/bj/1137421639>
- [27] Fan, J. and Jiang, J. (2005) Nonparametric Inferences for Additive Models. *Journal of the American Statistical Association*, **100**, 890-907. <http://dx.doi.org/10.1198/016214504000001439>
- [28] Fan, J. and Jiang, J. (2007) Nonparametric Inference with Generalized Likelihood Ratio Tests (with Discussion). *Test*, **16**, 409-444. <http://dx.doi.org/10.1007/s11749-007-0080-8>
- [29] Hui, Y.V. and Jiang, J. (2005) Robust Modelling of DTARCH Models. *The Econometrics Journal*, **8**, 143-158. <http://dx.doi.org/10.1111/j.1368-423X.2005.00157.x>
- [30] Brunner, E., Domhof, S. and Langer, F. (2002) Nonparametric Analysis of Longitudinal Data in Factorial Experiments. Wiley, New York.