Published in final edited form as:

Comput Stat Data Anal. 2017 December ; 116: 49–66. doi:10.1016/j.csda.2017.07.005.

A Continuous Threshold Expectile Model

Feipeng Zhang^{a,b} and Qunhua Li^{a,*}

^aDepartment of Statistics, Pennsylvania State University, PA, 16802, USA

^bDepartment of Statistics, Hunan University, Changsha, 410082, China

Abstract

Expectile regression is a useful tool for exploring the relation between the response and the explanatory variables beyond the conditional mean. A continuous threshold expectile regression is developed for modeling data in which the effect of a covariate on the response variable is linear but varies below and above an unknown threshold in a continuous way. The estimators for the threshold and the regression coefficients are obtained using a grid search approach. The asymptotic properties for all the estimators are derived, and the estimator for the threshold is shown to achieve root- n consistency. A weighted CUSUM type test statistic is proposed for the existence of a threshold at a given expectile, and its asymptotic properties are derived under both the null and the local alternative models. This test only requires fitting the model under the null hypothesis in the absence of a threshold, thus it is computationally more efficient than the likelihood-ratio type tests. Simulation studies show that the proposed estimators and test have desirable finite sample performance in both homoscedastic and heteroscedastic cases. The application of the proposed method on a Dutch growth data and a baseball pitcher salary data reveals interesting insights. The proposed method is implemented in the **R** package *cthreshER*.

Keywords

Expectile regression; Threshold; Weighted CUSUM test; Grid search method

1. Introduction

Expectile regression, first introduced by Aigner et al. (1976) and Newey and Powell (1987), has become popular in the last decades. Analogous to quantile regression (Koenker and Bassett, 1978), expectile regression draws a complete picture of the conditional distribution of the response variable given the covariates, making it a useful tool for modeling data with heterogeneous conditional distributions. As modeling tools, quantile regression and expectile regression both have advantages over the other in certain aspects: quantile regression is more robust to outliers than expectile regression, whereas expectile regression is more sensitive to

*Corresponding author at: Department of Statistics, The Pennsylvania State University, PA, 16802, USA. Fax: 814-863-7114. qunhua.li@psu.edu. (Qunhua Li).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

the extreme values in the response variable than quantile regression. However, expectile regression has certain computational advantages over quantile regression (Newey and Powell, 1987). First, unlike quantile regression, the loss function of expectile regression is everywhere differentiable, thus its estimation is more straightforward and much faster. Second, the computation of the asymptotic covariance matrix of the expectile regression estimator does not involve estimating the density function of the errors. Because of these features, many nonparametric or semiparametric expectile regression have been developed in recent years, for example, Yao and Tong (1996), De Rossi and Harvey (2009), Kuan et al. (2009), Schnabel and Eilers (2009), Kneib (2013), Sobotka et al. (2013), Xie et al. (2014), Schulze Waltrup et al. (2015), Kim and Lee (2016), and among others. These models greatly improve the flexibility of expectile regression for modeling nonlinear relationships.

However, some natural phenomena call for nonlinear regression forms that exhibit structural changes, sometimes in the form of two line segments with different slopes. For example, a child's height increases rapidly with age before and during puberty and then stops increasing in late teens. This implies that the growth curve of height may be described as two line segments with different slopes intersecting at a threshold. Another example arises from a study of the salaries of major league baseball players in 1987 (Hoaglin and Velleman, 1995). The data shows a positive correlation between salaries and years of experience for less experienced pitchers but a negative correlation for more experienced pitchers. In these instances, besides the regression coefficients, the onset of the transition point is often of great research interest, for example, when a child reaches his/her full adult height or when is the prime time for pitchers' salaries. Although the existing spline-based models (e.g., Schnabel and Eilers, 2009; Kim and Lee, 2016) or varying-coefficient expectile models (e.g., Xie et al., 2014) can capture the nonlinear relationship between the response variable and the predictors, they cannot provide information on the location of the threshold. This issue motivates us to consider a continuous threshold model for expectile regression. Continuous threshold regression, also called segmented regression or bent line regression, has been studied in the context of least squares regression (Quandt, 1958, 1960; Hinkley, 1969; Feder, 1975; Chappell, 1989; Chan and Tsay, 1998; Chiu et al., 2006; Hansen, 2017), quantile regression (Li et al., 2011), and rank-based regression (Zhang and Li, 2017). However, no literature has investigated continuous threshold models in the context of expectile regression.

In this article, we develop a continuous threshold expectile regression model. The contribution of this article is twofold. First, we propose a grid search method to estimate the unknown threshold and other regression coefficients. We derive the asymptotic properties for all the parameters including the threshold, and show that the estimator for the threshold achieves $\sqrt{n}$ -consistency. Second, we develop a testing procedure for the existence of structural change at a given expectile, based on a weighted CUSUM type statistic. This test only requires fitting the model under the null hypothesis in the absence of a threshold, thus it is computationally efficient. The limiting distribution of the test statistic is also established. The comparison with a sup-likelihood-ratio-type test (Lee et al., 2011) in simulation studies shows that our test has a better control of type I error and a higher power in a wide range of

scenarios. All the estimation and testing procedures are implemented in **R** code, which is available from the **R** package *ctthreshER* (Zhang and Li, 2016).

The remainder of the article is organized as follows. In Section 2, we describe the continuous threshold expectile regression model, and develop a grid search method for estimating the unknown threshold and regression coefficients. A testing procedure for the structural change at a given expectile level is also proposed. In Section 3, we conduct simulation studies and two real data analyses. Section 4 provides the conclusion with possible future extensions. Technical proofs are presented in the Appendix.

2. Methodology

2.1. Model

Let (Y_i, X_i, Z_i) , $i = 1, \dots, n$, be a sequence of independent and identically distributed sample from the population (Y, X, Z) . We assume that Y is the response variable, Z is a vector of covariates, and X is a scalar variable, whose relationship with Y changes at an unknown location. The population τ -expectile of Y , $\nu_\tau(Y)$, minimizes the loss function $E[\rho_\tau(Y - \nu)]$, where

$$\rho_\tau(u) = \omega_\tau(u)u^2 = \begin{cases} (1 - \tau)u^2, & u \leq 0, \\ \tau u^2, & u > 0, \end{cases}$$

is the asymmetric squared error loss function, and $0 < \tau < 1$ is the parameter that controls the degree of loss asymmetry. Clearly, when $\tau = 0.5$, the τ -expectile corresponds to the mean of Y .

In this paper, we model the conditional τ -th expectile of Y using the continuous threshold model

$$\nu_\tau(Y|X, Z) = \beta_0 + \beta_1 X + \beta_2 (X - t)_+ + \gamma^\top Z, \quad (1)$$

where t is the unknown location of the threshold, $\xi = (\beta_0, \beta_1, \beta_2, \gamma^\top)^\top$ is the vector of parameters excluding t , and γ is a $p \times 1$ vector of parameters. Here, $a_+ = aI(a > 0)$ is the positive part of a , where $I(\cdot)$ is the indicator function. Clearly, the linear expectile regression is continuous on X at t , but has different slopes on either side of the threshold t . In other words, β_1 is the slope of the left line segment for $X < t$ and $\beta_1 + \beta_2$ is the slope of the right line segment for $X > t$.

2.2. Estimation procedure

To estimate $\theta = (\xi^\top, t)^\top$ at a given expectile τ , we minimize the objective function

$$M_{n,\tau}(\theta) = n^{-1} \sum_{i=1}^n \rho_\tau(Y_i - \beta_0 - \beta_1 X_i - \beta_2 (X_i - t)_+ - \gamma^\top Z_i). \quad (2)$$

However, due to the existence of the threshold t , the objective function (2) is convex in ξ but non-convex in t , making it difficult to obtain its minimizer. One possible approach is to use the grid search strategy, which is commonly used for bent line mean regression (Quandt, 1958; Chappell, 1989). To proceed, we re-write the objective function (2) with respect to ξ and t as

$$M_{n,\tau}(\theta) \equiv M_{n,\tau}(\xi, t) = n^{-1} \sum_{i=1}^n \rho_{\tau}(Y_i - \xi^{\top} V_i(t)), \quad (3)$$

where $V_i(t) = (1, X_i, (X_i - t)_+, Z_i^{\top})^{\top}$. The minimization can be carried out in two steps:

1. for each $t \in \mathcal{T}$, where $\mathcal{T}$ is the range set of all t 's, obtain a profile estimate of ξ by $\hat{\xi}(t) = \arg \min_{\xi} M_{n,\tau}(\xi, t)$.
2. obtain the threshold estimate of t by $\hat{t} = \arg \min_{t \in \mathcal{T}} M_{n,\tau}(\hat{\xi}(t), t)$.

The estimate for θ then is $\hat{\theta} = (\hat{\xi}(\hat{t})^{\top}, \hat{t})^{\top}$.

2.3. Asymptotic properties

Because the objective function is not differentiable with respect to θ , it is impossible to obtain the asymptotic properties of $\hat{\theta}$ using the standard theory. Here, we derive the asymptotic properties using the modern empirical processes theory. We first introduce some notations. Denote the true parameters as θ_0 . Let $M_{\tau}(\theta) = E \rho_{\tau}(Y - \xi^{\top} V(t))$, where $V(t) = (1, X, (X - t)_+, Z^{\top})^{\top}$. Using the notation of empirical process, one can write

$$M_{n,\tau}(\theta) = \mathbb{P}_n m_{\theta} \quad \text{and} \quad M_{\tau}(\theta) = P m_{\theta},$$

where $\mathbb{P}_n = n^{-1} \sum_{i=1}^n \delta_{\mathcal{X}_i}$ is the empirical measure, and $m_{\theta}(\mathcal{X}) = \rho_{\tau}(Y - \xi^{\top} V(t)) = \omega_{\tau}[Y - \xi^{\top} V(t)]^2$ with the weights

$$\omega_{\tau} \equiv \omega_{\tau}(\mathcal{X}) = |\tau - I(Y - \xi^{\top} V(t) \leq 0)| = \begin{cases} (1 - \tau), & Y - \xi^{\top} V(t) \leq 0, \\ \tau, & Y - \xi^{\top} V(t) > 0. \end{cases}$$

Here, $\mathcal{X}$ is the observed data (Y, X, Z) .

In Lemma A.1 in the Appendix, we show that $\sup_{\theta \in \Theta} |M_{n,\tau}(\theta) - M_{\tau}(\theta)|$ converges to zero in probability, as n goes to infinity. Furthermore, we establish the consistency of $\hat{\theta}$.

Theorem 2.1—Under the regularity conditions in the Appendix, as $n \rightarrow \infty$, we have

$$\hat{\theta} \xrightarrow{P} \theta_0.$$

We prove the asymptotic normality by using Theorem 5.23 in Van der Vaart (2000), which establishes the asymptotic normality of M-estimators when the criterion function is

Lipschitz continuous and its limiting function admits a second order Taylor expansion. To proceed, define the matrix $\sum(\boldsymbol{\theta}) = E \dot{m}_{\boldsymbol{\theta}} \dot{m}_{\boldsymbol{\theta}}^{\top}$, where $\dot{m}_{\boldsymbol{\theta}}$ is

$$\dot{m}_{\boldsymbol{\theta}} = \begin{bmatrix} -2\omega_{\tau} \mathbf{V}(t) \{Y - \boldsymbol{\xi}^{\top} \mathbf{V}(t)\} \\ 2\beta_2 E \left\{ \omega_{\tau} [Y - \boldsymbol{\xi}^{\top} \mathbf{V}(t)] \middle| X \right\} I(X > t) \end{bmatrix}.$$

Define the Hessian matrix of $M_{\tau}(\boldsymbol{\theta})$ as

$$\begin{aligned} H(\boldsymbol{\theta}) &\equiv \frac{\partial^2}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^{\top}} M_{\tau}(\boldsymbol{\theta}) \\ &= 2E \left(\omega_{\tau} \begin{bmatrix} \mathbf{V}(t) \mathbf{V}(t)^{\top} & -\beta_2 I(X > t) \mathbf{V}(t) + \{Y - \boldsymbol{\xi}^{\top} \mathbf{V}(t)\} \mathbf{U}(t) \\ -\beta_2 I(X > t) \mathbf{V}(t)^{\top} + \{Y - \boldsymbol{\xi}^{\top} \mathbf{V}(t)\} \mathbf{U}(t)^{\top} & \beta_2^2 I(X > t) \end{bmatrix} \right) \\ &\quad + 2E \left(\begin{bmatrix} \mathbf{0}_{(p+3) \times (p+3)} & \mathbf{0}_{(p+3) \times 1} \\ \mathbf{0}_{1 \times (p+3)} & -\beta_2 E \left\{ \omega_{\tau} [Y - \boldsymbol{\xi}^{\top} \mathbf{V}(t)] \middle| X = t \right\} f_X(t) \end{bmatrix} \right), \end{aligned}$$

where $\mathbf{U}(t) = [0, 0, I(X > t), \mathbf{0}_{p \times 1}]^{\top}$.

Theorem 2.2—Under the regularity conditions in the Appendix, $\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)$ is asymptotically normally distributed with mean zero and covariance matrix $H(\boldsymbol{\theta}_0)^{-1} \Sigma(\boldsymbol{\theta}_0) H(\boldsymbol{\theta}_0)^{-1}$, as $n \rightarrow \infty$.

It is worthwhile to emphasize that the regression coefficients and threshold estimators $(\hat{\boldsymbol{\xi}}^{\top}, \hat{t})^{\top}$ are jointly asymptotically normal with a convergence rate of $\sqrt{n}$, and have non-zero asymptotic covariance in our model setting. This is different from the model with a discontinuous threshold. In the latter situation, the estimators of the regression coefficients $\hat{\boldsymbol{\xi}}$ are still $\sqrt{n}$ -consistent, but the threshold estimator $\hat{t}$ is n -consistent with a non-standard asymptotic distribution. The $\sqrt{n}$ -convergence rate of $\hat{t}$ in our model is due to the continuity of $M_{n,\tau}(\boldsymbol{\theta})$ at t .

The asymptotic covariance matrix can be estimated by $\hat{H}_n(\hat{\boldsymbol{\theta}})^{-1} \hat{\Sigma}(\hat{\boldsymbol{\theta}}) \hat{H}_n(\hat{\boldsymbol{\theta}})^{-1}$, where

$$\hat{\Sigma}_n(\hat{\boldsymbol{\theta}}) = n^{-1} \sum_{i=1}^n \hat{G}_n(\hat{\boldsymbol{\theta}}) \hat{G}_n(\hat{\boldsymbol{\theta}})^{\top},$$

$$\hat{G}_n(\hat{\boldsymbol{\theta}}) = \begin{bmatrix} -2\hat{\omega}_{\tau,i} \mathbf{V}_i(t) \left\{ Y_i - \hat{\boldsymbol{\xi}}^{\top} \mathbf{V}_i(t) \right\} \\ 2\hat{\beta}_2 \hat{\omega}_{\tau,i} \left\{ Y_i - \hat{\boldsymbol{\xi}}^{\top} \mathbf{V}_i(t) \right\} I(X_i > t) \end{bmatrix},$$

and

$$\hat{H}_n(\hat{\theta}) = \frac{2}{n} \sum_{i=1}^n \hat{\omega}_{\tau,i} \begin{bmatrix} \mathbf{V}_i(t) \mathbf{V}_i(t)^\top, & -\hat{\beta}_2 I(X_i > t) \mathbf{V}_i(t) + \{Y_i - \hat{\xi}^\top \mathbf{V}_i(t)\} \mathbf{U}_i(t) \\ -\hat{\beta}_2 I(X_i > t) \mathbf{V}_i(t)^\top + \{Y_i - \hat{\xi}^\top \mathbf{V}_i(t)\} \mathbf{U}_i(t)^\top, & \hat{\beta}_2^2 I(X_i > t) \end{bmatrix} \\ + \begin{bmatrix} \mathbf{0}_{(p+3) \times (p+3)}, & \mathbf{0}_{(p+3) \times 1} \\ \mathbf{0}_{1 \times (p+3)}, & -\frac{2}{nh} \sum_{i=1}^n \hat{\beta}_2 \hat{\omega}_{\tau,i} \left\{ Y_i - \hat{\xi}^\top \mathbf{V}_i(t) \right\} K\left(\frac{X_i - t}{h}\right) \end{bmatrix}.$$

Here, $\hat{\omega}_{\tau,i} = |\tau - \mathbb{I}(Y_i - \hat{\xi}^\top \mathbf{V}_i(t) \leq 0)|$, and $K(\cdot)$ is a kernel function with a bandwidth $h > 0$. In practice, we use the Epanechnikov kernel $K(u) = 3/4(1 - u^2)\mathbb{I}(|u| \leq 1)$ and obtain the optimal bandwidth by the Silverman's rule of thumb (Silverman, 1986), $h = 1.06\hat{\sigma}n^{-1/5}$, where $\hat{\sigma}$ is the standard deviation of X .

2.4. Testing the existence of a threshold at a given expectile

An important question before fitting model (1) is whether there exists a threshold at a pre-specified expectile. If a threshold does not exist, then t is unidentifiable and the estimation procedure in the last section is ill-conditioned. To test the existence of a threshold, we test

$$H_0: \beta_2 = 0 \text{ for any } t \in \mathcal{T} \text{ v.s. } H_1: \beta_2 \neq 0 \text{ for some } t \in \mathcal{T},$$

where $\mathcal{T}$ is the range set of all t 's.

Tests for structural changes have been developed in conditional mean regression (Andrews, 1993; Bai, 1996; Hansen, 1996, 2017), quantile regression (Qu, 2008; Li et al., 2011), transformation models (Kosorok and Song, 2007), time series models (Chan, 1993; Cho and White, 2007), and among others. To construct our test statistic, we take an approach in spirit similar to the test for structural changes in quantile regression in Qu (2008). This test is constructed by sequentially evaluating the subgradients of the objective function under H_0 for a subsample, in a fashion similar to the CUSUM statistic. An advantage of this test is that it only requires fitting the model under the null hypothesis. Thus, it is computationally more efficient than the likelihood-ratio type tests, such as the sup-likelihood-ratio-type test for testing threshold effects in regression models in Lee et al. (2011), which requires fitting the models under both null and alternative hypotheses.

To proceed, we define the following statistic,

$$R_n(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^n |\tau - \mathbb{I}(Y_i \leq \hat{\alpha}^\top \mathbf{W}_i)| (Y_i - \hat{\alpha}^\top \mathbf{W}_i) (X_i - t) \mathbb{I}(X_i \leq t),$$

where $\mathbf{W}_i = (1, X_i, \mathbf{Z}_i^\top)^\top$, and $\hat{\alpha}$ is the estimator of the coefficients $\alpha = (\beta_0, \beta_1, \gamma^\top)^\top$ under the null hypothesis H_0 , that is,

$$\hat{\alpha} = \arg \min_{\alpha} \frac{1}{\sqrt{n}} \sum_{i=1}^n |\tau - \mathbb{I}(Y_i \leq \alpha^\top \mathbf{W}_i)| (Y_i - \alpha^\top \mathbf{W}_i)^2.$$

An intuitive interpretation for $R_n(t)$ is given as follows. If there is not a threshold, $\hat{\alpha}$ is a good estimate of its population value, and hence, the estimated residual $e_i = Y_i - \hat{\alpha}^T \mathbf{W}_i$ would show a random pattern against X_i , leading to a small $R_n(t)$. On the other hand, if there exists a threshold, the estimate $\hat{\alpha}$ would differ significantly from the true value, and the estimated residuals would depart from zero in a systematic fashion related to X_i , resulting in a large absolute value of $R_n(t)$. Because the location of the threshold is unknown, we need search through all the possible locations. Therefore, we propose the test statistic

$$T_n = \sup_{t \in \mathcal{T}} |R_n(t)|.$$

This statistic can be viewed as a weighted CUSUM statistic based on the estimated residuals under the null hypothesis. Intuitively, it is plausible to reject H_0 when T_n is too large. This intuition will be formally verified by Theorem 2.5. It implies that $R_n(t)$ converges to a Gaussian process with mean zero, and the size of such a process can be used to test for a threshold effect.

In order to derive the large-sample inference for T_n , we consider the local alternative model,

$$Y_i = \beta_0 + \beta_1 X_i + n^{-1/2} \beta_2 (X_i - t)_+ + \gamma^\top \mathbf{Z}_i + e_i, \quad (4)$$

where t is the location of threshold, $\beta_2 \neq 0$, and the τ -expectile of e_i is zero. We first introduce some notations

$$\hat{S}_{\text{wn}}(\hat{\alpha}) = n^{-1} \sum_{i=1}^n |\tau - I(Y_i \leq \hat{\alpha}^\top \mathbf{W}_i)| \mathbf{W}_i \mathbf{W}_i^\top,$$

$$S_w(\alpha) = E[|\tau - I(Y \leq \alpha^\top \mathbf{W}_1)| \mathbf{W}_1 \mathbf{W}_1^\top],$$

$$\hat{S}_{1n}(\hat{\alpha}, t) = n^{-1} \sum_{i=1}^n |\tau - I(Y_i \leq \hat{\alpha}^\top \mathbf{W}_i)| \mathbf{W}_i (X_i - t) I(X_i \leq t),$$

$$S_1(\alpha, t) = E[|\tau - I(Y \leq \alpha^\top \mathbf{W}_1)| \mathbf{W}_1 (X - t) I(X \leq t)],$$

$$\hat{S}_{2n}(\hat{\alpha}, t) = n^{-1} \sum_{i=1}^n |\tau - I(Y_i \leq \hat{\alpha}^\top \mathbf{W}_i)| \mathbf{W}_i \beta_2 (X_i - t) I(X_i \geq t),$$

$$S_2(\boldsymbol{\alpha}, t) = E[|\tau - I(Y \leq \boldsymbol{\alpha}^\top \mathbf{W}_1)| \mathbf{W}_1 \beta_2(X - t) I(X \geq t)],$$

$$\text{and } q(t) = S_1(\boldsymbol{\alpha}, t)^\top S_w(\boldsymbol{\alpha})^{-1} S_2(\boldsymbol{\alpha}, t).$$

Theorem 2.3—Under the regularity conditions in the Appendix, for the local alternative model (4), $R_n(t)$ has the asymptotic representation

$$R_n(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^n e_i [\tau - I(Y_i - \boldsymbol{\alpha}^\top \mathbf{W}_i \leq 0)] [(X_i - t) I(X_i \leq t) - S_1(\boldsymbol{\alpha}, t)^\top S_w(\boldsymbol{\alpha})^{-1} \mathbf{W}_i] - q(t) + o_p(1).$$

(5)

Furthermore, T_n converges weakly to the process $\sup_t |R(t) - q(t)|$, where $R(t)$ is the Gaussian process with mean zero and covariance function

$$E \left[e_1^2 |\tau - I(Y_1 - \boldsymbol{\alpha}^\top \mathbf{W}_1 \leq 0)| \{ (X_1 - t_1) I(X_1 \leq t_1) - S_1(\boldsymbol{\alpha}, t_1)^\top S_w(\boldsymbol{\alpha})^{-1} \mathbf{W}_1 \} \right. \\ \left. \times \{ (X_1 - t_2) I(X_1 \leq t_2) - S_1(\boldsymbol{\alpha}, t_2)^\top S_w(\boldsymbol{\alpha})^{-1} \mathbf{W}_1 \} \right].$$

Corollary 2.4—Under the regularity conditions in the Appendix, for the local alternative model, $Y_i = \beta_0 + \beta_1 X_i + n^{-1/2} a_n \beta_2 (X_i - t)_+ + \boldsymbol{\gamma}^\top \mathbf{Z} + e_i$ for any increasing sequence a_n goes to infinite, we have that $\lim_{n \rightarrow \infty} P(|T_n| \geq t) = 1$ for any $t > 0$.

Because the limiting null distribution of T_n is nonstandard, we resort to the Gaussian multiplier method (Van der Vaart, 2000) to calculate the critical values, based on the asymptotic representation (5). The procedure is described in Algorithm 1. In the Appendix, we prove the following result, which implies the validity of the bootstrap resampling scheme.

Theorem 2.5—Under both the null and the local alternative hypotheses, $R_n^*(\tau)$ (defined in Algorithm 1) converges to the Gaussian process $R(t)$ as $n \rightarrow \infty$.

Algorithm 1

- 1 Generate iid $\{v_1, \dots, v_n\}$ from $N(0, 1)$.
- 2 Calculate the test statistic $T_n^* = \sup_{t \in \mathcal{T}} |R_n^*(t)|$, where

$$R_n^*(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^n v_i \hat{e}_i |\tau - I(\hat{e}_i \leq 0)| \\ \times \left[(X_i - t) I(X_i \leq t) - \hat{S}_{1n}(\hat{\alpha}, t)^\top \hat{S}_{wn}(\hat{\alpha})^{-1} \mathbf{W}_i \right],$$

with the estimated residuals $\hat{e}_i = Y_i - \hat{\alpha}^\top \mathbf{W}_i$ under the null hypothesis.

3 Repeat Steps 1–2 with NB times to obtain $T_n^{*(1)}, \dots, T_n^{*(NB)}$. Calculate the p-value as

$$\hat{p}_n = \text{NB}^{-1} \sum_{j=1}^{\text{NB}} I\{T_n^{*(j)} \geq T_n\}.$$

3. Simulation Studies and Applications

3.1. Simulation studies

In this section, we conduct simulation studies for assessing the finite sample performance of the proposed methods. In particular, we evaluate the accuracy of estimation, model fitting, and type I error and power of the proposed test. We consider the following two scenarios:

- i. Independent and identically distribution (IID): $Y = \beta_0 + \beta_1 X + \beta_2 (X - t)_+ + \gamma Z + e$,
- ii. Heteroscedasticity: $Y = \beta_0 + \beta_1 X + \beta_2 (X - t)_+ + \gamma Z + (1 + 0.2X)e$,

where x is generated from a uniform distribution $U(-2, 4)$, z is generated from a normal distribution $N(1, 0.5^2)$, and the parameters are $(\beta_0, \beta_1, \beta_2, \gamma, t)^\top = (1, 3, -2, 1, 1.5)^\top$. For each scenario, we consider three error distributions: (1) $e \sim N(0, 1)$, (2) $e \sim t_4$, and (3) a mixture distribution $e \sim 0.9N(0, 1) + 0.1t_4$, where t_4 is the t -distribution with four degrees of freedom. For each case, we conduct 1000 repetitions with sample sizes $n = 200$ and 1000, and fit the model at $\tau = 0.05, 0.1, 0.25, 0.5, 0.75, 0.9$ and 0.95.

Accuracy of estimation—The bias, empirical standard error, average estimated standard errors and 95% coverage probabilities (CP) are summarized in the Supplementary Tables C.1–C.2. Figures 1–2 display the average of the estimates and the 95% average confidence intervals (CI), which are constructed using the average of the estimate and the estimated standard error over all repetitions, as well as the coverage probabilities. As shown in Figures 1–2, all the estimates show little bias in all cases. In general, the CPs of the regression parameters $(\beta_0, \beta_1, \beta_2, \gamma)$ are close to the nominal level 95%, except when the sample size is small ($n = 200$) and the expectile is close to 0 or 1. The relatively low coverage in those cases is likely due to very few observations at extreme expectiles. When the sample size increases to $n = 1000$, their CPs increase to the nominal level. The CPs of threshold t are generally lower than those of the regression parameters, with most values in the range of 80–90% when $n = 200$ (Supplementary Tables C.1–C.2). A close examination shows that the relatively low coverage tends to occur in the heteroscedastic cases or the error cases with long tails. This is expected, as the estimates in these cases tend to have larger variation than

their IID or shorter-tail counterparts. When the sample size increases to $n = 1000$, the coverage probabilities of t also improve, with most CPs above 90%. In summary, the proposed estimate has a good finite sample performance.

Prediction accuracy—Next, we evaluate the performance of model fitting by comparing our method with the smooth expectile regression in Schnabel and Eilers (2009). The smooth expectile regression fits smooth expectile curves by combining the least asymmetrically weighted squares with the P-splines. Although it does not provide information on the existence or the location of the threshold, it is flexible in fitting nonlinear trend and can be used to fit segmented relationship.

To proceed, we use the same simulation setting as the above and perform 1000 repetitions with a sample size of $n = 200$. The smooth expectile regression is fitted using the **R** package *expectreg* (Sobotka et al., 2016). To evaluate model fitting, we compute the prediction errors for each method using a K-fold cross-validation. Specifically, we divide the data into K equal-sized subgroups, denoted as D_k for $k = 1, \dots, K$. The k th prediction error is given by

$$PE_k = \sum_{i \in D_k} \rho_\tau \left(Y_i - \hat{Y}_i^{(-k)} \right),$$

where $\hat{Y}_i^{(-k)}$ are fitted values from the model estimated based on the data without D_k . The total prediction error is defined as $PE = \sum_{k=1}^K PE_k$. Here we choose $K = 10$. As shown in Table 1, our method has smaller average PEs than the smoothed expectile regression in all cases.

Performance of the threshold test—To evaluate the performance of the testing procedure, we conduct another set of simulations using the simulation model above. In this simulation, the threshold effects are set at $\beta_2 = -1, -0.5, -0.2, -0.1, 0, 0.1, 0.2, 0.5, 1$ and all the other parameters are kept as before. For each case, we perform 500 repetitions with sample sizes of $n = 200$ and obtain p-values with 500 bootstrap replicates (i.e., $NB = 500$ in Algorithm 1). We also derive a sup-likelihood-ratio-type test proposed by Lee et al. (2011) in the context of the continuous threshold expectile model (See details in Appendix B) and compare its performance with our method.

The results are summarized in Supplementary Tables C.3–C.4 and illustrated in Figure 3. For better visualization, we only show a subset of results for $\beta_2 = 0, 0.2, 0.5$ and 1. The results from negative β_2 's follow a similar trend. As shown in Figure 3, both methods have type I errors reasonably close to the nominal level and a reasonable power to detect threshold effects. The control of type I error rate and the power are generally better for non-extreme expectiles than extreme expectiles. This is due to few observations at extreme expectiles and it can be improved with a larger sample size. As expected, the power for the heteroscedastic cases or the errors with long tails is lower than that for their IID or shorter-tailed counterparts. Compared with Lee's test, our test not only shows a higher power in all cases, but also has a better control of type I errors, especially at extreme expectiles. Our

method is especially advantageous in heteroscedastic cases. In addition, our test is computationally more efficient than Lee's test. On average it is about 6 times faster than Lee's test.

3.2. Applications

3.2.1. Fourth Dutch growth data—We first apply our method to the Fourth Dutch Growth data, which was collected by the Fourth Dutch Growth study (van Buuren, 2007) and is available in the **R** package *expectreg* (Sobotka et al., 2016). This dataset has the height, weight and head circumference of Dutch children between the ages of 0 and 21 years (van Buuren and Fredriks, 2001). A primary interest of this study concerns the relation between age and height. Figure 4a shows the relationship between age and height for a subset of 6,848 boys. Clearly, there is a nonlinear trend between height and age, with a steep curvature before age three due to rapid growth in early childhood, and a bent in the late teens due to reaching full adult height. This dataset has been analyzed by Schnabel and Eilers (2009). In their analysis, they took a square root transformation on age. While this transformation effectively removes the curvature at early childhood, the nonlinearity in the late teens still exists after transformation (Figure 4b). Then they fitted the transformed data using the smoothed expectile regression to take account of this nonlinearity. Though the smoothed expectile curves fit the data well, they do not provide any information on the location of the threshold, i.e., the age to stop growing.

Here, we re-analyze the data using the continuous threshold model and estimate the location of the threshold. Following Schnabel and Eilers (2009), we take a square root transformation on the age prior to model fitting. This transformation helps ensure the validity of the linearity assumption before and after the threshold. Specifically,

$$\nu_{\tau}(Y_i|X_i, \mathbf{Z}) = \beta_0 + \beta_1 X_i + \beta_2 (X_i - t)_+, \quad (6)$$

where Y_i is the height of the i th boy, X_i is the square root of his age, $\theta = (\beta_0, \beta_1, \beta_2)$ are the unknown parameters of interest, and t is the unknown location of threshold. We fit the model with $\tau = 0.05, 0.10, 0.25, 0.50, 0.75, 0.90$, and 0.95 , and perform the threshold test with 1000 bootstrap replicates.

For all the expectile levels we fit, the p-values from our threshold test are nearly 0, indicating a highly significant continuous threshold pattern. The estimated regression parameters for different expectile levels are reported in Table 2. The estimated coefficients show that the height first increases rapidly with age (roughly 31–35 cm per square root of age), and then the growth becomes very limited or nearly stops after age 17–18. Interestingly, the estimated thresholds show a general trend that shorter boys seem to stop growth later than taller boys, with a 95% confidence interval (CI), [18.56, 19.04] at the expectile level $\tau = 0.05$, and [17.24, 17.59] at $\tau = 0.95$. Figure 4b illustrates these results.

As a comparison, we also fit the data with the bent-line quantile regression (Li et al., 2011). Figure 4c shows the fitted quantile curves of Y on X . Though the overall trend of these curves is similar to that of the continuous threshold expectile regression, we observe that the

estimated threshold in the quantile regression is far more dependent on the quantile level than that in the expectile regression on the expectile level. This can be explained by how data are used in the fitting of these two models. In quantile regression, given a quantile, only the data below the quantile are used in the estimation; whereas, in expectile regression, all data are involved in the estimation at any expectile level. Therefore, a higher level of dependence is observed in the quantile regression than in the expectile regression.

3.2.2. Baseball pitcher salary—Our second example concerns the salaries of major league baseball (MLB) players for the 1987 baseball season (Hoaglin and Velleman, 1995). The full dataset was analyzed as an ASA data analysis challenge in 1989 by fifteen competing groups. Here we consider a subset with $n = 176$ pitchers, which was analyzed by Hettmansperger and McKean (2011) using a rank-based regression. This subset of data is available in the **R** package *Rfit* (Kloke and McKean, 2016). It consists of the 1987 beginning salary, the number of years of experience, and several performance measures in 1986, including average wins per year, average losses per year, earned run average, average games per year, average innings per year and average saves per year.

Visually, the scatter plot of log salary and log number of years of experience (Figure 5a) suggests that the salaries are first positively correlated with the years of experience, but then decline after about $\exp(2.3) = 10$ years. This pattern is somewhat unusual, because it is generally expected that salaries will grow with years of experience in players' early careers and after they become free agents (i.e., players whose initial 6 year contracts expire). Although salaries can decrease after players pass their prime time, the onset of decline is expected to happen much later. For example, Hauptert and Murray (2012) estimated the salary decline for MLB players occurs after 22 years, based on the MLB data for over 60 years (1911 through 1973). Interestingly, among the analyses conducted by the 15 ASA groups, the model with the best predictive performance is a segmented mean regression model with a fixed threshold at 7 years. However, the threshold in this model was chosen according to the length of the initial professional baseball contracts (6 years), rather than determined by the data. It is of interest to formally test if the visually observed transition is significant and to estimate the onset of the salary decline from the data. Furthermore, the salaries at a given number of years of experience show considerable heterogeneity. Previous analyses only focused on the mean regression model (Hettmansperger and McKean, 2011; Hoaglin and Velleman, 1995), ignoring the heterogeneity. A regression model based on the conditional distribution of the response variable would provide a more complete picture and a better exploration of the heterogeneity than a mean regression model.

Here we fit the data using the continuous threshold expectile regression. Following Hettmansperger and McKean (2011), we first take logarithm transformation for salaries and years of experience to ensure the validity of the linearity assumption before and after the threshold. Specifically,

$$\nu_{\tau}(Y_i|X_i, \mathbf{Z}_i) = \beta_0 + \beta_1 X_i + \beta_2 (X_i - t)_+ + \boldsymbol{\gamma}^T \mathbf{Z}_i, \quad (7)$$

where Y_i is the log (salary) of the i th pitcher, X_i is log (years of experience), Z_i is the vector of performance statistics, $\theta = (\beta_0, \beta_1, \beta_2, \gamma^T)^T$ are the unknown parameters of interest, t is the unknown location of threshold, and $\tau = 0.05, 0.10, 0.25, 0.50, 0.90$ and 0.95 . We first fit this model (7) with all the performance statistics. The estimated coefficients show that only average wins per year (Z_{i1}) and earned run average (ERA) (Z_{i2}) are statistically significant among all performance statistics (p-values < 0.05). Hence, only these two performance statistics are kept in our final model, i.e. $Z_i = (Z_{i1}, Z_{i2})$. We perform the threshold test with 1000 bootstrap times.

Table 3 reports the estimated coefficients and their standard errors. The coefficients show that average wins per year positively affects salaries and ERA negatively affects salaries for all pitchers. This indicates both team performance (average wins per year) and individual performance (ERA) affect salaries. Interestingly, individual performance (ERA) has a much stronger effect on salaries among high-salary pitchers (-0.405 at $\tau = 0.95$) than low-salary pitchers (-0.282 at $\tau = 0.05$); whereas team performance (average wins per year) has a bigger effect on salaries among low-salary pitchers (0.115 at $\tau = 0.05$) than high-salary pitchers (0.062 at $\tau = 0.95$). These agree with general salary rules in sports.

Besides the performance statistics, salaries also depend on years of experience: it first increases with years of experience ($\beta_1 > 0$ at all expectiles) and then decreases ($\beta_2 < 0$ at all expectiles), with a highly significant threshold effect (p-value < 0.05) at all the expectile levels considered. The coefficients show that salaries decline for pitchers with about 10 or more years of experience (range: (9.35, 10.14) for different expectiles), at all the expectile levels we fitted. Figure 5a displays the fitted expectile curves of Y on X at the mean values of Z_1 and Z_2 (7.64 and 3.67, respectively).

This raises two questions: why did the salaries decline rather earlier than the prime time estimated from the data in 1931–1973 (Hauptert and Murray, 2012)? and why did the decrease occur at around 10 years for all salary levels? The history of MLB shows that, in the time period of 1985 to 1987, the MLB team owners colluded in an effort to decrease salaries for free agents after their initial contracts expired. Pitchers with 10 or more years of experience were all free agents. Their salary decrease is a reflection of owners trying to control salaries. The reason that the observed threshold (10 years) is later than the start of free agents (7 years), is that some pitchers had become free agents before the collusion, thus they had more than 7 years of experience when the collusion occurred.

As a comparison, we also fit the data with the bent-line quantile regression (Li et al., 2011). Figure 5b shows the fitted quantile curves of Y on X at the mean values of Z_1 and Z_2 (7.64 and 3.67, respectively). Similar to the previous example, we also observe that the estimated threshold is more dependent on the quantile level than that in the expectile regression on the expectile level. In addition, we noticed that there are more crossings between quantile curves than between expectile curves. This agrees with the observation of Schnabel and Eilers (2009) and Schulze Waltrup et al. (2015) that expectile regression tends to have fewer crossings than quantile regression.

4. Concluding Remarks

In this article, we develop the continuous threshold expectile regression model. This model allows the expectiles of the response variable to be piecewise linear but still continuous in covariates. We develop a grid search method to estimate the unknown threshold and the regression coefficients. A weighted CUSUM type test statistics is proposed to test the structural change at a given expectile level. Our numerical studies show that both the proposed estimator and the proposed test have good finite sample performance.

Our work may be extended in several ways. First, although generally there are fewer crossings in expectile regression than in quantile regression (Schnabel and Eilers, 2009), the expectile crossings may happen. It will be worthwhile to extend our model to non-crossing continuous threshold expectile estimation and to develop tests for structural change across expectiles. Another interesting extension is to consider more than one threshold for a covariate. In such a situation, the estimation and test of the thresholds would be more complicated, and further investigation is needed.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors are grateful to the three anonymous referees for many helpful comments. The authors thank Dr. Andrew Wiesner for the assistance in the interpretation of the baseball data. Li and Zhang are supported in part by NIH R01GM109453. Zhang is also partially supported by the National Natural Science Foundation of China (11401194), the Natural Science Foundation of Hunan Province, China (2017JJ3021) and the Fundamental Research Funds for the Central Universities (227201305039).

Appendix A

Proofs

Regularity Conditions.

- (A1) $t_0 = \arg \min_{t \in \mathcal{T}} M_{\tau}(\hat{\xi}(t), t)$ is unique, where $\mathcal{T}$ is a compact set in $\mathbb{R}^1$.
- (A2) θ is in Θ , and Θ is a compact subset of $\mathbb{R}^{p+4}$.
- (A3) The scalar variable X has an absolutely continuous distribution with density function f_X , which is strictly positive, bounded and continuous for any t in a neighborhood of t_0 .
- (A4) Both $E[\sup_{\theta} |V(t)[Y - \xi^{\top} V(t)]|^2]$ and $E[\sup_{\theta} |\beta_2 I(X > t)[Y - \xi^{\top} V(t)]|^2]$ are finite, and $E|Y|^4 < \infty$, $E|X|^4 < \infty$, and $E\|Z\|^4 < \infty$.
- (A5) Given $\beta_2 = 0$, the Hessian matrix $H(\theta_0)$ is nonsingular.

Condition (A1) is the identifiability condition of the estimation. Conditions (A1)–(A5) are used for the consistency and the asymptotic normality of the estimates, under both homoskedasticity and heteroskedasticity cases.

We first provide the following uniformly convergence results.

Lemma A.1

Under the regular conditions, $\sup_{\theta \in \Theta} |M_{n,\tau}(\theta) - M_\tau(\theta)| \xrightarrow{P} 0$, as $n \rightarrow \infty$.

Proof of Lemma A.1—To show that the class of functions $\{m_\theta: \theta \in \Theta\}$ is Glivenko-Cantelli, it is sufficient to use Lemma 2.9 in Newey and McFadden (1994). Recalling that $\theta = (\xi^\top, t)^\top$, and the derivatives

$$\frac{\partial m_\theta(\mathcal{X})}{\partial \xi} = -2\omega_\tau \mathbf{V}(t)[Y - \xi^\top \mathbf{V}(t)],$$

$$\frac{\partial m_\theta(\mathcal{X})}{\partial t} = 2\omega_\tau \beta_2 I(X > t)[Y - \xi^\top \mathbf{V}(t)].$$

Note that $w_\tau = \max(\tau, 1 - \tau) < 1$ for any $\tau \in (0, 1)$. By the Conditions (A2) and (A4), we can show that $E \sup_{\theta} |\mathbf{V}(t)[Y - \xi^\top \mathbf{V}(t)]|$ and $E \sup_{\theta} |\beta_2 I(X > t)[Y - \xi^\top \mathbf{V}(t)]|$ are finite. Hence, applying the mean-value theorem, there is a $\theta^* = (\xi^{*\top}, t^*)^\top$, such that

$$M_{n,\tau}(\theta_1) - M_{n,\tau}(\theta_2) = \frac{1}{n} \sum_{i=1}^n \begin{bmatrix} \mathbf{V}_i(t^*)[Y_i - \xi^{*\top} \mathbf{V}_i(t^*)] \\ \beta_2^* I(X_i > t^*)[Y_i - \xi^{*\top} \mathbf{V}_i(t^*)] \end{bmatrix} (\theta_1 - \theta_2),$$

for every $\mathcal{X}$. Note that

$$\begin{aligned} & E \left\| \frac{1}{n} \sum_{i=1}^n \begin{bmatrix} \mathbf{V}_i(t^*)[Y_i - \xi^{*\top} \mathbf{V}_i(t^*)] \\ \beta_2^* I(X_i > t^*)[Y_i - \xi^{*\top} \mathbf{V}_i(t^*)] \end{bmatrix} \right\| \\ & \leq E \begin{bmatrix} \sup_{\theta} |\mathbf{V}(t)[Y - \xi^\top \mathbf{V}(t)]| \\ \sup_{\theta} |\beta_2 I(X > t)[Y - \xi^\top \mathbf{V}(t)]| \end{bmatrix} < \infty, \end{aligned}$$

by Conditions (A2) and (A4). Thus,

$$B_n \equiv \begin{bmatrix} \sup_{\theta} |\mathbf{V}(t)[Y - \xi^\top \mathbf{V}(t)]| \\ \sup_{\theta} |\beta_2 I(X > t)[Y - \xi^\top \mathbf{V}(t)]| \end{bmatrix} = O_P(1).$$

Then, $|M_{n,\tau}(\theta_1) - M_{n,\tau}(\theta_2)| \leq B_n \|\theta_1 - \theta_2\|$ for every $\mathcal{X}$, where $B_n = O_P(1)$. It is easy to show that $M_{n,\tau}(\theta) \xrightarrow{P} M_\tau(\theta)$ for all $\theta \in \Theta$. The conclusion follows from Lemma 2.9 in Newey and McFadden (1994).

Proof of Theorem 2.1—By Lemma A.1, $\sup_{\theta \in \Theta} |M_{n,\tau}(\theta) - M_\tau(\theta)| \xrightarrow{P} 0$ as n goes to infinity. Since Θ is compact and the uniqueness of the minimum θ_0 (by Conditions A1 and A2), along with that $M_{n,\tau}(\theta)$ is continuous with respect to θ , then we can establish that $\hat{\theta} \xrightarrow{P} \theta_0$ by Theorem 2.1 of Newey and McFadden (1994).

Proof of Theorem 2.2—Firstly, by Condition (A3), the function $\mathcal{X} \mapsto m_\theta(\mathcal{X})$ is measurable, and the function $\theta \mapsto m_\theta(\mathcal{X})$ is differentiable at θ_0 for P-almost every $\mathcal{X}$. We can show that $|m_{\theta_1}(\mathcal{X}) - m_{\theta_2}(\mathcal{X})| \leq m(\mathcal{X}) \|\theta_1 - \theta_2\|$ for every $\mathcal{X}$, where $E m(\mathcal{X})^2 < \infty$. Indeed, recalling that $\theta = (\xi^\top, t)^\top$, the derivatives

$$\frac{\partial m_\theta}{\partial \xi} = -2\omega_\tau V(t)[Y - \xi^\top V(t)],$$

$$\frac{\partial m_\theta}{\partial t} = 2\omega_\tau \beta_2 I(X > t)[Y - \xi^\top V(t)].$$

Note that $w_\tau = \max(\tau, 1 - \tau) < 1$ for any $\tau \in (0, 1)$. By the Conditions (A2) and (A4), it is easy to show that $E |\max V(t) [Y - \xi^\top V(t)]|$ and $E \max |\beta_2 I(X > t)|$ are finite. Hence, applying the mean-value theorem, $|m_{\theta_1}(\mathcal{X}) - m_{\theta_2}(\mathcal{X})| \leq m(\mathcal{X}) \|\theta_1 - \theta_2\|$ for every $\mathcal{X}$, where

$$E m(\mathcal{X})^2 = E \left[\frac{(\sup |V(t)[Y - \xi^\top V(t)]|)^2}{(\sup |\beta_2 I(X > t)[Y - \xi^\top V(t)]|)^2} \right] < \infty,$$

by Conditions (A2) and (A4).

Secondly, the map $\theta \mapsto M_\tau(\theta) = E m_\theta$ admits a second order Taylor expansion at θ_0 , with a nonsingular symmetric Hessian matrix $H(\theta_0)$. We can verify that $H(\theta)$ is continuous in θ . Indeed, the elements of $H(\theta)$ are quadratic functions of ξ , and hence $H(\theta)$ is continuous in ξ . It is sufficient to show that $H(\theta)$ is continuous in t . Note that the first term of $H(\theta)$ is a function of t through moments of the form $E[V(t)I(X > t)]$. By Condition (A4), $(E\|V(t)\|^2)^{1/2} \leq C_1$ for some constant $C_1 < \infty$. By Condition (A3), for any $t_1 < t_2$, we have $|F_X(t_1) - F_X(t_2)| \leq \max_x |f_X(x)| |t_2 - t_1| \leq C_2 |t_2 - t_1|$ for some constant $C_2 < \infty$. Then, by Cauchy-Schwartz inequality,

$$\begin{aligned} E\|V(t)I(t_1 \leq X \leq t_2)\|^2 &\leq (E\|V(t)\|^2)^{1/2} (E|t_1 \leq X \leq t_2|^2)^{1/2} \\ &\leq C_1 C_2 |t_1 - t_2|^{1/2}, \end{aligned}$$

is uniformly continuous in t . Hence, the first term of $H(\boldsymbol{\theta})$ is continuous in t . On the other hand, since $E\omega_\tau = \tau(1 - F_Y(\boldsymbol{\xi}^\top \mathbf{V}(t))) + (1 - \tau)F_Y(\boldsymbol{\xi}^\top \mathbf{V}(t))$ is continuous in t , then the second term of $H(\boldsymbol{\theta})$ is continuous in t . Thus, $H(\boldsymbol{\theta})$ is continuous in t .

Finally, by Theorem 2.1, $\hat{\boldsymbol{\theta}}$ is consistent for $\boldsymbol{\theta}_0$ in a neighborhood of $\boldsymbol{\theta}_0$, it follows that

$\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)$ is asymptotically normal with mean zero and covariance matrix $H(\boldsymbol{\theta}_0)^{-1}\Sigma(\boldsymbol{\theta}_0)H(\boldsymbol{\theta}_0)^{-1}$, by Theorem 5.23 in Van der Vaart (2000).

Lemma A.2

Under the regularity conditions, as $n \rightarrow \infty$, we have

- i. $\hat{S}_{\text{wn}}(\hat{\boldsymbol{\alpha}}) \xrightarrow{P} S_w(\boldsymbol{\alpha})$
- ii. $\sup_t |\hat{S}_{1n}(\hat{\boldsymbol{\alpha}}, t) - S_1(\boldsymbol{\alpha}, t)| \xrightarrow{P} 0$.
- iii. $\sup_t |\hat{S}_{2n}(\hat{\boldsymbol{\alpha}}, t) - S_2(\boldsymbol{\alpha}, t)| \xrightarrow{P} 0$.

Proof of Lemma A.2—It is easily obtained by using the law of large number for (i). To establish (ii) and (iii), note that $\hat{S}_{1n}(\hat{\boldsymbol{\alpha}}, t)$ and $\hat{S}_{2n}(\hat{\boldsymbol{\alpha}}, t)$ are sums of indicator functions and Lipschitz functions, hence they are Glivenko-Cantelli class, which implies that both (ii) and (iii) holds.

Proof of Theorem 2.3—Recall that

$\hat{\boldsymbol{\alpha}} = \arg \min_{\boldsymbol{\alpha}} n^{-1/2} \sum_{i=1}^n |\tau - I(Y_i \leq \boldsymbol{\alpha}^\top \mathbf{W}_i)| (Y_i - \boldsymbol{\alpha}^\top \mathbf{W}_i)^2$, which is equivalent to the solution of the estimating equation

$$U_n(\boldsymbol{\alpha}) = \frac{1}{\sqrt{n}} \sum_{i=1}^n |\tau - I(Y_i \leq \boldsymbol{\alpha}^\top \mathbf{W}_i)| \mathbf{W}_i (Y_i - \boldsymbol{\alpha}^\top \mathbf{W}_i).$$

Then, under the local alternative model (4), the estimating equation can be written as

$$\begin{aligned} U_n(\boldsymbol{\alpha}) &= \frac{1}{\sqrt{n}} \sum_{i=1}^n |\tau - I(Y_i \leq \boldsymbol{\alpha}^\top \mathbf{W}_i)| \mathbf{W}_i e_i \\ &+ \frac{1}{n} \sum_{i=1}^n |\tau - I(Y_i \leq \boldsymbol{\alpha}^\top \mathbf{W}_i)| \mathbf{W}_i \beta_2 (X_i - t) I(X_i > t) + o_p(1). \end{aligned}$$

By the mean-value theorem, we have

$$\begin{aligned} -U_n(\boldsymbol{\alpha}) &= U_n(\hat{\boldsymbol{\alpha}}) - U_n(\boldsymbol{\alpha}) \\ &= -\frac{1}{\sqrt{n}} \sum_{i=1}^n \left| \tau - I(Y_i \leq \hat{\boldsymbol{\alpha}}^{*\top} \mathbf{W}_i) \right| \mathbf{W}_i \mathbf{W}_i^\top (\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}) + o_p(1) \\ &= -\hat{S}_{\text{wn}}(\hat{\boldsymbol{\alpha}}^*) \sqrt{n}(\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}) + o_p(1). \end{aligned}$$

where $\hat{\alpha}^*$ lies in the line between $\hat{\alpha}$ and α . By Lemma A.2, $\hat{S}_{wn}(\hat{\alpha}) \xrightarrow{P} S_w(\alpha)$, and under the local alternative model 4, it yields that

$$\begin{aligned}\sqrt{n}(\hat{\alpha} - \alpha) &= \frac{1}{\sqrt{n}} S_w(\alpha)^{-1} \sum_{i=1}^n |\tau - I(Y_i \leq \alpha^\top \mathbf{W}_i)| \mathbf{W}_i (Y_i - \alpha^\top \mathbf{W}_i) + o_P(1) \\ &= \frac{1}{\sqrt{n}} S_w(\alpha)^{-1} \sum_{i=1}^n |\tau - I(Y_i \leq \alpha^\top \mathbf{W}_i)| \mathbf{W}_i e_i \\ &\quad + \frac{1}{\sqrt{n}} S_w(\alpha)^{-1} \sum_{i=1}^n |\tau - I(Y_i \leq \alpha^\top \mathbf{W}_i)| \mathbf{W}_i \beta_2(X_i - t) I(X_i > t) + o_P(1).\end{aligned}$$

Thus, under the local alternative model 4, by plugging in the representation for $\sqrt{n}(\hat{\alpha} - \alpha)$ and some algebraic manipulation, we have

$$\begin{aligned}R_n(t) &= \frac{1}{\sqrt{n}} \sum_{i=1}^n |\tau - I(Y_i \leq \alpha^\top \mathbf{W}_i)| e_i \left[(X_i - t) I(X_i \leq t) - \hat{S}_{1n}(\alpha, t)^\top \hat{S}_{wn}(\alpha)^{-1} \mathbf{W}_i \right] \\ &\quad - \hat{S}_{1n}(\alpha, t)^\top \hat{S}_{wn}(\alpha)^{-1} \hat{S}_{2n}(t, \alpha) + o_P(1) \\ &= \frac{1}{\sqrt{n}} \sum_{i=1}^n |\tau - I(Y_i \leq \alpha^\top \mathbf{W}_i)| e_i \left[(X_i - t) I(X_i \leq t) - \hat{S}_{1n}(\alpha, t)^\top \hat{S}_{wn}(\alpha)^{-1} \mathbf{W}_i \right] \\ &\quad - q(t) + o_P(1)\end{aligned}$$

It is easy to derive the remainder conclusion for weak convergence of $R_n(\hat{\alpha}, t)$ by following the proofs in Stute (1997).

Proof of Corollary 2.4—The proof follows the same lines as that for Theorem 2.3 and thus is omitted.

Proof of Theorem 2.5—We divide the proof into three steps. Firstly, we show that the covariance function of R_n^* converges to that of R . Define

$$\begin{aligned}R_n^*(t) &= \frac{1}{\sqrt{n}} \sum_{i=1}^n v_i (Y_i - \alpha^\top \mathbf{W}_i) |\tau - I(Y_i \leq \alpha^\top \mathbf{W}_i)| \\ &\quad \times [(X_i - t) I(X_i \leq t) - S_1(\alpha, t)^\top S_w(\alpha)^{-1} \mathbf{W}_i].\end{aligned}$$

By the fact that the consistency of $\hat{\alpha} - \alpha$, along with the uniform convergence of $\hat{S}_{1n}(\hat{\alpha}, t) - S_1(\alpha, t)$ and $\hat{S}_{wn}(\hat{\alpha}) - S_w(\alpha)$, one can easily show $R_n^*(t)$ and $R_n^{**}(t)$ are asymptotically equivalent in the sense that $\sup_t \|R_n^*(t) - R_n^{**}(t)\| = o_P(1)$. Note that v_i 's are independent of (Y_i, X_i, Z_i) , and $E v_i = 0$, $\text{Var}(v_i) = 1$. Then, for any t_1, t_2 , the covariance function of R_n^{**} is

$$\begin{aligned}
& \text{Cov}(R_n^{**}(t_1), R_n^{**}(t_2)) \\
&= \frac{1}{n} \sum_{i=1}^n \mathbb{E} \left(v_i^2 e_i^2 | \tau - I(e_i \leq 0) |^2 \{ (X_i - t_1) I(X_i \leq t_1) - S_1(\boldsymbol{\alpha}, t_1)^T S_w(\boldsymbol{\alpha})^{-1} \mathbf{W}_i \} \right. \\
&\quad \times \left. \{ (X_i - t_2) I(X_i \leq t_2) - S_1(\boldsymbol{\alpha}, t_2)^T S_w(\boldsymbol{\alpha})^{-1} \mathbf{W}_i \} \right) \\
&= \mathbb{E} \left[e^2 | \tau - I(e \leq 0) |^2 \{ (X - t_1) I(X \leq t_1) - S_1(\boldsymbol{\alpha}, t_1)^T S_w(\boldsymbol{\alpha})^{-1} \mathbf{W} \} \right. \\
&\quad \times \left. \{ (X - t_2) I(X \leq t_2) - S_1(\boldsymbol{\alpha}, t_2)^T S_w(\boldsymbol{\alpha})^{-1} \mathbf{W} \} \right],
\end{aligned}$$

which is the same as the covariance of $R(t)$.

Secondly, it is easily to show that any finite-dimensional projection of $R_n^{**}(t)$ converges to that of $R(t)$, by the central limit theorem.

Thirdly, $R_n^{**}(t)$ is uniformly tight. Note that the class of all indicator functions $I(X \leq t)$ is a Vapnik-Chervonenskis (VC) class of functions. Then, the class of functions $\mathcal{F}_n = \{ (X_i - t) I(X_i \leq t) - S_{1n}(t) S_w^{-1} \mathbf{W}_i : t \in \mathbb{R}^1 \}$ is a VC class of functions. Thus, by the equicontinuity lemma 15 of (Pollard, 1984), one can show that $R_n^{**}(\tau)$ is uniformly tight. Then, by the Cramer-Wold device, the proof of Theorem 2.5 is completed.

Appendix B

Sup-likelihood-ratio-type test

In this section, we derive a competing test for structural change at a given expectile, based on the general testing framework proposed by Lee et al. (2011). Recall that $\boldsymbol{\xi} = (\beta_0, \beta_1, \beta_2, \boldsymbol{\gamma}^\top)^\top$ and $\boldsymbol{\alpha} = (\beta_0, \beta_1, \boldsymbol{\gamma}^\top)^\top$. According to Lee et al. (2011), the quasi-likelihood-ratio (QLR) test statistic takes the form of

$$\text{QLR} = n \left(-\hat{M}_{n,\tau} - (-\tilde{M}_{n,\tau}) \right),$$

where $\hat{M}_{n,\tau} = M_{n,\tau}(\hat{\boldsymbol{\theta}})$, and $\tilde{M}_{n,\tau} = M_{n,\tau}(\tilde{\boldsymbol{\xi}}, t)$ with $\tilde{\boldsymbol{\xi}} = \arg \max_{\boldsymbol{\alpha}: \beta_2=0} [-M_{n,\tau}(\boldsymbol{\xi}, t)]$. Note that $M_{n,\tau}(\boldsymbol{\xi}, t)$ does not depend on t under $H_0 : \beta_2 = 0$. Based on Corollary 1 of Lee et al. (2011), we can show that the limiting null distribution of the QLR statistic for the continuous threshold expectile model is

$$\frac{1}{2} \left[\sup_t \mathcal{G}(t)^\top B(t)^{-1} \mathcal{G}(t) - \mathcal{G}_1^\top B_1^{-1} \mathcal{G}_1 \right],$$

where $\mathcal{G}$ is a Gaussian process with mean zero and covariance matrix

$$2E[\omega_{\tau}(t_1)\omega_{\tau}(t_2)V(t_1)V(t_2)^{\top}(Y - \xi^{\top}V(t_1))(Y - \xi^{\top}V(t_2))],$$

$B(t) = 2E[\omega_{\tau}(t)V(t)V(t)^{\top}]$ with $\omega_{\tau}(t) = |\tau - I(Y < \xi^{\top}V(t))|$, and $\mathcal{G}_1$ and B_1 are the parts of $\mathcal{G}(t)$ and the submatrix of $B(t)$ corresponding to $\beta_2 = 0$, respectively.

We summarize the computation procedure for the p-values in Algorithm 2.

Algorithm 2

1 Generate iid $\{v_1, \dots, v_n\}$ from $N(0, 1)$.

2 Calculate the test statistic

$$QLR^* = \frac{1}{2} \left[\sup_t G_n(t)^{\top} \hat{B}(t)^{-1} G_n(t) - G_{n1}^{\top} \hat{B}_1^{-1} G_{n1} \right],$$

where

$$G_n(t) = \frac{2}{\sqrt{n}} \sum_{i=1}^n V_i(t) \left[Y_i - \hat{\xi}(t)^{\top} V_i(t) \right] |\tau - I(Y_i < \hat{\xi}(t)^{\top} V_i(t))| v_i,$$

$$G_{n1}(t) = \frac{2}{\sqrt{n}} \sum_{i=1}^n W_i(t) [Y_i - \hat{\alpha}^{\top} W_i(t)] |\tau - I(Y_i < \hat{\alpha}^{\top} W_i(t))| v_i,$$

$$\hat{B}(t) = \frac{2}{\sqrt{n}} \sum_{i=1}^n V_i(t) V_i(t)^{\top} |\tau - I(Y_i < \hat{\xi}(t)^{\top} V_i(t))|,$$

$$\hat{B}_1(t) = \frac{2}{\sqrt{n}} \sum_{i=1}^n W_i(t) W_i(t)^{\top} |\tau - I(Y_i < \hat{\alpha}^{\top} W_i(t))|.$$

$\hat{\alpha}$ is estimated under the null hypothesis, i.e., $\hat{\alpha}$ is the subvector of $\hat{\xi}$ after deleting the element of β_2 .

3 Repeat Steps 1–2 with NB times to obtain $QLR^{*(1)}, \dots, QLR^{*(NB)}$. Calculate the p-value as

$$\hat{p}_n = NB^{-1} \sum_{j=1}^{NB} I\{QLR^{*(j)} \geq QLR\}.$$

Appendix C

Supplementary data

Supplementary tables can be found online at <http://dx.doi.org/>.

References

- Aigner D, Amemiya T, Poirier DJ. On the estimation of production frontiers: Maximum likelihood estimation of the parameters of a discontinuous density function. *International Economic Review*. 1976; 17:377–96.
- Andrews D. Tests for parameter instability and structural change with unknown change point. *Econometrica*. 1993; 61:821–856.
- Bai J. Testing for parameter constancy in linear regressions: an empirical distribution function approach. *Econometrica*. 1996; 64:597–622.
- van Buuren S. Worm plot to diagnose fit in quantile regression. *Statistical Modelling*. 2007; 7:363–376.
- van Buuren S, Fredriks M. Worm plot: a simple diagnostic device for modelling growth reference curves. *Statistics in Medicine*. 2001; 20:1259–1277. [PubMed: 11304741]
- Chan KS. Consistency and limiting distribution of the least squares estimator of a threshold autoregressive model. *Annals of Statistics*. 1993; 21:520–533.
- Chan KS, Tsay RS. Limiting properties of the least squares estimator of a continuous threshold autoregressive model. *Biometrika*. 1998; 85:413–426.
- Chappell R. Fitting bent lines to data, with applications to allometry. *Journal of Theoretical Biology*. 1989; 138:235–256. [PubMed: 2607772]
- Chiu G, Lockhart R, Routledge R. Bent-cable regression theory and applications. *Journal of the American Statistical Association*. 2006; 101:542–553.
- Cho JS, White H. Testing for regime switching. *Econometrica*. 2007; 75:1671–1720.
- De Rossi G, Harvey A. Quantiles, expectiles and splines. *Journal of Econometrics*. 2009; 152:179–185.
- Efron B. Regression percentiles using asymmetric squared error loss. *Statistica Sinica*. 1991; 1:93–125.
- Feder PI. On asymptotic distribution theory in segmented regression problems—identified case. *The Annals of Statistics*. 1975; 3:49–83.
- Hansen BE. Inference when a nuisance parameter is not identified under the null hypothesis. *Econometrica*. 1996; 64:413–430.
- Hansen BE. Regression kink with an unknown threshold. *Journal of Business and Economic Statistics*. 2017:228–240.
- Hauptert M, Murray J. Regime switching and wages in major league baseball under the reserve clause. *Biometrika*. 2012; 6:143–162.
- Hettmansperger, T., McKean, JW. *Robust Nonparametric Statistical Methods*. 2. New York: Chapman; 2011.
- Hinkley DV. Inference about the intersection in two-phase regression. *Biometrika*. 1969; 56:495–504.
- Hoaglin DC, Velleman PF. A critical look at some analyses of major league baseball salaries. *The American Statistician*. 1995; 49:277–285.
- Kim M, Lee S. Nonlinear expectile regression with application to value-at-risk and expected shortfall estimation. *Computational Statistics and Data Analysis*. 2016; 94:1–19.
- Kloke, J., McKean, J. Rfit: Rank estimation for linear models. R package version 1.0. 2016. URL: <https://CRAN.R-project.org/package=Rfit>
- Kneib T. Beyond mean regression. *Statistical Modelling*. 2013; 13:275–303.
- Koenker R, Bassett JG. Regression quantiles. *Econometrica*. 1978; 46:33–50.
- Kosorok MR, Song R. Inference under right censoring for transformation models with a change-point based on a covariate threshold. *The Annals of Statistics*. 2007; 35:957–989.
- Kuan CM, Yeh JH, Hsu YC. Assessing value at risk with care, the conditional autoregressive expectile models. *Journal of Econometrics*. 2009; 150:261–270.
- Lee S, Seo MH, Shin Y. Testing for threshold effects in regression models. *Journal of the American Statistical Association*. 2011; 106:220–231.

- Li C, Wei Y, Chappell R, He X. Bent line quantile regression with application to an allometric study of land mammals' speed and mass. *Biometrics*. 2011; 67:242–249. [PubMed: 20528859]
- Newey WK, McFadden D. Large sample estimation and hypothesis testing. *Handbook of Econometrics*. 1994; 4:2111–2245.
- Newey WK, Powell JL. Asymmetric least squares estimation and testing. *Econometrica*. 1987; 55:819–847.
- Pollard, D. *Convergence of Stochastic Processes*. Springer Science & Business Media; 1984.
- Qu Z. Testing for structural change in regression quantiles. *Journal of Econometrics*. 2008; 146:170–184.
- Quandt RE. The estimation of the parameters of a linear regression system obeying two separate regimes. *Journal of the American Statistical Association*. 1958; 53:873–880.
- Quandt RE. Tests of the hypothesis that a linear regression system obeys two separate regimes. *Journal of the American Statistical Association*. 1960; 55:324–330.
- Schnabel SK, Eilers PH. Optimal expectile smoothing. *Computational Statistics and Data Analysis*. 2009; 53:4168–4177.
- Schulze Waltrup L, Sobotka F, Kneib T, Kauermann G. Expectile and quantile regression—david and goliath? *Statistical Modelling*. 2015; 15:433–456.
- Silverman, BW. *Density Estimation for Statistics and Data Analysis*. Vol. 26. CRC press; 1986.
- Sobotka F, Kauermann G, Schulze Waltrup L, Kneib T. On confidence intervals for semiparametric expectile regression. *Statistics and Computing*. 2013; 23:135–148.
- Sobotka, F., Schnabel, S., Schulze Waltrup, L., Eilers, P., Kneib, T., Kauermann, G. *expectreg: expectile and quantile regression*. R package version 0.39. 2014. URL: <https://CRAN.R-project.org/package=expectreg>
- Stute W. Nonparametric model checks for regression. *The Annals of Statistics*. 1997; 25:613–641.
- Van der Vaart, AW. *Asymptotic Statistics*. Cambridge University Press; 2000.
- Xie S, Zhou Y, Wan A. A varying-coefficient expectile model for estimating value at risk. *Journal of Business and Economic Statistics*. 2014; 32:576–592.
- Yao Q, Tong H. Asymmetric least squares regression estimation: A nonparametric approach. *Journal of Nonparametric Statistics*. 1996; 6:273–292.
- Zhang, F., Li, Q. *cthreshER: Continuous threshold expectile regression*. R package version 1.0. 2016. URL: <https://CRAN.R-project.org/package=cthreshER>
- Zhang F, Li Q. Robust bent line regression. *Journal of Statistical Planning and Inference*. 2017; 185:41–55. [PubMed: 28943710]

Figure 1.

Mean and 95% averaged confidence interval (CI) of the estimates for the IID case in a simulation study with 1000 repetitions. “*” indicates the coverage probability is less than 90%.

Figure 2.

Mean and 95% averaged confidence interval (CI) of the estimates for the heteroscedastic case in a simulation study with 1000 repetitions. “***” indicates the coverage probability is less than 90%.

Figure 3.

Comparison of the proposed test ("Our") and the sup-likelihood-ratio-type test ("Lee"). Type I error and power are calculated at the significance level of 5% from 500 simulated samples of 200 observations, with 500 bootstrap replicates. The red line represents power = 0.05.

(a) Scatter plot of height (in cm) against age (in years).

(b) Fitted expectile curves for the data with square root transformed age.

(c) Fitted quantile curves for the data with square root transformed age.

Figure 4.

Analysis of Dutch boys data from the Fourth Dutch Growth Study.

(a) Fitted expectile curves for the data.

(b) Fitted quantile curves for the data.

Figure 5.
Analysis of baseball salaries data.

Table 1

Averaged prediction errors for our method and Schnabel’s method, based on 1000 simulated samples of $n = 200$ observations. The cases with a lower prediction error in the comparison are in boldface.

Error	τ	IID		Heteroscedastic	
		Our	Schnabel	Our	Schnabel
1	0.05	34.82	36.97	54.77	56.90
	0.10	53.27	56.01	83.59	86.27
	0.25	85.57	89.69	134.12	138.17
	0.50	102.50	107.27	160.65	165.39
	0.75	85.41	89.27	133.81	137.66
2	0.90	53.07	55.45	83.28	85.61
	0.95	34.71	36.45	54.56	56.37
	0.05	95.48	100.27	148.48	155.06
	0.10	127.41	131.05	198.49	201.90
	0.25	179.89	184.46	280.44	283.70
3	0.50	206.74	211.84	322.03	326.23
	0.75	180.75	184.52	281.28	284.07
	0.90	129.34	132.14	200.36	203.08
	0.95	97.99	101.74	150.27	156.18
	0.05	42.21	44.65	68.63	71.13
	0.10	61.61	64.41	98.87	101.45
	0.25	95.33	99.41	150.89	154.70
	0.50	112.81	117.61	177.74	182.23
	0.75	94.88	98.61	149.40	152.90
	0.90	61.01	63.22	96.46	98.35
	0.95	41.61	43.38	66.04	67.87

Table 2

The estimated parameters and their standard errors (listed in parentheses) for Dutch boys data. The p-values are computed from our proposed threshold test based on 1000 bootstrap replicates.

τ	p-value	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	t
0.05	0	40.875 (0.188)	31.019 (0.080)	-30.268 (5.187)	4.337 (0.014)
0.10	0	41.356 (0.153)	31.480 (0.069)	-31.624 (3.983)	4.316 (0.013)
0.25	0	42.011 (0.121)	32.217 (0.0593)	-30.252 (3.099)	4.276 (0.0118)
0.50	0	42.751 (0.106)	32.956 (0.056)	-32.305 (2.764)	4.255 (0.012)
0.75	0	43.485 (0.100)	33.722 (0.057)	-31.886 (2.138)	4.214 (0.010)
0.90	0	44.235 (0.103)	34.456 (0.062)	-31.624 (2.013)	4.173 (0.011)
0.95	0	44.780 (0.110)	34.887 (0.067)	-34.831 (1.996)	4.173 (0.011)

Table 3

The estimated parameters and their standard errors (listed in parentheses) for base-ball salaries data. The p-values are computed from our threshold test based on 1000 bootstrap replicates.

z	p-value	β_0	β_1	β_2	γ_1	γ_2	t
0.05	0.000	4.291 (0.140)	0.962 (0.067)	-3.323 (0.398)	0.115 (0.016)	-0.282 (0.041)	2.235 (0.065)
0.10	0.000	4.305 (0.156)	0.989 (0.066)	-3.150 (0.568)	0.110 (0.013)	-0.262 (0.034)	2.255 (0.074)
0.25	0.000	4.452 (0.203)	1.021 (0.049)	-2.673 (0.619)	0.102 (0.013)	-0.258 (0.040)	2.256 (0.066)
0.50	0.000	4.772 (0.241)	1.007 (0.045)	-2.360 (0.405)	0.091 (0.012)	-0.281 (0.043)	2.276 (0.056)
0.75	0.001	5.115 (0.295)	0.975 (0.054)	-2.128 (0.323)	0.083 (0.013)	-0.309 (0.047)	2.296 (0.055)
0.90	0.001	5.570 (0.558)	0.917 (0.088)	-1.998 (0.424)	0.073 (0.018)	-0.350 (0.077)	2.316 (0.102)
0.95	0.007	6.047 (0.780)	0.872 (0.099)	-1.979 (0.343)	0.062 (0.025)	-0.405 (0.111)	2.317 (0.089)