American Review of Mathematics and Statistics
December 2016, Vol. 4, No.2, pp. 30-38
ISSN: 2374-2348 (Print), 2374-2356 (Online)

Published by American Research Institute for Policy Development
DOI: 10.15640/arms.v4n1a4
URL: https://doi.org/10.15640/arms.v4n1a4

Asymptotic Properties of Local Composite Quantile Regression Estimation for Time-Varying Diffusion Model[1]

Ji-xia Wang[2] & Yuewen Xiao[3*]

Abstract

Based on discretely observed samples, this paper proposes local linear composite quantile regression estimation for time-dependent drift parameter of diffusion models. We verify the asymptotic bias, asymptotic variance and asymptotic normality of the local estimation proposed. The asymptotic relative efficiency of the local estimation with respect to local least squares estimation is discussed. The results show that the estimation proposed can be more efficient than the local least squares estimation for many commonly seen error distributions

Keywords: time-dependent parameter; composite quantile regression estimation; local linear fitting; diffusion model; asymptotic normality.

Subclass: [2010]{62M05; 62F12}

1. Introduction

Composite quantile regression (CQR) is proposed by Zou and Yuan (2008) for estimating regression coefficients in classical linear regression models. More recently, Kai el.(2010) considers a general non-parametric regression models by using CQR method. However, to our knowledge, little literature has researched parameter estimation by CQR in diffusion models. This motivates us to consider estimating regression coefficients under the framework of diffusion models. In this paper, we consider the diffusion model on a filtered probability space $(\Omega, \mathrm{F}, (\mathrm{F}_t)_{t\geq 0}, P)$,

$$(1.1) \qquad dX_t = \beta(t)b(X_t)dt + \sigma(X_t)dW_t,$$

where $\beta(t)$ is a time-dependent drift function and W_t is the standard Brownian motion. $b(\cdot)$ and $\sigma(\cdot)$ are known functions. Model (1.1) includes many famous option pricing models and interest rate term structure models, such as Black and Scholes(1973), Vasicek(1977), Ho and Lee(1986), Black, Derman and Toy (1990) and so on.

---

[1]This work was supported by grants from the National Natural Science Foundation of China(No. U1504701) and Program for Innovative Research Team(in Science and Technology)in University of Henan Province(No. 14IRTSTHN023).
[2] College of Mathematics and Information Science, Henan Normal University, 453007, Henan Province, P. R. China.
[3]Corresponding author. Rm 711, Siyuan Bldg, 670 Guoshun Rd, Shanghai, P.R.China, 200433. TEL: 86-21-25011234. FAX: 86-21-65642351. Email: yuewenxiao@fudan.edu.cn

We allow $\beta(t)$ being smooth in time. The techniques that we employ here are based on local linear fitting (see Fan and Gijbels(1996)) for the time-dependent parameter. The rest of this paper is organized as follows. In Section 2, we propose the local linear composite quantile regression estimation for the drift parameter and study its asymptotic properties. The asymptotic relative efficiency of the local estimation with respect to local least squares estimation is discussed in Section 3. The proof of result is given in Section 4.

2. Local estimation of the time-dependent parameter

Let the data $\{X_{t_i}, i=1,2,\cdots,n+1\}$ be equally sampled at discrete time points, $t_1 < t_2 < \cdots < t_{n+1}$. Denote $Y_{t_i} = X_{t_{i+1}} - X_{t_i}, \varepsilon_{t_i} = W_{t_{i+1}} - W_{t_i}$, and $\Delta_i = t_{i+1} - t_i$. Due to the independent increment property of Brownian motion W_t, ε_{t_i} are independent and normally distributed with mean zero and variance Δ_i. Thus, the discretized version of the model (1.1) can be expressed as

$$(2.1) \qquad Y_{t_i} = \beta(t_i) b(X_{t_i}) \Delta_i + \sigma(X_{t_i}) \Delta_i Z_{t_i},$$

where Z_{t_i} are independent and normally distributed with mean zero and variance $1/\Delta_i$.The first-order discretized approximation error to the continuous-time model is extremely small according to the findings in Stanton (1997) and Fan and Zhang(2003), this simplifies the estimation procedure.

Suppose the drift parameter $\beta(t)$ to be twice continuously differentiable in t. We can take $\beta(t)$ to be local linear fitting. That is, for a given time point t_0, we use the approximation

$$(2.2) \qquad \beta(t) \approx \beta(t_0) + \beta'(t_0)(t - t_0)$$

for t in a small neighborhood of t_0. Let h denote the size of the neighborhood and $K(\cdot)$ be a nonnegative weighted function. h and $K(\cdot)$ are the bandwidth parameter and kernel function, respectively. Denoting $\beta_0 = \beta(t_0)$ and $\beta_1 = \beta'(t_0)$, (2.2) can be expressed as

$$(2.3) \qquad \beta(t) \approx \beta_0 + \beta_1(t - t_0).$$

Now we propose the local linear CQR estimation of the drift parameter $\beta(t)$. Let $\rho_{\tau_k}(r) = \tau_k r - \tau I_{\{r<0\}}, k = 1,2,\cdots,q$,which are q check loss functions at q quantile positions: $\tau_k = \frac{k}{q+1}$. Thus, following the local CQR technique, $\beta(t)$ can be estimated via minimizing the locally weighted CQR loss

$$(2.4) \qquad \sum_{k=1}^{q} \{\sum_{i=1}^{n} \rho_{\tau_k} \{\frac{Y_{t_i}}{\Delta_i}[b(X_{t_i})]^{-1} - \beta_{0k} - \beta_1(t_i - t_0)\} K_h(t_i - t_0)\},$$

where $K_h(t_i - t_0){=}K(\frac{t_i - t_0}{h})$ and h is a properly selected bandwidth. Denote the minimizer of the locally weighted CQR loss (2.4) by $(\hat{\beta}_{01}, \hat{\beta}_{02}, \cdots, \hat{\beta}_{0q}, \hat{\beta}_1)^T$. Then, we let

$$(2.5) \quad \hat{\beta}(t_0) = \frac{1}{q}\sum_{k=1}^{q} \hat{\beta}_{0k}$$

We refer to $\hat{\beta}(t_0)$ as the local linear CQR estimation of $\beta(t_0)$, for a given time point t_0. To obtain the estimated function $\hat{\beta}(\cdot)$, we usually evaluate the estimations at hundreds of grid points.

In order to discuss the asymptotic properties of the estimation, we introduce the following assumptions. Throughout this paper, M denotes a positive generic constant independent of all other variables.

(A1) The functions $b(\cdot)$ and $\sigma(\cdot)$ in model (1.1) are continuous.

(A2) The kernel function $K(\cdot)$ is a symmetric and Lipschitz continuous function with finite support $[-M, M]$.

(A3) The bandwidth $h{=}h(n) \to 0$ and $nh \to 0$.

Let $F(\cdot)$ and $f(\cdot)$ be the cumulative density function and probability density function of the error, respectively. $g(\cdot)$ denotes the density function of time, usually a uniform distribution on time interval $[a,b]$. Define

$$\mu_j = \int u^j K(u)du, \quad \nu_j = \int u^j K^2(u)du, \quad j = 1, 2, \cdots$$

and

$$(2.6) \quad R(q) = \frac{1}{q^2}\sum_{k=1}^{q}\sum_{k'=1}^{q} \frac{\tau_{kk'}}{f(c_k) f(c_{k'})}$$

where $c_k = F^{-1}(\tau_k)$ and $\tau_{kk'}{=}\tau_k \wedge \tau_{k'} - \tau_k \tau_{k'}$.

Theorem 2.1 Under assumptions (A1)-(A3), for a given time point t_0, the local CQR estimation $\hat{\beta}(t_0)$ from (2.5) satisfies,

$$(2.7) \quad E[\hat{\beta}(t_0)] - \beta(t_0) = \frac{1}{2}\beta''(t_0)\mu_2 h^2 + o(h^2)$$

$$(2.8) \quad Var[\hat{\beta}(t_0)] = \frac{1}{nh}\frac{\nu_0 \sigma^2(X_{t_0})}{g(t_0) b^2(X_{t_0})} R(q) + o(\frac{1}{nh})$$

and, as $n \to \infty$,

$$(2.9)\quad \sqrt{nh}\{\hat{\beta}(t_0)-\beta(t_0)-\frac{1}{2}\beta''(t_0)\mu h^2\}\rightarrow_L N(0,\frac{v_0\sigma^2(X_{t_0})}{g(t_0)b^2(X_{t_0})}R(q))$$

where $\rightarrow_L$ means convergence in distribution.

3. Asymptotic relative efficiency

We discuss the asymptotic relative efficiency(ARE) of the local linear CQR estimation with respect to the local linear least squares estimation(see Fan and Gijbels(1996)) by comparing their mean-squared errors(MSE).From theorem 2.1, we obtain the MSE $\hat{\beta}(t_0)$. That is,

$$(3.1)\quad MSE[\hat{\beta}(t_0)\,]=[\frac{1}{2}\beta''(t_0)\mu_2]^2+\frac{1}{nh}\frac{v_0\sigma^2(X_{t_0})}{g(t_0)b^2(X_{t_0})}R(q)+o(h^4+\frac{1}{nh})$$

We obtain the optimal bandwidth via minimizing the MSE (3.1), denoted by

$$h^{opt}(t_0)\,]=[\frac{v_0\sigma^2(X_{t_0})R(q)}{g(t_0)b^2(X_{t_0})[\beta''(t_0)\mu_2]^2}]^{\frac{1}{5}}n^{-\frac{1}{5}}.$$

The MSE of the local linear least squares estimation of $\beta(t_0)$, denoted by $\hat{\beta}_{LS}(t_0)$, is

$$(3.2)\quad MSE[\hat{\beta}_{LS}(t_0)\,]=[\frac{1}{2}\beta''(t_0)\mu_2]^2h^4+\frac{1}{nh}\frac{v_0\sigma^2(X_{t_0})}{g(t_0)b^2(X_{t_0})}+o(h^4+\frac{1}{nh})$$

and the optimal bandwidth is

$$h^{opt}_{LS}(t_0)\,]=[\frac{v_0\sigma^2(X_{t_0})}{g(t_0)b^2(X_{t_0})[\beta''(t_0)\mu_2]^2}]^{\frac{1}{5}}n^{-\frac{1}{5}}.$$

By straightforward calculations, we have, as $n\rightarrow\infty$,

$$\frac{MSE[\hat{\beta}_{LS}(t_0)\,]}{MSE[\hat{\beta}(t_0)\,]}\rightarrow[R'(q)]^{-\frac{4}{5}}$$

Thus, the ARE of the local linear CQR estimation with respect to the local linear least squares estimation is

$$(3.3)\quad ARE(\hat{\beta}(t_0),\hat{\beta}_{LS}(t_0))=[R(q)]^{-\frac{4}{5}}$$

(3.3) reveals that the ARE depends only on the error distribution. The ARE we obtained is equal to that in Kai el.(2010).

Table 3.1 displays $ARE(\hat{\beta}(t_0), \hat{\beta}_{LS}(t_0))$ for some commonly seen error distributions. Table 1 in Kai el.(2010) can be seen as ARE for more error distributions.

Table 3.1: Comparisons of $ARE(\hat{\beta}(t_0), \hat{\beta}_{LS}(t_0))$ for the values of q

Error	$q=1$ $q=5$	$q=9$ $q=19$	$q=99$
$N(0,1)$	0.6968 0.9339	0.96590.9858	0.9980
Laplace	1.7411 1.2199	1.1548 1.0960	1.0296
$0.9N(0,1)+0.1N(0,10^2)$	4.0505 4.9128	4.70693.5444	1.1379

From Table 3.1, we can see that the local linear CQR estimation is more efficient than the local linear least squares estimation when the error distribution is not standard normal distribution. When the error distribution is $N(0,1)$ and $q=1,5,9,19,99$, the $ARE(\hat{\beta}(t_0), \hat{\beta}_{LS}(t_0))$ is very close to 1, which demonstrates that the local linear CQR estimation performs well when the error conforms to the standard normal distribution too.

4. Proof of result

In order to prove theorem 2.1, we first give some notations and lemmas. Let $S=\begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}$, and $\Sigma=\begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$, where S_{11} is a $q\times q$ diagonal matrix with diagonal elements $f(c_k), k=1,2,\cdots,q$, $S_{12}=(\mu_1 f(c_1), \mu_1 f(c_2), \cdots, \mu_1 f(c_q))^T$, $S_{21}=S_{12}{}^T$ and $S_{22}=\mu_2\sum_{k=1}^{q} f(c_k)$. Σ_{11} is a $q\times q$ matrix with (k,k')-element $\nu_0\tau_{k,k'}, k,k'=1,2,\cdots,q,$ $\Sigma_{12}=(\nu_1\Sigma_{k'=1}^{q}\tau_{1k'}, \nu_1\Sigma_{k'=1}^{q}\tau_{2k'}, \cdots, \nu_1\Sigma_{k'=1}^{q}\tau_{qk'})^T$, $\Sigma_{21}=\Sigma_{12}{}'$ and $\Sigma_{22}=\nu_2\Sigma_{k,k'=1}^{q}\tau_{kk'}$.

Furthermore, let $u_k=\sqrt{nh}\left\{\beta_{0k}-\beta(t_0)-\dfrac{\sigma(X_{t_0})}{b(X_{t_0})}c_k\right\}, v=h\sqrt{nh}\left\{\beta_1-\beta'(t_0)\right\}$,

and $\Delta_{i,k}=\dfrac{1}{\sqrt{nh}}\left(u_k+\dfrac{t_i-t_0}{h}v\right)$.Write $d_{i,k}=c_k\left\{\dfrac{\sigma(X_{t_i})}{b(X_{t_i})}-\dfrac{\sigma(X_{t_0})}{b(X_{t_0})}\right\}+r_i$ with $r_i=\beta(t_i)-\beta(t_0)-\beta'(t_0)(t_i-t_0)$.

Define $\eta_{i,k}^*$ to be $I_{\{Z_{t_i}\le c_k-d_{i,k}b(X_{t_i})/\sigma(X_{t_i})-\tau_k\}}$. Let $W_n^*=(w_{11}^*, w_{12}^*, \cdots, w_{1q}^*, w_{1(q+1)}^*)^T$ with $w_{1k}^*=\dfrac{1}{\sqrt{nh}}\sum_{i=1}^{n}\eta_{i,k}^* K_h(t_i-t_0), k=1,2,\cdots,q$, and $w_{1(q+1)}^*=\dfrac{1}{\sqrt{nh}}\sum_{k=1}^{q}\sum_{i=1}^{n}\eta_{i,k}^* K_h(t_i-t_0)\dfrac{t_i-t_0}{h}$.

Lemma 4.1 Under assumption (A1)-(A3), minimizing (2.4) is equivalent to minimizing the following term:

$$L_n(\theta)=\sum_{k=1}^{q}u_k\left\{\sum_{i=1}^{n}\frac{\eta_{i,k}^* K_h(t_i-t_0)}{\sqrt{nh}}\right\}+v\sum_{k=1}^{q}\sum_{i=1}^{n}\frac{\eta_{i,k}^* K_h(t_i-t_0)(t_i-t_0)}{h\sqrt{nh}}+\sum_{k=1}^{q}B_{n,k}(\theta)$$

$$=\frac{1}{2}\theta^T S_n\theta + (W_n^*)^T\theta + o_{\mathrm{p}}(1)$$

with respect to $\theta=(u_1,u_2,\cdots,u_q,v)$, where

$$B_{n,k}(\theta)=\sum_{i=1}^{n}\left(K_h(t_i-t_0)\int_0^{\Delta_{i,1}}\left[I_{\left\{Z_{t_i}\le c_k-\frac{d_{i,1}b(X_{t_i})}{\sigma(X_{t_i})}+\frac{zb(X_{t_i})}{\sigma(X_{t_i})}\right\}}-I_{\left\{Z_{t_i}\le c_k-\frac{d_{i,1}b(X_{t_i})}{\sigma(X_{t_i})}\right\}}\right]dz\right),\quad S_n=\begin{pmatrix}S_{n,11} & S_{n,12}\\ S_{n,21} & S_{n,22}\end{pmatrix},$$

with $S_{n,11}=\left[\sum_{i=1}^{n}K_h(t_i-t_0)\frac{b(X_{t_i})}{nh\sigma(X_{t_i})}\right]S_{11}$, $S_{n,21}=S_{n,12}^T$,

$$S_{n,12}=\left[\sum_{i=1}^{n}K_h(t_i-t_0)\frac{t_i-t_0}{h}\frac{b(X_{t_i})}{nh\sigma(X_{t_i})}\right]\left(f(c_1),f(c_2),\cdots,f(c_q)\right)^T,$$

and $S_{n,22}=\sum_{k=1}^{q}f(c_k)\sum_{i=1}^{n}\left[K_h(t_i-t_0)\frac{(t_i-t_0)^2}{h^2}\frac{b(X_{t_i})}{nh\sigma(X_{t_i})}\right]$.

The proof of lemma 4.1 is similar to lemma 2 and lemma 3 in Kai el.(2010).

Proof of theorem 2.1

Using the results of Parzen(1962), we have

$$\frac{1}{nh}\sum_{i=1}^{n}K_h(t_i-t_0)\frac{(t_i-t_0)^j}{h^j}\to_P g(t_0)u_j$$

where $\to_P$ means convergence in probability. Thus,

$$S_n\to_P\frac{g(t_0)b(X_{t_0})}{\sigma(X_{t_0})}S=\frac{g(t_0)b(X_{t_0})}{\sigma(X_{t_0})}\begin{pmatrix}S_{11} & S_{12}\\ S_{21} & S_{22}\end{pmatrix}.$$

According to lemma 4.1, we have

$$L_\theta(\theta)=\frac{1}{2}\frac{g(t_0)b(X_{t_0})}{\sigma(X_{t_0})}\theta^T S\theta+(W_n^*)^T\theta+o_p(1).$$

Since the convex function $L_n(\theta)-(W_n^*)^T\theta$ converges in probability to the convex function $\frac{1}{2}\frac{g(t_0)b(X_{t_0})}{\sigma(X_{t_0})}\theta^T S\theta$, according to the convexity lemma in Pollard(1991), for any compact set, the quadratic approximation to $L_\theta(\theta)$ holds uniformly for θ. Thus, we have

$$\hat{\theta}_n = -\frac{g(t_0)b(X_{t_0})}{\sigma(X_{t_0})}S^{-1}W_n^* + o_p(1).$$

Define $\eta_{i,k} = I_{\{z_{t_i}\le c_k\}} - \tau_k$ and $W_n = (w_{11}, w_{12}, \cdots w_{1q}, w_{1(q+1)})^T$ with $w_{1k} = \frac{1}{\sqrt{nh}}\sum_{i=1}^{n}\eta_{i,k}K_h(t_i-t_0), k=1,2,\cdots,q$, and $w_{1(q+1)} = \frac{1}{\sqrt{nh}}\sum_{k=1}^{q}\sum_{i=1}^{n}\eta_{i,k}K_h(t_i-t_0)\frac{t_i-t_0}{h}$.

By using the central limit theorem and the Cramer-Wald theorem, we have

$$(4.1)\quad \frac{W_n - E(W_n)}{\sqrt{Var(W_n)}} \to_L N(0, \mathrm{I}_{(q+1)\times(q+1)}).$$

Notice that $Cov(\eta_{i,k},\eta_{i,k'}) = \tau_{kk'}$ and $Cov(\eta_{i,k},\eta_{j,k'}) = 0$ If $i \ne j$. We have

$$\frac{1}{nh}\sum_{i=1}^{n}K_h^2(t_i-t_0)\frac{(t_i-t_0)^j}{h^j} \to_P g(t_0)v_j.$$

Thus, $Var(W_n) \to g(t_0)\Sigma.$. Combining the result (4.1), we have $W_n \to_L N(0, g(t_0)\Sigma)$. Moreover, we have

$$Var(w_{1k}^* - w_{1k}) = \frac{1}{nh}\sum_{i=1}^{n}K_h^2(t_i-t_0)Var(\eta_{i,k}^* - \eta_{i,k})$$

$$\le \frac{1}{nh}\sum_{i=1}^{n}K_h^2(t_i-t_0)[F(c_k + \frac{|d_{i,k}|b(X_{t_i})}{\sigma(X_{t_i})}) - F(c_k)] = o_p(1)$$

And

$$Var(w_{1(q+1)}^* - w_{1(q+1)}) = \frac{1}{nh}\sum_{i=1}^{n}K_h^2(t_i-t_0)\frac{t_i-t_0}{h}Var(\sum_{k=1}^{q}\eta_{i,k}^* - \eta_{i,k})$$

$$\le \frac{q^2}{nh}\sum_{i=1}^{n}K_h^2(t_i-t_0)\frac{t_i-t_0}{h}\max_k[F(c_k + \frac{|d_{i,k}|b(X_{t_i})}{\sigma(X_{t_i})}) - F(c_k)] = o_p(1).$$

Therefore, $Var(w_n^* - w_n) = o_p(1)$. Using Slutsky's theorem yields $w_n^* \to_L N(0, g(t_0)\Sigma)$.

Thus,

$$\hat{\theta}_n+\frac{\sigma(X_{t_0})}{g(t_0)b(X_{t_0})}S^{-1}E(W_n^*)\rightarrow_L N(0,\frac{\sigma^2(X_{t_0})}{g(t_0)b^2(X_{t_0})}S^{-1}\Sigma S^{-1})$$

So the asymptotic bias of $\hat{\beta}(t_0)$ is:

$$bias(\hat{\beta}(t_O))=\frac{1}{q}\frac{\sigma(X_{t_0})}{b(X_{t_0})}\sum_{k=1}^{q}c_k-\frac{1}{q\sqrt{nh}}\frac{\sigma(X_{t_0})}{g(t_0)b(X_{t_0})}e_{q\times1}^T(S_{11})^{-1}E(W_{1n}^*)$$

$$=\frac{1}{q}\frac{\sigma(X_{t_0})}{b(X_{t_0})}\sum_{k=1}^{q}c_k-\frac{1}{q\sqrt{nh}}\frac{\sigma(X_{t_0})}{g(t_0)b(X_{t_0})}\sum_{i=1}^{n}K_i\sum_{k=1}^{q}\frac{1}{f(c_k)}\left[F(c_k-\frac{d_{i,k}b(X_{t_0})}{\sigma(X_{t_i})})-F(c_k)\right],$$ where

$K_i=K_h(t_i-t_0), e_{q\times1}=(1,1,\cdots,1)^T$ and $W_{1n}^* = (w_{11}^*, w_{12}^*, \ldots w_{1q}^*)^T$.

Note that Z_{ti} is symmetric, thus $\sum_{k=1}^{q}c_k = 0$, and

$$\frac{1}{q}\sum_{k=1}^{q}\frac{1}{f(c_k)}\left[F(c_k-\frac{d_{i,k}b(X_{t_i})}{\sigma(X_{t_i})})-F(c_k)\right]=-\frac{r_ib(X_{t_i})}{\sigma(X_{t_i})}(1+o_P(1)).$$

Therefore,

$$bias(\hat{\beta}(t_O))=\frac{1}{nh}\frac{\sigma(X_{t_0})}{g(t_0)b(X_{t_0})}\sum_{i=1}^{n}K_i\frac{r_ib(X_{t_i})}{\sigma(X_{t_i})}(1+o_P(1)).$$ Since

$$\frac{1}{nh}\sum_{i=1}^{n}K_i\frac{r_ib(X_{t_i})}{\sigma(X_{t_i})}=\frac{g(t_0)\beta''(t_0)b(X_{t_0})}{2\sigma(X_{t_0})}\mu_2h^2(1+o_P(1)).$$ We have

$$bias(\hat{\beta}(t_0))=\frac{1}{2}\beta''(t_0)\mu_2h^2+o_P(h^2).$$ and

$$Var[\hat{\beta}(t_0)]=\frac{1}{nh}\frac{\sigma^2(X_{t_0})}{g(t_0)b^2(X_{t_0})}\frac{1}{q^2}e_{q\times1}^T(S^{-1}\sum S^{-1})_{11}e_{q\times1}+o_p(\frac{1}{nh})$$

$$=\frac{1}{nh}\frac{v_0\sigma^2(X_{t_0})}{g(t_0)b^2(X_{t_0})}R(q)+o_p(\frac{1}{nh}).$$

This completes the proof.

References

Black, F., Derman, E. and Toy, W.(1990).*A one-factor model of interest rate and its application to treasury bond options.* Finan. Analysts' J., **46:**33-39.

Black,F. and Scholes, M. (1973). *The pricing of options and corporate liabilities.* J. Polit. Economy, **81:**637-654.

Fan, J. and Gijbels, I.(1996).*Local polynomial modelling and its applications.* Chapman and Hall, London.

Fan, J., Jiang, J., Zhang, C. and Zhou, Z.(2003).*Time-dependent diffusion models for term structure dynamics and the stock price volatility.* Statistica Sinica, **13:**965-992.

Kai,B.,Li,R. and Zou, H.(2010).*Local composite quantile regression smoothing: an efficient and safe alternative to local polynomial regression.*J.R. Statist. Soc. B,**72:**49-69.

Parzen, E.(1962).*On estimation of a probability density function and model.* Ann. Math. Statist., **33:**1065-1076.

Pollard, D.(1991).*A symptotics for least absolute deviation regression estimations}.Econometr.*Theory,**7:**186-199.

Stanton,R.(1997).*A nonparametric model of term structure dynamics and the market price of interest rate risk.* J. Finance, **52:**1973-2002.

Vasicek,O.A.(1977).*An equilibrium characterization of the term structure.* J. Finan. Econom., **5:**177-188.

Zou,H. and Yuan, M.(2008).*Composite quantile regression and the oracle model selection theory***.**Ann. Statist.,**36:**1108-1126.

T.A.Louis, *Finding observed information using the EMI gorithm*, J.Royal Stat.Soc, 1982, B44: 226-233.