Regensburger
DISKUSSIONSBEITRÄGE
zur Wirtschaftswissenschaft

CONSISTENCY OF NONLINEAR REGRESSION QUANTILES UNDER TYPE I CENSORING WEAK DEPENDENCE AND GENERAL COVARIATE DESIGN

Walter Oberhofer and Harry Haupt
University of Regensburg

Regensburger Diskussionsbeiträge zur Wirtschaftswissenschaft 406
University of Regensburg Discussion Papers in Economics 406

UNIVERSITÄT REGENSBURG
Wirtschaftswissenschaftliche Fakultät

CONSISTENCY OF NONLINEAR REGRESSION QUANTILES UNDER TYPE I CENSORING, WEAK DEPENDENCE AND GENERAL COVARIATE DESIGN

Walter Oberhofer and Harry Haupt

Address correspondence to:
Harry Haupt, Department of Economics and Econometrics, University of Regensburg,
Universitaetsstr. 31, 93053 Regensburg, Germany, e-mail: harald.haupt@wiwi.uni-r.de.

Abstract:
For both deterministic or stochastic regressors, as well as parametric nonlinear or linear regression functions, we prove the weak consistency of the coefficient estimators for the Type I censored quantile regression model under different censoring mechanisms with censoring points depending on the observation index (in a nonstochastic manner) and a weakly dependent error process. Our argumentation is based on an exposition of the connection between the residuals of the economically relevant model at the outset of the censored regression problem, and the residuals which are subject to the corresponding optimization process of censored quantile regression.

Kurzfassung:
In dieser Arbeit wird die schwache Konsistenz der Koeffizientenschätzer für das zensierte (Typ I) Quantilsregressionsmodell unter sehr allgemeinen Bedingungen – lineare und nichtlineare Regressionsfunktionen, deterministische und stochastische Regressoren, Zensierungsgrenzen die (in nichtstochastischer Weise) vom Beobachtungsindex abhängen sowie schwach abhängige Fehlerterme – bewiesen. Die Argumentation basiert dabei auf dem Zusammenhang zwischen den ökonomischen relevanten Residuen des Ausgangsmodells und den Residuen die Gegenstand der Zielfunktion des Optimierungskalküls der zensierten Quantilsregression sind.

JEL classification: C22, C24.

1 Introduction

From an econometric point of view, median, or, more general, quantile restrictions have been introduced to address problems in dealing with non-normality and robustness issues. From an economic point of view, the consideration of unconditional and/or conditional quantiles could also be of paramount interest in quite a lot of applications. In addition, many of these problems are subject to (fixed or random) censoring of one or more variables.

In two seminal papers on Type I linear censored regression models, Powell (1984, 1986) proved the root-n-consistency of coefficient estimators under a zero median restriction by introducing the CLAD (censored least absolute deviations) estimator, and subsequently generalized his results to conditional quantile restrictions (censored quantiles). These papers gave rise to a large amount of literature concerning several further aspects of the problem. For example, computational issues (see, e.g., Buchinsky and Hahn, 1998; Bilias et al., 2000; Khan and Powell, 2001), semi- and nonparametric modelling of the regression function (see, e.g., Newey and Powell, 1990; Chen and Khan, 2000; Lewbel and Linton, 2002), and extensions to the case of random censoring (see Honore and Powell, 2003, and the literature cited therein). Parts of this literature have been surveyed and discussed in Powell (1994), and Pagan and Ullah (1999).

The objective of the present paper is to generalize existing results to the case of dependent errors for both cases of deterministic or stochastic covariates, and for both linear and nonlinear parametric regression functions. The generalization to censored nonlinear regression quantiles seems quite natural, and has, to the best knowledge of the authors, not been addressed in the literature so far. Nonlinear quantile regression models have been discussed in Oberhofer (1982), Koenker and Park (1994), and Mukherjee (2000). Quantile regression under (weakly) dependent errors has been discussed in Oberhofer and Haupt (2005) for unconditional quantiles in a parametric context, and Cai (2002), De Gooijer and Zerom (2003), and Ioannides (2004), in a nonparametric context. We explicitly elaborate the connection between the residuals of the economically relevant model at the outset of the censored regression problem, and the residuals which are subject of the corresponding optimization process. Though these issues have not been discussed in the extant literature – maybe due to the fact that neglecting them does not flaw the asymptotic results – these considerations seem to be vital for a proper understanding of distributional restrictions caused by censoring.

The remainder of the paper is organized as follows: in the following Section 2 we introduce the censored quantile regression model with fixed censoring points depending on the observation index, a general nonlinear regression function, and weakly dependent errors. In the first part of Section 3 we provide a thorough discussion of the model assumptions when the regressors are fixed, followed by a proof of weak consistency in the second part. Then, in an analogous manner, we study the assumptions and consistency for the case of stochastic regressors in Section 4.

2 Censored quantile regression

Given a complete probability space $(\Omega, \wp, P)$ we consider the nonlinear regression model

$$y_t^* = g(x_t, b_0) + u_t, \qquad 1 \le t \le T, \tag{2.1}$$

where $b_0 \in K \subset \mathbb{R}^p$ is a vector of unknown parameters, $x_t \in D_x \subset \mathbb{R}^m$ denote the row vectors of covariates, the disturbances u_t are unobservable weakly dependent random variables, the response variables y_t^* are (Type I) censored, and g is a function: $D_x \times K \to \mathbb{R}$.

Denoting the observed response variable by y_t, the censoring leads to

$$y_t = y_t^* = g(x_t, b_0) + u_t, \quad \text{if} \quad y_t^* \in C_t, \quad \text{and} \quad y_t \in A_t \quad \text{if} \quad y_t^* \notin C_t. \tag{2.2}$$

Usually, A_t contains only one or two elements (i.e. the censoring points). We consider the following case of censoring from above and below:

(C) $C_t = \{z \big| c_{1,t} < z < c_{2,t}\}$, $A_t = \{c_{1,t}, c_{2,t}\}$ and $y_t = c_{1,t}$, if $g(x_t, b_0) + u_t \le c_{1,t}$, and $y_t = c_{2,t}$, if $g(x_t, b_0) + u_t \ge c_{2,t}$.

Obviously, different cases of one-sided censoring are nested in (C). We assume that the censoring points $c_{1,t}$ and $c_{2,t}$, respectively are fixed and known. Treating the case of LAD and quantile estimation of the linear regression model under fixed censoring, Powell (1984, 1986) considered the case $c_{1,t} = 0$, $c_{2,t} = \infty$. In the case of random censoring, varying censoring points typically are observed only when the observation is censored.

It proves useful to define the censoring function $\text{cens}_t(z) = z$ if $z \in C_t$, $\text{cens}_t(z) = c_{1,t}$ if $z \le c_{1,t}$, and $\text{cens}_t(z) = c_{2,t}$ if $z \ge c_{2,t}$. Thus, (2.2) can be rewritten as $y_t = \text{cens}_t[g(x_t, b_0) + u_t]$. According to Powell (1984), we will consider the following nonlinear regression model

$$y_t = \text{cens}_t[g(x_t, b_0)] + v_t, \tag{2.3}$$

where $\text{cens}_t[g(x_t, b)]$ is the regression function and

$$v_t = \text{cens}_t[g(x_t, b_0) + u_t] - \text{cens}_t[g(x_t, b_0)] \tag{2.4}$$

is the error term with distribution $G_t(z)$. The distribution function of u_t is denoted by $F_t(z)$, where we assume $F_t^{-1}(\vartheta) = 0$ for a fixed ϑ with $0 < \vartheta < 1$ and all t. Consequently b_0 in (2.1) corresponds to the parameter vector of the ϑ regression quantile.

Given observations on $y = (y_1, \ldots, y_T)'$ and $x = (x_1, \ldots, x_T)$, any vector b minimizing the loss function

$$S_T(b) = \sum_{t=1}^{T} \vartheta \left| y_t - \text{cens}_t[g(x_t, b)] \right|_+ + (1 - \vartheta) \left| y_t - \text{cens}_t[g(x_t, b)] \right|_- \tag{2.5}$$

leads to an estimator of b_0 in (2.3) and will be denoted by $\hat{b}_T$, where for $z \in \mathbb{R}$ we define

$$|z|_+ = \begin{cases} z & \text{if} \quad z \geq 0, \\ 0 & \text{if} \quad z < 0, \end{cases} \qquad \text{and} \qquad |z|_- = \begin{cases} 0 & \text{if} \quad z \geq 0, \\ -z & \text{if} \quad z < 0. \end{cases}$$

The purpose of this paper is to set forth conditions for the weak consistency of $\hat{b}_T$. These conditions are very easy to understand and typical for regression quantiles under censoring. First, if properly normalized, $S_T(b)$ obeys a law of large numbers (LLN) and, second $S_T(b) = \lim_{T\to\infty} E[S_T(b)]$ has a unique minimizer.

Following this rather well known model preliminaries, a few remarks which might not only be of some pedagogical interest but also concern the mathematical correctness of further argumentation, seem to be in order. Despite the fact that the assumptions stated so far follow common practice, from a conceptual point of view it is rather surprising that the argumentation is based on the distribution and density of u_t. The residual, which should be the subject of minimization, however, is $y_t - \mathrm{cens}_t[g(x_t, b)]$. Thus, it proves useful to display the deviations in loss function (2.5) in terms of the errors v_t in the regression model (2.3). By defining

$$h_t(b) \equiv \mathrm{cens}_t[g(x_t, b)] - \mathrm{cens}_t[g(x_t, b_0)], \tag{2.6}$$

from (2.3) follows $y_t - \mathrm{cens}_t[g(x_t, b)] = y_t - \mathrm{cens}_t[g(x_t, b_0)] - h_t(b) = v_t - h_t(b)$, and consequently the loss function (2.5) can be rendered to

$$S_T(b) = \sum_{t=1}^{T} \vartheta \left|v_t - h_t(b)\right|_+ + (1-\vartheta) \left|v_t - h_t(b)\right|_- . \tag{2.7}$$

3 Deterministic regressors

3.1 Discussion of assumptions

We employ the following assumptions:

(A1) b_0 is an inner point of a compact set $K \subset \mathbb{R}^p$ and D_x is a measurable subspace of $\mathbb{R}^m$.

(A2) The input vectors $x_t = (x_{t,1}, \ldots, x_{t,m})$ are non-random and given.

(A3) By taking the supremum over all F elements of the σ-algebra $\sigma(u_t)$ and all G elements of the σ-algebra $\sigma(u_{t+k})$, the coefficients $\alpha_0(k|\, u) = \sup |P(F \cap G) - P(F)P(G)|$, $k = 0, 1, 2, \ldots$, converge to zero.

(A4) The distribution function of u_t is denoted by $F_t(z)$, where $F_t^{-1}(\vartheta) = 0$ for a fixed ϑ with $0 < \vartheta < 1$ and all t.

(A5) There exist $\alpha > 0$ and $\epsilon > 0$ such that for all z, where $|z| \leq \alpha$, and all t

$$|F_t(z) - F_t(0)| \geq \epsilon |z|.$$

(A6) For every $\delta > 0$ there exist $\beta > 0$ and T_0 such that for all $T \geq T_0$

$$\inf_{||b-b_0|| \geq \delta} \frac{1}{T} \sum_{t=1}^{T} |h_t(b)| \geq \beta.$$

(A7) There exists a constant $M < \infty$ such that for all $T \geq 1$

$$\frac{1}{T} \sum_{t=1}^{T} (c_{2,t} - c_{1,t})^2 \leq M.$$

Assumptions (A1) and (A2) imply the existence of a measurable estimator $\hat{b}_T$ due to theorem 3.10 of Pfanzagl (1969). Note that the notion of weak dependence introduced in assumption (A3) is significantly weaker compared to mixing concepts (see the discussion in Oberhofer and Haupt, 2005). The normalization in assumption (A4) is required to define the ϑ-quantile regression function, and (A5) is typical for quantile regression. Usually instead of (A5) the existence of the density $f_t(z)$ and its positivity in the near of $z = 0$ will be assumed. Assumption (A6) deserves some special attention. It is a natural identification criterion and guarantees under L_1 norm that an arbitrary regression function $\text{cens}_t[g(x_t, b)]$ and the true regression function $\text{cens}_t[g(x_t, b_0)]$ differ for $\|b - b_0\| \geq \delta > 0$. Assumption (A7) implies $\sup_{b \in K} T^{-1} \sum_{t=1}^{T} h_t(b)^2 \leq M < \infty$.

3.2 Consistency

For technical reasons we will consider the loss function

$$Q_T(b) = \frac{1}{T}[S_T(b) - S_T(b_0)] \tag{3.1}$$

equivalent to $S_T(b)$ from an estimation point of view. By using $Q_T(b)$ it is not necessary to assume the existence of moments of the errors u_t. For notational convenience $Q_T(b)$ will be rewritten as $Q_T(b) = T^{-1} \sum_{t=1}^{T} a_t(b)$, where, in the light of the preceding discussion

$$a_t(b) = \vartheta |v_t - h_t(b)|_+ + (1 - \vartheta)|v_t - h_t(b)|_- - \vartheta |v_t|_+ - (1 - \vartheta)|v_t|_-. \tag{3.2}$$

In order to prove the validity of a LLN for $Q_T(b)$, we calculate the expected value of $a_t(b)$ for the censoring mechanism (C). Again, for ease of notation, we abbreviate $a_t(b)$ by a_t, and $h_t(b)$ by h_t, respectively, throughout the remainder of the paper.

LEMMA 1. *For $h_t > 0$ follows*

$$E(a_t) = \int_0^{h_t} (h_t - z) dG_t(z) + h_t[G_t(0) - \vartheta], \tag{3.3}$$

and, for $h_t \leq 0$ follows

$$E(a_t) = \int_{h_t}^{0} (z - h_t) dG_t(z) + h_t[G_t(0) - \vartheta]. \tag{3.4}$$

PROOF. By definition

$$a_t = (1-\vartheta)h_t + \begin{cases} 0 & \text{if } \ v_t \leq \min(0, h_t), \\ -v_t & \text{if } \ 0 < v_t \leq h_t, \\ v_t - h_t & \text{if } \ h_t < v_t \leq 0, \\ -h_t & \text{if } \ v_t > \max(0, h_t). \end{cases} \tag{3.5}$$

Since we get $E(a_t) = 0$ for $h_t = 0$, we consider the case $h_t \neq 0$. Firstly, for $h_t > 0$ we get

$$E(a_t) = (1-\vartheta)h_t - \int_0^{h_t} z dG_t(z) - h_t P(v_t > h_t). \tag{3.6}$$

A possible discontinuity of $G_t(z)$ at $z = 0$ does not cause any problems here, and due to $P(v_t > h_t) = 1 - G_t(h_t)$, (3.3) follows from (3.6). Secondly, for $h_t < 0$ we get

$$E(a_t) = (1-\vartheta)h_t + \int_{h_t}^{0} (z - h_t) dG_t(z) - h_t P(v_t > 0). \tag{3.7}$$

Analogously a possible discontinuity of $G_t(z)$ at $z = h_t$ does not cause any problems here, and by noting that $P(v_t > 0) = 1 - G_t(0)$, (3.4) follows from (3.7), which completes the proof. ∎

LEMMA 2. *Under censoring mechanism* (C)*, we get*

$$G_t(z) = \left\{ \begin{array}{cl} 0 & \textit{if} \ \ z < 0, \\ F_t[z + c_{1,t} - g(x_t, b_0)] & \textit{if} \ \ 0 \leq z < c_{2,t} - c_{1,t}, \\ 1 & \textit{if} \ \ c_{2,t} - c_{1,t} \leq z, \end{array} \right\} \quad \textit{for} \quad g(x_t, b_0) \leq c_{1,t}, \tag{3.8}$$

$$G_t(z) = \left\{ \begin{array}{cl} 0 & \textit{if} \ \ z < c_{1,t} - g(x_t, b_0), \\ F_t(z) & \textit{if} \ \ c_{1,t} - g(x_t, b_0) \leq z < c_{2,t} - g(x_t, b_0), \\ 1 & \textit{if} \ \ c_{2,t} - g(x_t, b_0) \leq z, \end{array} \right\} \quad \textit{for} \quad c_{1,t} < g(x_t, b_0) < c_{2,t} \tag{3.9}$$

and

$$G_t(z) = \left\{ \begin{array}{cl} 0 & \textit{if} \ \ z < c_{1,t} - c_{2,t}, \\ F_t[z + c_{2,t} - g(x_t, b_0)] & \textit{if} \ \ c_{1,t} - c_{2,t} \leq z < 0, \\ 1 & \textit{if} \ \ 0 \leq z, \end{array} \right\} \quad \textit{for} \quad c_{2,t} \leq g(x_t, b_0). \tag{3.10}$$

PROOF. First we consider the case $g(x_t, b_0) \leq c_{1,t}$, where $v_t = \text{cens}_t[g(x_t, b_0)] + u_t - c_{1,t}$. Obviously we have $0 \leq v_t \leq c_{2,t} - c_{1,t}$ leading to $G_t(z) = 0$ if $z < 0$, $G_t(0) = F_t[c_{1,t} - g(x_t, b_0)]$, and $G_t(z) = 1$ if $c_{2,t} - c_{1,t} \leq z$. For $0 < v < c_{2,t} - c_{1,t}$ follows $v \leq z$ if $g(x_t, b_0) + u_t - c_{1,t} \leq z$, implying $G_t(z) = F_t[z + c_{1,t} - g(x_t, b_0)]$. Thus (3.8) is shown, and (3.9) and (3.10) follow from analogous argumentation. ■

LEMMA 3. *Under censoring mechanism* (C), *we get*

$$E(a_t) = h_t\{F_t[\text{cens}_t[g(x_t, b_0)] - g(x_t, b_0)] - \theta\} + \begin{cases} \int_0^{h_t}(h_t - z)dF_t(z), & if \quad h_t > 0, \\ \int_{h_t}^0 (z - h_t)dF_t(z), & if \quad h_t \leq 0, \end{cases} \tag{3.11}$$

PROOF. (i) The assertion is trivial for $h_t = 0$.

(ii) For $c_{1,t} < g(x_t, b_0) < c_{2,t}$, due to Lemma 2 the discontinuity points are given by $z_0 = c_{1,t} - g(x_t, b_0) < 0$ and $z_1 = c_{2,t} - g(x_t, b_0) > 0$. By definition $c_{1,t} - g(x_t, b_0) \leq h_t \leq c_{2,t} - g(x_t, b_0)$. A possible discontinuity of $G_t(z)$ in h_t does not affect the integrals in (3.3) and (3.4), respectively. In addition z_0 lies outside the range of integration in (3.3) and z_1 lies outside the range of integration in (3.4). Consequently, due to Lemma 2, $dG_t(z)$ in (3.3) and (3.4) can be replaced by $dF_t(z)$, respectively. Then the assertion follows from $G_t(0) = F_t(0) = \vartheta$ and $\text{cens}_t[g(x_t, b_0)] = g(x_t, b_0)$.

(iii) For $g(x_t, b_0) \leq c_{1,t}$, by definition we have $0 \leq h_t \leq c_{2,t} - c_{1,t}$. The discontinuity points of $G_t(z)$ are given by $z_0 = 0$ and $z_1 = c_{2,t} - c_{1,t}$, not affecting the integral in (3.3). Therefore the assertion follows from $G_t(0) - \vartheta = F[c_{1,t} - g(x_t, b_0)] - \vartheta \geq 0$ according to Lemma 2 and $\text{cens}_t[g(x_t, b_0)] = c_{1,t}$.

(iv) For $g(x_t, b_0) \geq c_{2,t}$, by definition we have $c_{1,t} - c_{2,t} \leq h_t \leq 0$. The discontinuity points of $G_t(z)$ are given by $z_0 = 0$ and $z_1 = c_{1,t} - c_{2,t}$. Obviously, both points do not lie in the interior of the range of integration in (3.4), but z_0 lies on the upper boundary of the integral in (3.4) leading to a contribution $-h_t\{1 - F_t[c_{2,t} - g(x_t, b_0)]\}$. The second summand on the right-hand-side of (3.4) equals $h_t(1 - \theta)$. Consequently, the assertion is proved, since $\text{cens}_t[g(x_t, b_0)] = c_{2,t}$. ■

REMARK. Considering the three cases $g(x_t, b_0) \leq c_{1,t}$, $c_{1,t} < g(x_t, b_0) < c_{2,t}$, and $c_{2,t} \leq g(x_t, b_0)$ follows that the first summand on the right hand side of (3.11) is non-negative.

LEMMA 4. *Under the censoring mechanism* (C)*, and assumptions* (A2)*,* (A4)*-*(A6)*, the following holds true: For every $\delta > 0$ there exists a T_0, such that for all $T \geq T_0$*

$$\inf_{||b-b_0|| \geq \delta} E[Q_T(b)] = \inf_{||b-b_0|| \geq \delta} \frac{1}{T} \sum_{t=1}^{T} E(a_t) \geq \eta,$$

where $\eta = (\epsilon\beta/4)\min(\alpha, \beta/4)$.

PROOF. For $h_t \geq 0$, we get from Lemma 3

$$E(a_t) \geq \int_0^{h_t} [h_t - z] dF_t(z) \geq \int_0^{h_t/2} [h_t - z] dF_t(z) \geq \frac{h_t}{2} [F_t(h_t/2) - F_t(0)].$$

Then, due to (A5)

$$E(a_t) \geq \begin{cases} \frac{h_t}{2} \epsilon \frac{h_t}{2} & \text{if} \quad \frac{h_t}{2} \leq \alpha, \\ \frac{h_t}{2} \epsilon \alpha & \text{if} \quad \frac{h_t}{2} > \alpha. \end{cases} \tag{3.12}$$

Inequality (3.12) holds analogously for $h_t < 0$ and summation over t leads to

$$\frac{1}{T} \sum_{t=1}^{T} E(a_t) \geq \frac{\epsilon \alpha}{2} \frac{1}{T} \sum_{\substack{t=1 \\ |h_t| > 2\alpha}}^{T} |h_t| + \frac{\epsilon}{4} \frac{1}{T} \sum_{\substack{t=1 \\ |h_t| \leq 2\alpha}}^{T} h_t^2. \tag{3.13}$$

Due to (A6), for $||b - b_0|| \geq \delta$ we obtain

$$\frac{1}{T} \sum_{\substack{t=1 \\ |h_t| > 2\alpha}}^{T} |h_t| \geq \frac{\beta}{2}, \qquad \text{or} \qquad \frac{1}{T} \sum_{\substack{t=1 \\ |h_t| \leq 2\alpha}}^{T} |h_t| \geq \frac{\beta}{2}.$$

In the first case the assertion follows from (3.13) by setting $\eta = (\epsilon \alpha \beta)/4$. Applying the Cauchy-Schwarz inequality to the second case leads to

$$\frac{1}{T} \sum_{\substack{t=1 \\ |h_t| \leq 2\alpha}}^{T} h_t^2 \geq \left(\frac{1}{T} \sum_{\substack{t=1 \\ |h_t| \leq 2\alpha}}^{T} |h_t| \right)^2 \geq \left(\frac{\beta}{2} \right)^2. \tag{3.14}$$

Due to (3.14) the assertion follows from (3.13) by setting $\eta = (\epsilon \beta^2)/16$. ■

In the following Theorem it will be shown that the loss function $Q_T(b)$ obeys a LLN. Often a LLN for the loss function is assumed and it will not be deduced from properties of the error process.

THEOREM 1. *For the nonlinear regression model (2.3) under assumptions (A1)-(A7), the estimator $\hat{b}_T$ of the parameter b_0 of the ϑth regression quantile is weakly consistent.*

PROOF. Due to (3.5) we get $|a_t| \leq |h_t|$. Therefore, according to Doukhan (1994, 1.2.2)

$$|cov(a_s, a_t)| \leq 4|h_s||h_t| \, \alpha(s, t| \, a), \tag{3.15}$$

where $\alpha(s, t| \, a)$ is the mixing coefficient of the random variables a_s and a_t.

According to (3.5) and the definition of v_t, a_t is a function of u_t implying $\alpha(s,t|\,a) \leq \alpha_0(|s-t|\,|\,u)$, where $\alpha_0(|s-t|\,|\,u)$ is defined in assumption (A3). Consequently from (3.15) follows

$$var[Q_T(b)] = var\left(\frac{1}{T}\sum_{t=1}^{T} a_t\right) \leq 8\frac{1}{T}\sum_{k=0}^{T-1}\alpha_0(k|\,u)\frac{1}{T}\sum_{t=1}^{T-k}|h_t||h_{t+k}|. \tag{3.16}$$

Due to (A7), the second factor of the sum on the right hand side of (3.16) can be bounded by

$$\frac{1}{T}\sum_{t=1}^{T-k}|h_t||h_{t+k}| \leq \frac{1}{T}\sqrt{\sum_{t=1}^{T-k}h_t^2\sum_{s=1}^{T-k}h_{s+k}^2} \leq \frac{1}{T}\sum_{t=1}^{T}h_t^2 \leq M. \tag{3.17}$$

For $\eta = (\epsilon\beta)/4 \cdot \min(\alpha, \beta/4)$ used in Lemma 4, from the Tchebichev inequality follows for all b with $||b-b_0|| \geq \delta$

$$P\left(|Q_T(b) - E[Q_T(b)]| \leq \frac{\eta}{2}\right) \geq 1 - \frac{4}{\eta^2}var[Q_T(b)],$$

and consequently, by virtue of Lemma 4, (3.16), and (3.17)

$$P\left(|Q_T(b)| \geq \frac{\eta}{2}\right) \geq 1 - 32\left(\frac{1}{T}\sum_{k=0}^{T-1}\alpha_0(k|\,u)\right)M\frac{1}{\eta^2}. \tag{3.18}$$

Since the $\alpha_0(k|\,u)$ constitute a null sequence, follows $T^{-1}\sum_k \alpha_0(k|\,u) \to 0$, and the right hand side of (3.18) converges to 1 for every fixed $\eta > 0$, and $\delta > 0$ for $T \to \infty$. The interpretation of (3.18) is then that, due to $Q_T(b_0) = 0$, the minimum of $Q_T(b)$ cannot be attained for a b with $||b-b_0|| \geq \delta$ asymptotically, whereas δ can be chosen arbitrarily small. ■

REMARK. The special case of a linear regression function, resulting from setting $g(x_t, b_0) = x_t b_0$, and $p = m$ in section 2, leads to the usual Type I linear censored regression model

$$y_t^* = x_t b_0 + u_t, \qquad 1 \leq t \leq T. \tag{3.19}$$

Assumption (A6) is a natural identification criterion, though unusual for the linear case. Usually the positive definiteness of $\lim_{T\to\infty} T^{-1}\sum_{t=1}^{T} x_t' x_t$ will be assumed, implying identifiability. At this point it proves useful to note that it makes few sense to consider trending regressors when using a censoring mechanism independent from t. For illustration, consider for example the simple linear regression $y_t^* = b_1 + b_2 x_t + u_t$, where x_t is assumed to be trending. Then, asymptotically, the share of censored realizations is either 0 or 1, and an asymptotic analysis in both cases is more or less senseless. If, however, we admit a censoring mechanism depending on t, in order to avoid the problem mentioned above, the censoring points have to be adapted to the evolution of the unknown trend. In what follows we will rule out the case of trending regressors and consider the case, where $\|x_t\|^2 \leq M_1 < \infty$. The considerations are similar to those of Powell (1984).

For further notational convenience we define the dummy variables

$$d_{0,t}(\rho) = \begin{cases} 1 & \text{if} \quad c_{1,t} + \rho \le x_t b_0 \le c_{2,t} - \rho, \\ 0 & \text{else,} \end{cases}$$

and

$$d_t(\rho) = \begin{cases} 1 & \text{if} \quad |x_t(b - b_0)| \ge \rho, \\ 0 & \text{else,} \end{cases}$$

and we assume

(A6.1) There exists an $M_1 < \infty$, such that $\|x_t\|^2 \le M_1$ for all t.

(A6.2) There exists a $\lambda > 0$, such that $T^{-1} \sum_{t=1}^{T} d_{0,t}(\lambda)(x_t' x_t)$ converges to a nonsingular matrix $\Omega(\lambda)$.

Then for $\rho = 2\alpha$ and all t, where $d_{0,t}(\rho) = 1$ and $d_t(\rho) = 1$, we have

$$|h_t| = \begin{cases} |x_t(b - b_0)| \ge \rho & \text{if} \quad x_t b \in C_t, \\ |\text{cens}_t[x_t b] - x_t b_0| \ge \rho & \text{if} \quad x_t b \notin C_t. \end{cases}$$

Thus, according to $E(a_t) \ge 0$, (A5), and $[x_t(b - b_0)]^2 \le \|x_t\|^2 \|b - b_0\|^2$ analogously to (3.13) follows

$$\begin{aligned} \frac{1}{T} \sum_{t=1}^{T} E(a_t) &\ge \frac{\epsilon\alpha}{2} \frac{1}{T} \sum_{t=1}^{T} d_{0,t}(\rho) d_t(\rho) |h_t| \\ &\ge \frac{\epsilon\alpha\rho}{2M_1} \frac{1}{T} \sum_{t=1}^{T} d_{0,t}(\rho) d_t(\rho) \|x_t\|^2 \\ &\ge \frac{\epsilon\alpha\rho}{2M_1 \|b - b_0\|^2} \frac{1}{T} \sum_{t=1}^{T} d_{0,t}(\rho) d_t(\rho) [x_t(b - b_0)]^2 \\ &\ge \frac{\epsilon\alpha\rho}{2M_1 \|b - b_0\|^2} [\frac{1}{T} \sum_{t=1}^{T} d_{0,t}(\rho)[x_t(b - b_0)]^2 - \frac{1}{T} \sum_{t=1}^{T} d_{0,t}(\rho)[1 - d_t(\rho)]\rho^2]. \end{aligned} \tag{3.20}$$

Then by denoting the lowest eigenvalue of the limit matrix $\Omega(\lambda)$ as λ_1, we obtain asymptotically

$$\inf_{b \ne b_0} \frac{1}{\|b - b_0\|^2} \frac{1}{T} \sum_{t=1}^{T} d_{0,t}(\rho)[x_t(b - b_0)]^2 \ge \lambda_1 / 2$$

and for all b, where $\|b - b_0\| \ge \delta$,

$$\frac{1}{\|b - b_0\|^2} \frac{1}{T} \sum_{t=1}^{T} d_{0,t}(\rho)(1 - d_t(\rho))\rho^2 \le \delta^{-2} \rho^2.$$

Consequently, because $\rho = 2\alpha < \lambda$ can be chosen arbitrarily close to zero, the right hand side of (3.20) is bounded from below asymptotically by a positive number. Note that if assumption (A5) is fulfilled for an α it remains valid for any positive $\alpha_1 < \alpha$ and if $\Omega(\lambda)$ is nonsingular for a λ, it remains nonsingular for any positive $\lambda_0 < \lambda$. Thus Lemma 4 is shown for T large enough assuming (A6.1) and (A6.2) instead of (A6). Obviously (A6.2) corresponds to the well known identification criterion used for the linear model and least squares estimation. At this point it is important to keep in mind that b_0 is unknown, with the consequence that we also do not know (and do not need to know) the summation area in (A6.2) for estimation purposes.

4 Stochastic regressors

4.1 Discussion of assumptions

Now assumption (A2), that the input vectors x_t are non-random and given, will be dropped and we will consider the case of stochastic regressors, which we assume to be independent of the errors. As a consequence, in place of assumptions (A6) and (A7) we introduce assumptions required for the weak consistency in this case.

In what follows, we will rule out the case of trending regressors, and consider the sequence $\{x_t|\, t = 1, 2, \ldots\}$ as a realization of a stochastic process $\{X_t|\, t = 1, 2, \ldots\}$. Please note that the small greek letters used in this section have another meaning than in the former sections. We start with an enumeration of the assumptions we require to prove consistency of $\hat{b}_T$ in the case of stochastic regressors.

(B1) $\{X_t\}$ is a sequence of $1 \times m$ random vectors with existing first and second moments.

(B2) All X_s are independent of all u_t.

(B3) For every $\delta > 0$ there exists a $\beta > 0$ such that for all $T > T_0$

$$\inf_{||b-b_0||\geq\delta} \frac{1}{T}\sum_{t=1}^{T} E[|\, h_t|\,] \geq \beta.$$

(B4) By taking the supremum over all F elements of the σ-algebra $\sigma(X_t)$ and all G elements of the σ-algebra $\sigma(X_{t+k})$, the coefficients $\alpha(k|\, X) = \sup|P(F \cap G) - P(F)P(G)|$, $k = 1, 2, \ldots$, converge to zero.

Due to (B2) assumptions (A3)-(A5) are valid independently of the regressors.

4.2 Consistency

The previous Lemmata 1-3 allow in Lemma 4 the calculation of $E[Q_T(b|\, x_t, \ldots, x_T)]$. In Lemma 5 we will apply this idea analogously to the calculation of the unconditional expectation $E[Q_T(b)]$ for the stochastic case.

LEMMA 5. *Under assumptions (*A1*), (*A3*)-(*A5*), (*A7*), and (*B1*)-(*B3*), for every $\delta > 0$ there exists an $\eta > 0$ such that for all $T \geq T_0$*

$$\inf_{||b-b_0||\geq\delta} E[Q_T(b)] = \inf_{||b-b_0||\geq\delta} \frac{1}{T}\sum_{t=1}^{T} E(a_t) \geq \eta.$$

PROOF. From (3.13) follows

$$\frac{1}{T}\sum_{t=1}^{T} E(a_t|\, x_1, \ldots, x_T) \geq \frac{\epsilon\alpha}{2}\frac{1}{T}\sum_{\substack{t=1\\|h_t|>2\alpha}}^{T} |h_t| + \frac{\epsilon}{4}\frac{1}{T}\sum_{\substack{t=1\\|h_t|\leq 2\alpha}}^{T} h_t^2, \tag{4.1}$$

where we consider h_t given $x_1, \ldots, x_T$, and therefore

$$\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T} E(a_t) &= \frac{1}{T}\sum_{t=1}^{T} E[E(a_t|\, x_1, \ldots, x_T)] \\
&\geq \frac{\epsilon\alpha}{2}\frac{1}{T}\sum_{\substack{t=1\\|h_t|>2\alpha}}^{T} E(|h_t|) + \frac{\epsilon}{4}\frac{1}{T}\sum_{\substack{t=1\\|h_t|\leq 2\alpha}}^{T} E(h_t^2)
\end{aligned} \tag{4.2}$$

Due to (B3) for $||b - b_0|| \geq \delta$ we get

$$\frac{1}{T}\sum_{\substack{t=1\\|h_t|>2\alpha}}^{T} E(|h_t|) \geq \frac{\beta}{2} \qquad \text{or} \qquad \frac{1}{T}\sum_{\substack{t=1\\|h_t|\leq 2\alpha}}^{T} E(|h_t|) \geq \frac{\beta}{2}. \tag{4.3}$$

In the first case the assertion follows from (4.2) by setting $\eta = (\epsilon\alpha\beta)/4$. Applying the Cauchy-Schwarz inequality twice to $T^{-1}\sum_{t=1}^{T} E(|h_t|)$ for $|h_t| \leq 2\alpha$ leads to

$$\frac{1}{T}\sum_{\substack{t=1\\|h_t|\leq 2\alpha}}^{T} E(|h_t|) \leq \frac{1}{T}\sum_{\substack{t=1\\|h_t|\leq 2\alpha}}^{T} \sqrt{E(h_t^2)} \leq \sqrt{\frac{1}{T}\sum_{\substack{t=1\\|h_t|\leq 2\alpha}}^{T} E(h_t^2)}, \tag{4.4}$$

and proves the assertion in the second case by setting $\eta = (\epsilon\beta^2)/16$. ■

To be in a position to prove Theorem 2, we have to generalize the condition

$$\lim_{T\to\infty} var\left[\frac{1}{T}\sum_{t=1}^{T} a_t\right] = 0,$$

which is central to Theorem 1.

THEOREM 2. *Under assumptions* (A1), (A3)-(A5), (A7), *and* (B1)-(B4), *the estimator* $\hat{b}_T$ *of the parameter* b_0 *of the* ϑ*th regression quantile is weakly consistent.*

PROOF. We have to show that $var(T^{-1}\sum_{t=1}^{T} a_t)$ converges to zero. Due to (3.15)

$$|cov(a_s|\, x_s, a_t|\, x_t)| \leq 4|h_s||h_t|\alpha_0(|t-s||\, u), \tag{4.5}$$

where h_t given $x_1 \dots x_T$ will be considered. Then, due to the identity $cov(X,Y) = E(XY) - E(X)E(Y)$ and the fundamental property of conditional expectation, we get

$$cov(a_s, a_t) = E[cov(a_s|\, x_s, a_t|\, x_t)] + cov[E(a_s|\, x_s), E(a_t|\, x_t)]. \tag{4.6}$$

Note that the second term on the right hand side of (4.6) vanishes, if X_s and X_t are independent. From (4.5) follows (by applying the same argument as in (3.15))

$$E[cov(a_s|\, x_s, a_t|\, x_t)] \leq 4E[|h_s||h_t|\,]\alpha_0(|t-s||\, u). \tag{4.7}$$

From $|a_t| \leq |h_t| \leq c_{2,t} - c_{1,t}$ follows

$$cov[E(a_s|\, x_s), E(a_t|\, x_t)] \leq 4|c_{2,s} - c_{1,s}||c_{2,t} - c_{1,t}|\alpha(|t-s||\, X). \tag{4.8}$$

and analogously to Theorem 1, by summation over s and t, due to (4.7) and (4.8) we get

$$var\left(\frac{1}{T}\sum_{t=1}^{T} a_t\right) \leq \frac{8}{T}\sum_{k=0}^{T-1}\alpha_0(k|\, u)M + \frac{8}{T}\sum_{k=0}^{T-1}\alpha(k|\, X)M. \tag{4.9}$$

Due to assumptions (A3) and (B4) this proves the assertion. ■

REMARK. Analogously to the deterministic case, for the linear model have to assume

(B3.1) There exists an $M_2 < \infty$, such that $E[\|x_t\|^2] \leq M_2$ for all t,

(B3.2) There exists a $\lambda > 0$, such that $T^{-1}\sum_{t=1}^{T} E[d_{0,t}(\lambda(x_t'x_t)]$ converges to a nonsingular matrix $\Omega(\lambda)$,

instead of (B3).

Literatur

[1] Bilias, Y., S. Chen, & Z. Ying (2000) Simple resampling methods for censored regression quantiles. *Journal of Econometrics* 103, 73-110.

[2] Buchinsky, M., & J. Hahn (1998) An alternative estimator for the censored quantile regression model. *Econometrica* 66, 653-671.

[3] Cai, Z. (2002) Regression quantiles for time series. *Econometric Theory* 18, 169-192.

[4] Chen, S., & S. Khan (2000) Estimating censored regression models in the presence of nonparameric multiplicative heteroskedasticity. *Journal of Econometrics* 98, 283- 316.

[5] De Gooijer, J.G., & D. Zerom (2003) On Additive Conditional Quantiles With High- Dimensional Covariates. *Journal of the American Statistical Association* 98, 135-146.

[6] Doukhan, P. (1994) *Mixing*. Springer Verlag, New York.

[7] Ioannides, D.A. (2004) Fixed design regression quantiles for time series. *Statistics and Probability Letters* 68, 235-245.

[8] Khan, S., & J.L. Powell (2001) Two-step estimation of semiparametric censored regression models. *Journal of Econometrics* 103, 73-110.

[9] Koenker, R., & G. Bassett (1978) Regression quantiles. *Econometrica* 46, 33-50.

[10] Koenker, R., & B. Park (1994) An interior point algorithm for nonlinear quantile regression. *Journal of Econometrics* 71, 265-283.

[11] Lewbel, A., & O. Linton (2002) Nonparametric censored and truncated regression. *Econometrica* 70, 765-779.

[12] Mukherjee, K. (2000) Linearization of Randomly Weighted Empiricals under Long Range Dependence with Applications to Nonlinear Regression Quantiles. *Econometric Theory* 16, 301-323.

[13] Newey, W., & J.L. Powell (1990) Efficient estimation of linear and type I censored regression models under conditional quantile restrictions. *Econometric Theory* 6, 295- 317.

[14] Oberhofer, W. (1982) The consistency of nonlinear regression minimizing the L_1 norm. *Annals of Statistics* 10, 316-319.

[15] Oberhofer, W., & H. Haupt (2005) The asymptotic distribution of the unconditional quantile estimator under dependence. *Statistics and Probability Letters*, forthcoming.

[16] Pagan, A.R., & A. Ullah (1999) *Nonparametric econometrics*. Cambridge University Press.

[17] Pfanzagl, J. (1969) On the measurability and consistency of minimum contrast estimator. *Metrika* 14, 249-272.

[18] Powell, J.L. (1984) Least absolute deviations estimation of the censored regression model. *Journal of Econometrics* 25, 303- 325.

[19] Powell, J.L. (1986) Censored regression quantiles. *Journal of Econometrics* 32, 143-155.

[20] Powell, J.L. (1994) Estimation of semiparametric models. In R.F. Engle and D.L. McFadden (eds.), *Handbook of Econometrics*, vol. 4, Elsevier Science, Amsterdam, 2443-2521.