Partial Linear Quantile Regression and Bootstrap Confidence Bands

Wolfgang K. Härdle
Ya’acov Ritov
Song Song

CASE - Center for Applied Statistics and Economics
Humboldt-Universität zu Berlin

Department of Statistics
The Hebrew University of Jerusalem
http://www.case.hu-berlin.de
http://www.stat.huji.ac.il
A 2000-year old story from India...
$\log(\text{Salary}) \sim N(\text{age})$  

Years

“The rich got richer and the poor got poorer!”

Yu et al. (2003)
Quantile Regression

- QR: conditional behavior of a response $Y$
- Median regression = mean regression (symmetric)
- “Gradually developing into a comprehensive strategy for completing the regression prediction” claimed by Koenker & Hallock (2001)
Figure 1: The 0.9-quantile curve, the 0.9-quantile smoother with $h_{0.9} = 1.25$ and 95% confidence bands. QR1

Partial Linear Quantile Regression and Bootstrap Confidence Bands
Example

- **Financial Market & Econometrics**
  - VaR (Value at Risk) tool to measure risk, Lauridsen (2000)
  - Detect conditional heteroscedasticity, Koenker & Bassett (1982)

- **Labor Market**
  - Analyse income of football players w.r.t. different ages, years, and countries, etc
  - Investigate discrimination effects, Buchinsky (1995)

\[
\log (\text{Income}) = A(\text{year, age, etc}) + \beta B(\text{education, gender, nationality, union status, etc}) + \varepsilon
\]

- Inequality analysis
- ...
Quantile Regression

- $l(x) = F_{Y|x}^{-1}(p)$ $p$-quantile regression curve
- $l(x)$ = linear (parametric) form, Koenker & Bassett (1978)
- $l_h(x)$ quantile-smoother

How to decide between functional forms? (global variability of the estimate, peak or valley really a feature?)
Theorem (Härdle and Song (2009))

An approximate $(1 - \alpha) \times 100\%$ confidence band over $[0, 1]$ is

$$l_h(t) \pm (nh)^{-1/2}\{p(1 - p)/\hat{f}_X(t)\}^{1/2}\hat{f}^{-1}\{l(t)|t\} \times \{d_n + c(\alpha)(2\delta \log n)^{-1/2}\} \cdot \{\lambda(K)\}^{1/2},$$

where $c(\alpha) = \log 2 - \log |\log(1 - \alpha)|$ and $\hat{f}_X(t), \hat{f}\{l(t)|t\}$ are consistent estimates for $f_X(t), f\{l(t)|t\}$.

*Emil Julius Gumbel* on BBI:
Challenges

\[ L(\|l_h - l\|_\infty) \]

\[ L^*(\|l_h - l\|_\infty) \]

\[ (\log n)^{-1} \]

\[ \exp\{-\exp(-x)\} \]

Partial Linear Quantile Regression and Bootstrap Confidence Bands
Opportunities

- “Hungarian machine gun", $x \in \mathbb{R}^1$ (KMT)
- Extend this to $x \in \mathbb{R}^d$ and improve band precision?
Outline

1. Motivation ✓
2. Bootstrap Confidence Bands
3. Bootstrap Confidence Bands in PLMs
4. Monte Carlo Study
5. Labour Market Applications
Quantile Regression

- \( \{(X_i, Y_i)\}_{i=1}^{n} \) i.i.d. rv’s, \( x \in J^* = (a, b) \) for some \( 0 < a < b < 1, y \in \mathbb{R} \)
- Suppose \( Y_i = l(X_i) + \varepsilon_i, \varepsilon_i \sim F(\cdot|X_i) \) with \( F(0|X_i) = p \). Both \( l \) & \( F \) are smooth.
- Estimator \( l_h(\cdot) \): the solution of

\[
\frac{\sum_{i=1}^{n} K_h(x - X_i) \mathbf{1}\{Y_i < l_h(x)\}}{\sum_{i=1}^{n} K_h(x - X_i)} < p \leq \frac{\sum_{i=1}^{n} K_h(x - X_i) \mathbf{1}\{Y_i \leq l_h(x)\}}{\sum_{i=1}^{n} K_h(x - X_i)}
\]

- \( S_n \): any slowly varying function (e.g., \( S_n^2 = S_n \) is valid...). \( \lambda_i \) and \( C_i \): generic constants.
Local rate of convergence of $l_h$:
$$\delta_n = h^2 + (nh)^{-1/2} = O(n^{-2/5}) \text{ with } h_n = O(n^{-1/5})$$

Auxiliary estimate $l_g$ with larger bandwidth $g_n = h_n n^{\zeta}$ ($\zeta$: 4/45).

$F(\cdot|X_i)$ estimates: $\hat{F}_i(\cdot) \{\sum_{j=1}^{n} K_h(X_j - X_i)\}^{-1} K_h(X_j - X_i)$ on $Y_j - l_h(X_i)$. 
Check Function

\[ \rho_p(u) = pu1\{u \in (0, \infty)\} - (1 - p)u1\{u \in (-\infty, 0)\} \]

Figure 2: Check function for \( p = 0.9 \), \( p = 0.5 \) and weight function in conditional mean regression.
The Quantile Curve

\[ \rho_p(u) = pu 1\{u \in (0, \infty)\} - (1-p)u 1\{u \in (-\infty, 0)\} \]

\[ l(x) = \arg \min_{\theta} E\{\rho_p(Y - \theta) | X = x\} \]

\[ l_h(x) = \arg \min_{\theta} n^{-1} \sum_{i=1}^{n} \rho_p(Y_i - \theta) K_h(x - X_i) \]

where \( K_h(u) = h^{-1} K(u/h) \) is a kernel (symmetric density function with compact support) with bandwidth \( h \)
Weight Function

\[ \psi(u) = p - 1\{u \in (-\infty, 0)\} \]

\( l_h(x) \) and \( l(x) \): treated as a zero of \( \tilde{H}_n\{l_h(x), x\} \) and \( \tilde{H}\{l(x), x\} \)

where:

\[ \tilde{H}_n\{l_h(x), x\} = 0 : \quad \tilde{H}_n(\theta, x) \overset{\text{def}}{=} n^{-1} \sum_{i=1}^{n} K_h(x - X_i)\psi(Y_i - \theta) \]

\[ \tilde{H}\{l(x), x\} = 0 : \quad \tilde{H}(\theta, x) \overset{\text{def}}{=} \int_{\mathbb{R}} f(x, y)\psi(y - \theta)dy \]
Lemma

[Franke and Mwita (2003), p14] If assumptions (A1, A2, A4) hold, then for any small enough $\varepsilon > 0$,

$$\sup_{|t|<\varepsilon, i=1,\ldots,n, X_i \in J^*} |\hat{F}_i(t) - F(t|X_i)| = O_p\{S_n\delta_n\varepsilon^{1/2} + \varepsilon^2\}. \quad (2)$$

Rationale: No error at $t = 0$ ($\hat{F}_i(0) = F(0|X_i) = p$). For $t \in (0, \varepsilon)$, $\hat{F}_i(t)$, like $l_h$, is based on sample size of $O_p(nh)$, hence the random error is $O_p\{(nh_n)^{-1/2} t^{1/2}\}$, while the bias is $O_p(\varepsilon h^2) = o_p(\delta_n)$. The $S_n$ term takes care of the maximization.
The Bootstrap couple

- $U_1, \ldots, U_n$: i.i.d. uniform [0, 1] rv’s
- Bootstrap sample

$$Y_i^* = l_g(X_i) + \hat{F}_i^{-1}(U_i), \quad i = 1, \ldots, n$$

- Couple with the true conditional distribution:

$$Y_i^\# = l(X_i) + F^{-1}(U_i|X_i), \quad i = 1, \ldots, n.$$

Given $X_1, \ldots, X_n$: $Y_1, \ldots, Y_n$ and $Y_1^\#, \ldots, Y_n^\#$ are equally distributed.
A very close couple

\[ Y_i^* = l_g(X_i) + \hat{F}_i^{-1}(U_i), \quad i = 1, \ldots, n \]

\[ Y_i^\# = l(X_i) + F^{-1}(U_i|X_i), \quad i = 1, \ldots, n. \]

Values of \( Y_i^\# \) and \( Y_i^* \) are meaningful only if \( |U_i - p| < S_n \delta_n \).

By the inverse function theorem around \( p \), we have

\[
\max_{i:|Y_i^\# - l(X_i)| < S_n \delta_n} |Y_i^\# - l(X_i) - Y_i^* + l_g(X_i)| = O_p\{S_n \delta_n^{-3/2}\}.
\]
**How close?**

- $q_{K_i}(Y_1, \ldots, Y_n)$: the local quantile at $X_i$. Assumption A3 gives:
  \[
  \max_{|X_i - x_j| < ch} \left| l_g(X_i) - l_g(x_j) - l(X_i) + l(x_j) \right| = O_p(\delta_n)
  \]

- $l^*_h$ and $l^\#_h$: local bootstrap quantile and its coupled sample analogue. Then
  \[
  l^*_h(X_i) - l_g(X_i) = q_{K_i}[\{ Y_j^* - l_g(x_j) + l_g(x_j) - l_g(X_i) \}]_{j=1}^n
  \]
  \[
  l^\#_h(X_i) - l(X_i) = q_{K_i}[\{ Y_j^\# - l(x_j) + l(x_j) - l(X_i) \}]_{j=1}^n
  \]
  Thus
  \[
  \max_i \left| l^*_h(X_i) - l_g(X_i) - l^\#_h(X_i) - l(X_i) \right| = O_p(\delta_n).
  \]
Bootstrap confidence bands

Bootstrap bootstrapping approximation rate

**Theorem**

*If assumptions *(A1)–(A3) hold, then*

\[
\sup_{x \in J^*} |l^*_h(x) - l_g(x) - l^#_h(x) - l(x)| = O_p(\delta_n) = O_p(n^{-2/5}).
\]

Bootstrap improves the rate of convergence.
Why oversmoothing?

- To handle the bias. Tuning parameter: $g$
- Härdle and Marron (1991), let

$$b_h(x) \overset{\text{def}}{=} E l_h^\#(x) - l(x)$$

$$\hat{b}_{h,g}(x) \overset{\text{def}}{=} E^* l_h^*(x) - l_g(x)$$

- Investigate MSE by decomposing into variance & squared bias

$$E \left[ \left\{ \hat{b}_{h,g}(x) - b_h(x) \right\}^2 \mid X_1, \ldots, X_n \right] = \underbrace{V_n^2}_{\text{Variance}} + \underbrace{B_n^2}_{\text{Bias}^2}$$
Oversmoothing

**Theorem**

*Under some assumptions, for any* $x \in J^*$

$$
E \left[ \left\{ \hat{b}_{h,g}(x) - b_h(x) \right\}^2 \mid X_1, \ldots, X_n \right] \sim h^4(C_1 g^4 + C_2 n^{-1} g^{-5})
$$

*in the sense that the ratio between the RHS and the LHS tends in probability to 1 for some constants* $C_1$, $C_2$.

To minimize MSE, $g = \mathcal{O}(n^{-1/9})$, $g \gg h$, where $h = \mathcal{O}(n^{-1/5})$
The multivariate case

\[ x = (u, v)^\top \in \mathbb{R}^d, \quad v \in \mathbb{R}: \]

\[ \tilde{l}(x) = u^\top \beta + l(v) \]

- Estimation idea: ANOVA
- Partition \([0, 1]\) (for \(v\)) in \(a_n\) intervals \(l_{ni}\) & regard \(l(v)\) as a constant item inside \(l_{ni}\).
Two stage estimation procedure

- Linear quantile regression inside each $l_{ni} +$ Weighted mean yields $\hat{\beta}$:

$$\hat{\beta} = \arg\min_\beta \min_{l_1, \ldots, l_n} \sum_{i=1}^n \psi\{ Y_i - \beta^T U_i - \sum_{j=1}^{a_n} l_j 1(V_i \in l_{ni}) \}$$

- Smooth quantile estimate $\hat{l}_h(v)$ from $(V_i, Y_i - U_i^T \hat{\beta})_{i=1}^n$.

**Theorem**

$\exists$ positive definite matrices $D_n$, $C_n$, s.t.

$$\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{\mathcal{L}} N\{0, p(1-p)D_n^{-1}C_nD_n^{-1}\} \text{ as } n \to \infty.$$
**Uniform consistency of $\hat{l}_h(v)$**

**Lemma**

*Under assumptions (A7) & (A8), we have a.s. as $n \to \infty$*

\[
\sup_{v \in J^*} |\hat{l}_h(v) - l(v)| \leq C_5 \max\{ (nh/\log n)^{-1/2}, h^{\bar{\alpha}} \} \tag{3}
\]

*with another constant $C_5$ not depending on $n$. If additionally $\bar{\alpha} \geq \{ \log(\sqrt{\log n}) - \log(\sqrt{nh}) \} / \log h$, (3) can be further simplified to:*

\[
\sup_{v \in J^*} |\hat{l}_h(v) - l(v)| \leq C_5\{ (nh/\log n)^{-1/2} \}.
\]
Multidimensional Uniform Confidence Bands

- Estimation error for parametric part: $O_p(n^{-1/2})$.
- Bootstrapping approximation error for nonparametric part: $O_p(n^{-2/5})$, dominating!

**Corollary**

Under the assumptions (A1) - (A8), an approximate $(1 - \alpha) \times 100\%$ confidence band over $\mathbb{R}^{d-1} \times [0, 1]$ is

$$ u^\top \hat{\beta} + l_h(v) \pm \left[ \hat{f}\{l(x)|x\} \sqrt{\hat{f}_X(x)} \right]^{-1} d_{\alpha}^*, $$

where $d_{\alpha}^*$ is based on the bootstrap sample (specify later).
How to Bootstrap?

1) Simulate $\{(X_i, Y_i)\}_{i=1}^{n}$, $n = 1000$ w.r.t. $f(x, y)$.

$$f(x, y) = f_{y|x}(y - \sin x)1(x \in [0, 1]),$$

where $f_{y|x}(x)$ is the pdf of $N(0, x)$.

2) Compute $l_h(x)$ of $Y_1, \ldots, Y_n$ and residuals

$\hat{\varepsilon}_i = Y_i - l_h(X_i), \ i = 1, \ldots, n$.

If we choose $p = 0.9$, then $\Phi^{-1}(p) = 1.2816$, $l(x) = \sin(x) + 1.2816\sqrt{x}$ and the bandwidth is $h = 0.05$. 
3) Compute the conditional edf:

\[ F_n(t|x) = \frac{\sum_{i=1}^{n} K_h(x - X_i) 1\{\hat{\varepsilon}_i \leq t\}}{\sum_{i=1}^{n} K_h(x - X_i)} \]

with the quartic kernel

\[ K(u) = \frac{15}{16} (1 - u^2)^2, \quad (|u| \leq 1). \]

4) Generate rv \( \varepsilon_{i,b}^* \sim F_{n|x}, b = 1, \ldots, B \) and construct the bootstrap sample \( Y_{i,b}^*, i = 1, \ldots, n, b = 1, \ldots, B \) as follows:

\[ Y_{i,b}^* = l_g(X_i) + \varepsilon_{i,b}^*, \]

with \( g = 0.2 \).
5) For each bootstrap sample \( \{(X_i, Y_{i,b}^*)\}_{i=1}^n \), compute \( l_h^* \) and the random variable

\[
d_b \overset{\text{def}}{=} \sup_{X \in J^*} \left[ \hat{f} \{ l(x) | x \} \sqrt{\hat{f}_X(x)} \frac{l_h^*(X) - l_g(X)}{\sqrt{\hat{f}_X(x)}} \right].
\] (5)

6) Calculate the \((1 - \alpha)\) quantile \( d_{\alpha}^* \) of \( d_1, \ldots, d_B \).

7) Construct the bootstrap uniform confidence band centered around \( l_h(x) \), i.e. \( l_h(x) \pm \left[ \hat{f} \{ l(x) | x \} \sqrt{\hat{f}_X(x)} \right]^{-1} d_{\alpha}^* \).
Figure 3: The real 0.9 quantile curve, 0.9 quantile estimate with corresponding 95% uniform confidence band from asymptotic theory and confidence band from bootstrapping.
PLM QR

- Bivariate data \( \{(U_i, V_i, Y_i)\}_{i=1}^{n}, n = 8000 \) with:

\[
y = 2u + v^2 + \varepsilon - \Phi(p)
\]

where \( u \in [0, 2], \ v \in [0, 1] \) and \( \varepsilon \) is the standard normal rv.

- The real 0.9-quantile curve \( \tilde{l}(x) = 2u + v^2 \).

- \( h = 0.2 \) & \( g = 0.7 \). For the following specific set of random variables, \( a_n = 20, \ \hat{\beta} = 2.016758 \)
**# of Partitions?**

Figure 4: $\hat{\beta}$ with respect to different $p$ for different # of observations, i.e. $n = 1000$, $n = 8000$, $n = 261148$. 

Partial Linear Quantile Regression and Bootstrap Confidence Bands — logosmall
<table>
<thead>
<tr>
<th>$a_n$</th>
<th>$n = 1000$</th>
<th>$n = 8000$</th>
<th>$n = 261148$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n^{1/3}/8$</td>
<td>5.4 $\times$ 10^{-1}</td>
<td>4.0 $\times$ 10^{-2}</td>
<td>3.6 $\times$ 10^{-3}</td>
</tr>
<tr>
<td>$n^{1/3}/4$</td>
<td>6.1 $\times$ 10^{-1}</td>
<td>3.5 $\times$ 10^{-2}</td>
<td>3.3 $\times$ 10^{-3}</td>
</tr>
<tr>
<td>$n^{1/3}/2$</td>
<td>6.2 $\times$ 10^{-1}</td>
<td>3.6 $\times$ 10^{-2}</td>
<td>3.2 $\times$ 10^{-3}</td>
</tr>
<tr>
<td>$n^{1/3}$</td>
<td>8.0 $\times$ 10^{-1}</td>
<td>3.9 $\times$ 10^{-2}</td>
<td>3.1 $\times$ 10^{-3}</td>
</tr>
<tr>
<td>$n^{1/3} \cdot 2$</td>
<td>4.9 $\times$ 10^{-1}</td>
<td>3.6 $\times$ 10^{-2}</td>
<td>2.9 $\times$ 10^{-3}</td>
</tr>
<tr>
<td>$n^{1/3} \cdot 4$</td>
<td>4.9 $\times$ 10^{-1}</td>
<td>3.6 $\times$ 10^{-2}</td>
<td>2.8 $\times$ 10^{-3}</td>
</tr>
<tr>
<td>$n^{1/3} \cdot 8$</td>
<td>3.4 $\times$ 10^{-3}</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1: SSE of $\hat{\beta}$ with respect to $a_n$ for different numbers of observations.

☐ Suggest $a_n = n^{1/3}$ (cost / performance)
Figure 5: Nonparametric part smoothing, real 0.9 quantile curve with respect to $v$, 0.9 quantile smoother with corresponding 95% bootstrap uniform confidence band.
How income depends on age w.r.t. different education levels?

Relation: $\log (\text{Wage}) \sim \beta \cdot \text{Education} + I(\text{Age})$

Administrative data from the German National Pension Office

Male, born 1939 $\sim$ 1942, sample 25 - 59, full-time, begin receiving a pension in 2004 $\sim$ 2005
- Education categories: -9 “no answer", 1 “low education", 2 “apprenticeship" and 3 “university"
- High quality data: true panel + \( n = 261148 \) observations!
- Quartic kernel, \( h = 0.059 \) (after rescaling)
Figure 6: \( \hat{\beta} \) corresponding to different quantiles.
Figure 7: 95% uniform confidence bands for 0.05-quantile smoothers with 4 different education levels
On average (Median) - no significant effect

Figure 8: 95% uniform confidence bands for 0.50-quantile smoothers with 4 different education levels
Many high income people - no high education

Figure 9: 95% uniform confidence bands for 0.90-quantile smoothers with 4 different education levels
Why?

- Smart, no need go to school
- Be scientist after PHD graduation
- Poor, not continue school, but hard working & know a lot from practice
- Education may make people less creative
- ...

Our normal impression:
\[ E(y|\nu, u = \text{Low education}) < E(y|\nu, u = \text{University}) \]
Real effect of education for income?

"Concentrated, away from risk" (the game 0/2 or 1)
A 2000-year old story from India . . .
Causality Test?

Risky to claim: Low education leads causes high income for high income labour!

- “High education - causality effect - low income labour - earn more"
- “Low education - not causality effect - high income labour earn more"
- Causality test in quantile for β-mixing time series: Jeong, Härdle and Song (2009).
- Panel data causality test (further research)
Drawbacks

- Very rich people maybe not recorded in the pension system
- Maybe not same retirement time
- Not exactly i.i.d. (further research)
- ...

Partial Linear Quantile Regression and Bootstrap Confidence Bands — logosmall
Sth must keep in mind!

- You are dealing with 70-year old people now!
- Time flies (technology level ↑), more and more high income jobs require high educated people. Time variation of the $\hat{\beta}$? further research.
References

- **J. Franke and P. Mwita**
  Nonparametric Estimates for conditional quantiles of time series
  *Report in Wirtschaftsmathematik 87, University of Kaiserslautern, 2003.*

- **P. Hall**
  On convergence rates of suprema

- **Jinyong Hahn**
  Bootstrapping Quantile Regression Estimators
References

W. Härdle and J. S. Marron
Bootstrap simultaneous error bars for nonparametric regression

W. Härdle and S. Song
Confidence bands in quantile regression
References

J. Horowitz
Bootstrap methods for median regression models

K. Jeong, W. Härdle and S. Song
A Consistent Nonparametric Test for Causality in Quantile
_Revise and resubmit to Econometric Theory_, 2009.

R. Koenker and K. F. Hallock
Quantile regression

G. Tusnady
A remark on the approximation of the sample distribution function in the multidimensional case
References

K. Yu, Z. Lu and J. Stander
Quantile regression: applications and current research areas
Appendix - Assumptions

A1. \( X_1, \ldots, X_n \) are an i.i.d. sample, and \( f_X(x) \geq \lambda_0 \). The quantile function satisfies: \( |l'(\cdot)| \leq \lambda_1 \), \( |l''(\cdot)| \leq \lambda_2 \).

A2. \( F(t|x) \) have a density, \( f(t|x) \geq \lambda_3 > 0 \), continuous in \( x \), and in \( t \) in the neighborhood of 0. That is, for some \( A(\cdot) \) and \( f_0(\cdot) \)

\[
F(t|x') = p + f_0(x)t + A(x)(x' - x) + R(t, x'; x),
\]

where \( \sup_{t,x,x'} \frac{|R(t,x';x)|}{t^2 + |x' - x|^2} < \infty \).
Note that by Assumption A1, $l_h(x)$ is the quantile of a discrete distribution.
This distribution is equivalent to a sample size of $O_p(nh)$ from a distribution with $p$-quantile whose biased is $O_p(h^2)$ relative to the true value.
Let $\delta_n$ be the local rate of convergence of the function $l_h$, essentially $\delta_n = h_n^2 + (nh_n)^{-1/2} = O(n^{-2/5})$, with $h_n = O(n^{-1/5})$. 
Appendix

A3. The estimate $l_g$, satisfies:

\[
\sup_{x \in J^*} |l''_g(x) - l''(x)| = o_p(1),
\]
\[
\sup_{x \in J^*} |l'_g(x) - l'(x)| = o_p(\delta_n/h)
\]  

(7)

Note that there is no $S_n$ term in (7) exactly because the bandwidth $g_n$ used to calculate $l_g$ is slightly larger than that used for $l_h$. As a result $l_g$ has a slightly worse rate of convergence (as an estimator of the quantile function), but its derivatives converge faster.

We assume:

(A4). $f_X(x)$ is twice continuously differentiable and $f(t|x)$ is uniformly bounded in $x$ and $t$ by, say, $\lambda_4$. 

Partial Linear Quantile Regression and Bootstrap Confidence Bands
(A7). The conditional densities $f(\cdot | y), y \in \mathbb{R}$, are uniformly local Lipschitz continuous of order $\tilde{\alpha} (\text{ulL}-\tilde{\alpha})$ on $J$, uniformly in $y \in \mathbb{R}$, with $0 < \tilde{\alpha} \leq 1$, and $(nh)/\log n \to \infty$.

(A8). \[ \inf_{v \in J^*} \left| \int \psi \{ y - l(v) + \varepsilon \} dF(y|v) \right| \geq \tilde{q}|\varepsilon|, \quad \text{for } |\varepsilon| \leq \delta_1, \]

where $\delta_1$ and $\tilde{q}$ are some positive constants, see also [?]. This assumption is satisfied if there exists a constant $\tilde{q}$ such that $f\{l(v)|v\} > \tilde{q}/p, x \in J$. 