Applied Quantile Regression: Microeconometric, Financial, and Environmental Analyses

Inaugural-Dissertation
zur Erlangung des Doktorgrades
der Wirtschaftswissenschaftlichen Fakultät
der Eberhard-Karls-Universität Tübingen

vorgelegt von
Niels Schulze
aus Kaiserslautern

Tübingen
2004

2

Dekan:	Prof. Dr. Jörg Baten
Erstkorrektor:	Prof. Dr. Gerd Ronning
Zweitkorrektor:	Prof. Dr. Joachim Grammig
Tag der mündlichen Prüfung:	25. November 2004

Abstract

In 1978, Roger Koenker and Gilbert Bassett, Jr. introduced a new econometric estimation method and entitled it quantile regression. Since then, many subsequent authors have elaborated and extended the underlying theoretical framework. Other contributions have successfully applied the procedure to a wide range of problems from a variety of scientific branches.

This study presents the basic features of quantile regression along with some important properties and a selection of significant extensions and applications. Subsequently, the procedure is used in three new and original empirical regression settings to demonstrate the universality and flexibility of the approach.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

What the regression curve does is give a grand summary for the averages of the distributions corresponding to the set of x's. We could go further and compute several different regression curves corresponding to the various percentage points of the distributions and thus get a more complete picture of the set. Ordinarily this is not done, and so regression often gives a rather incomplete picture. Just as the mean gives an incomplete picture of a single distribution, so the regression curve gives a correspondingly incomplete picture for a set of distributions.

Frederick Mosteller and John W. Tukey (1977)

Already a quick analysis of some arbitrarily chosen contributions to the applied econometrics literature immediately reveals the fact that the overwhelming majority of empirical regression studies are based on the analysis of the (conditional) mean of the regressand. However, as Mosteller and Tukey (1977) paraphrased above, this confinement might give an incomplete picture and thus can lead to possibly wrong conclusions as soon as not all assumptions of the classical linear regression model hold.

A solution to the risen question of how to "go further" was proposed by Roger Koenker and Gilbert W. Bassett, Jr. (1978). They introduced a new method labeled "quantile regression" that allows the estimation of the *entire* distribution of the response vari-

able conditional on any set of (linear) regressors. In other words, the calculation of a single value (the conditional mean) is replaced by the computation of a whole set of numbers (the conditional quantiles) which are able to give a more complete picture of the underlying interrelations.

Furthermore, the quantile regression procedure inherits some additional advantages over least squares regression, as for example equivariance to monotone transformations of the response variable, robustness against outliers of the regressand, and higher efficiency for a wide range of error distributions.

Of course, as Milton Friedman (1975) put it in his famous dictum, "there is no such thing as a free lunch". In the case of quantile regression, one can argue that there is (at least at first sight) a higher computational burden. In contrast to the least squares case, the objective function is not differentiable at the origin, so no general closed solution can be given. However, the quantile regression problem can be shown to have a linear programming representation which substantially eases the calculation. With some additional modifications, the method can be made competitive in computation time even for very large data sets. Another potential criticism point of Koenker and Bassett's (1978) original proposal could be the missing of a well-established general asymptotic theory and thus a lack of appropriate inference procedures. Yet, a variety of subsequent authors have addressed that issue and comprehensively closed the gap.

This study is organized as follows: chapter two starts with the presentation of some basic principles of the quantile regression approach along with a few remarks on the computation of the estimates. Subsequently, several important properties of conditional quantiles are discussed and two examples try to clarify the presentation and interpretation of the results. Next, the asymptotic theory and inferential strategies for quantile regression are addressed, including a comparison of different procedures. An introduction to several extensions and empirical applications of estimated conditional quantiles follows. Chapter two concludes with a short retrospective on the situation *before* the invention of quantile regression.

The subsequent chapters (three to five) contain three new and original empirical implementations of quantile regression in three different scientific branches to demonstrate the variability and vast application possibilities of Koenker and Bassett's (1978) seminal proposal.

Chapter three contains an econometric demand analysis of cross-section micro data taken from a consumer panel. The use of the quantile regression model enables the analysis of consumer behavior conditional on the intensity of consumption. Several interesting results are revealed.

In chapter four, a new measure for the degree of linkages between financial markets, called coexceedances, is introduced. Subsequently, these coexceedances are used in a quantile regression setting to contrast contagion against interdependence. Multiple new insights are gained and presented.

Chapter five applies the quantile regression procedure to the analysis of environmental connectivities. The impact of different meteorological influence factors on the conditional distribution of daily maximum ozone concentrations is considered. The obtained results vary significantly for different ozone regimes.

The conclusions in chapter six summarize our findings and give an outlook to future prospects on the successful application of quantile regression.

Chapter 2

Quantile Regression

2.1 Basics

In this section, we seek to present the fundamental principals of quantile regression. We start with some basic definitions around the quantile function. Next, we consider empirical quantiles and present an alternative formulation. Subsequently, the concept is extended to a regression setting. Finally, some computational issues of quantile regression are addressed.

2.1.1 The quantile function

For any τ in the interval $(0, 1)$ and any (discrete or continuous) random variable Y, the τ-th *quantile* of Y can be defined as any number $\xi_\tau \in \mathbb{R}$ that fulfils[1]

$$P(Y < \xi_\tau) \ \leq \ \tau \ \leq \ P(Y \leq \xi_\tau) \tag{2.1}$$

It can be seen that a solution to (2.1) always exists and that it is unique if Y is a continuous random variable (in this case the two probabilities given in (2.1) coincide).

---

[1]The underlying intuition is very simple: at least τ percent of the probability mass of Y is lower than or equal to ξ_τ, and at least $(1 - \tau)$ percent of the probability mass of Y is higher than or equal to ξ_τ.

Table 2.1: Distribution function $F_Y(y)$ and corresponding quantile function $Q_Y(\tau)$

Distribution	$F_Y(y)$	$Q_Y(\tau)$
Exponential	$1 - e^{-y}$	$-\ln(1 - \tau)$
Gaussian	$\Phi(y)$	$\Phi^{-1}(\tau)$
Gumbel	$1 - \exp(-e^z)$	$\ln\big(-\ln(1 - \tau)\big)$
Logistic	$\dfrac{e^y}{1 + e^y}$	$\ln\left(\dfrac{\tau}{1 - \tau}\right)$
Pareto	$1 - (\frac{\alpha}{y})^\beta$	$\alpha(1 - \tau)^{-\frac{1}{\beta}}$
Uniform	y	τ

If Y is a discrete variable, the solution to (2.1) is for some τ a closed interval of the real line. To circumvent any problems arising from this non-uniqueness, we will from now on always choose the smallest element of the solution set. With this convention, we can use the so-called (right-continuous) cumulative distribution function (CDF)

$$F_Y(y) = P(Y \leq y) \qquad (2.2)$$

to define the (left-continuous) *quantile function*

$$Q_Y(\tau) = F_Y^{-1}(\tau) = \inf\{y \mid F_Y(y) \geq \tau\} \qquad 0 < \tau < 1 \qquad (2.3)$$

In letters: for any τ in the interval $(0, 1)$, the quantile function $Q_Y(\tau)$ provides the τ-th (unconditional) quantile of Y. The function has several interesting properties, e.g. for any monotonically increasing and left-continuous function g, it can be shown (see e.g. Peracchi (2001)) that for every τ in the interval $(0, 1)$

$$P\Big(Y \leq Q_Y(\tau)\Big) = P\Big(g(Y) \leq g(Q_Y(\tau))\Big) = \tau \qquad (2.4)$$

Table 2.1 uses this property and exemplary lists some common distribution functions along with their quantile function analogues.

For a continuous random variable Y, the so-called probability density function (PDF) is

defined as the derivative of the distribution function:[2]

$$f_Y(y) = \frac{dF_Y(y)}{dy} \tag{2.5}$$

The analogue for the quantile function was (to our knowledge) first mentioned by Tukey (1965) and is defined as

$$s_Y(\tau) = \frac{dQ_Y(\tau)}{d\tau} \tag{2.6}$$

Tukey (1965) called $s_Y(\tau)$ the *sparsity function*. Parzen (1979) labeled it the *quantile-density function*. Note that with the derivative of the identity $F_Y(F_Y^{-1}(\tau)) = \tau$, we have

$$\frac{dF_Y(F_Y^{-1}(\tau))}{d\tau} = f_Y(F_Y^{-1}(\tau)) \cdot \frac{dF_Y^{-1}(\tau)}{d\tau} = 1 \tag{2.7}$$

and the sparsity function can thus also be expressed as the reciprocal of the density function, evaluated at the quantile of interest:[3]

$$s_Y(\tau) = \frac{dF_Y^{-1}(\tau)}{d\tau} = \frac{1}{f_Y(Q_Y(\tau))} \tag{2.8}$$

We will come back to the sparsity function in section 2.3.

2.1.2 Empirical quantiles

If a random sample $Y_1, Y_2, \ldots, Y_n$ is taken, the so-called *empirical distribution function* is defined as the quotient of the number of observations lower than or equal to the value of interest and the total number of observations:

$$\widehat{F}_Y(y) = \frac{\#(Y_i \leq y)}{n} \tag{2.9}$$

---

[2]In the discrete case, it is defined as $f(y) = P(Y = y)$.

[3]Parzen (1979) named $f_Y(Q_Y(\tau))$ the *density-quantile function* and additionally introduced its derivative $-df_Y(Q_Y(\tau))/d\tau$ as the *score function* of the probability density.

In analogy to (2.3) we can define the *empirical quantile function* as

$$\widehat{Q}_Y(\tau) \;=\; \widehat{F}_Y^{-1}(\tau) \;=\; \inf\{y \mid \frac{\#(Y_i \le y)}{n} \;\ge\; \tau\} \qquad\qquad 0 < \tau < 1 \qquad\qquad (2.10)$$

It can easily be seen from equation (2.10) that in order to obtain the desired quantile, one first has to sort and rank the observed sample and then check at which observation the threshold is reached. This procedure was state of the art until Koenker and Bassett (1978) proposed a complete new and different method to calculate the quantile in question:

$$\widehat{Q}_Y(\tau) \;=\; \operatorname*{argmin}_{\xi_\tau \in \mathbb{R}} \left\{ \sum_{i \in \{i \mid Y_i \ge \xi_\tau\}} \tau \,|Y_i - \xi_\tau| \;\;+\; \sum_{i \in \{i \mid Y_i < \xi_\tau\}} (1 - \tau)\,|Y_i - \xi_\tau| \right\} \qquad (2.11)$$

To paraphrase it, the concept of *sorting* has been replaced by *optimizing* a (weighted) loss function. All observations greater than the unknown optimal value (to be more precise the absolute differences between the observations and the optimum) are weighted with τ, all observations below the optimum are weighted with $(1-\tau)$. Koenker and Bassett (1978) remarked that "the case of the median ($\tau = 1/2$) is, of course, well known, but the general result has languished in the status of curiosum".[4]

For our further proceeding, we use the indicator function ($I(A) = 1$ if A is true, and $I(A) = 0$ otherwise) to introduce the so-called *check function*

$$\rho_\tau(u) = u(\tau - I(u < 0)) \qquad\qquad 0 < \tau < 1 \qquad\qquad (2.12)$$

The check function allows us to reformulate the objective function of (2.11) as a single expression:

$$\widehat{Q}_Y(\tau) \;=\; \operatorname*{argmin}_{\xi_\tau \in \mathbb{R}} \sum_i \rho_\tau(Y_i - \xi_\tau) \qquad\qquad (2.13)$$

Before we show that equations (2.11) and (2.13) really provide the desired quantile, we present a little example to illustrate the intuition of Koenker and Bassett's (1978)

---

[4]For $\tau = 1/2$, equation (2.11) simplifies to $\widehat{Q}_Y(0.5) = \operatorname*{argmin}_{\xi_{0.5} \in \mathbb{R}} \sum_i |Y_i - \xi_{0.5}|$.

Table 2.2: Example. The table depicts the evaluation of the objective function in equation (2.11) for a random sample taken from a standard normal distribution for $\tau = 0.25$.

i	y_i	leftsum	rightsum	objfun
1	-1.491088	6.691236	0	6.691236
2	-1.309872	5.830457	.1359124	5.966370
3	-1.114695	4.952164	.4286767	5.380841
4	-1.011688	4.514384	.6604424	5.174827
5	-.9411827	4.232362	.8719594	5.104321
6	-.6687527	3.210749	1.893572	5.104321
7	-.5695546	2.863556	2.339963	5.203519
8	-.2859412	1.941812	3.828934	5.770746
9	-.2598439	1.863521	3.985517	5.849038
10	-.2188201	1.750705	4.262428	6.013133
11	-.1958248	1.693217	4.434893	6.128110
12	-.0433352	1.350115	5.692932	7.043047
13	.0225686	1.218308	6.286066	7.504374
14	.1109507	1.063639	7.147792	8.211431
15	.2860689	.8009616	8.986532	9.787494
16	.4537083	.5914124	10.87248	11.46389
17	.8690140	.1761066	15.85614	16.03225
18	.9308267	.1297471	16.64426	16.77400
19	1.042856	.0737323	18.15666	18.23039
20	1.337785	0	22.35940	22.35940

seminal proposal. We drew twenty observations from a standard normal distribution and sorted them in ascending order. The second column of table 2.2 lists the drawn sample observations y_i. Columns three and four present the two sums in equation (2.11) evaluated at each observation for $\tau = 0.25$ (the first quartile). The last column finally lists the resulting value of the composite objective function. Figure 2.1 presents the outcome in a graphical manner.

Table 2.2 and Figure 2.1 clarify that, in the given example, any value in the interval $[-.9411827, -.6687527]$ minimizes the objective function in equation (2.11) for $\tau = 0.25$. This is completely in line with the demanded properties of a quantile stated in equation (2.1). As already noticed, we will always choose the smallest value of the interval to guarantee the left-continuity of the quantile function. Moreover, it can be seen that a non-unique solution to equation (2.11) only arises if $n\tau$ is an integer, so there is a good chance that it will not occur too often in practical applications.

Figure 2.1 further shows that the objective function is convex and piecewise linear with kinks at the observed y_i's. At each observation, the slope of the loss function changes
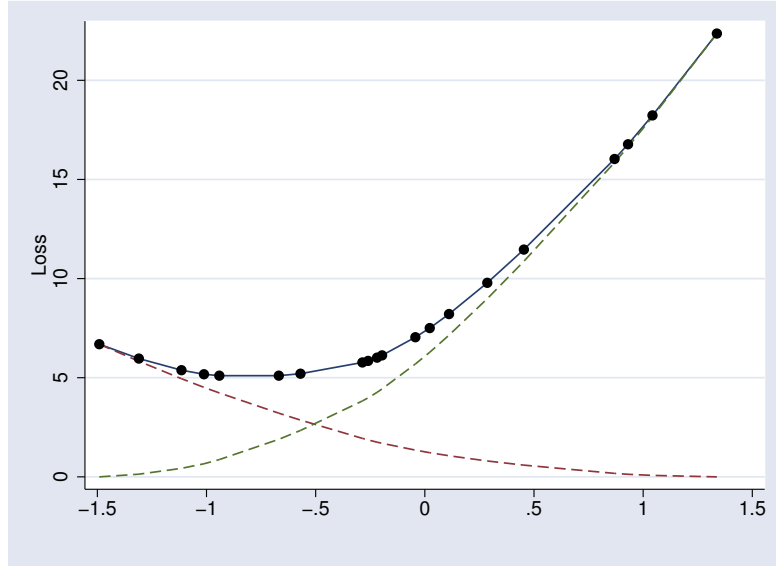
Figure 2.1: Example. The graph pictures the values of table 2.2. The two dashed lines show the two weighted sums of equation (2.11), the solid line presents the composed loss function, all for $\tau = 0.25$.

by exactly 1, as the contribution of the value in question changes from $-(1-\tau)$ to $+\tau$.

So, in our example, the slope of the objective function ranges from -5 to $+15$.

We still have omitted the general proof that the optimum of the loss function provides the desired quantile. Turning back to the theoretical case, we can express the expected loss as

$$E\Big(\rho_\tau(Y - \xi_\tau)\Big) \;=\; \tau \int_{\xi_\tau}^{\infty} (y - \xi_\tau)dF(y) \;-\; (1 - \tau)\int_{-\infty}^{\xi_\tau}(y - \xi_\tau)dF(y) \qquad (2.14)$$

Taking the derivative with respect to ξ_τ yields:[5]

$$
\begin{aligned}
\frac{\partial E\Big(\rho_\tau(Y - \xi_\tau)\Big)}{\partial \xi_\tau} \;&=\; \tau\,\frac{\partial \int_{\xi_\tau}^{\infty}(y - \xi_\tau)\partial F(y)}{\partial \xi_\tau} \;-\; (1 - \tau)\,\frac{\partial \int_{-\infty}^{\xi_\tau}(y - \xi_\tau)\partial F(y)}{\partial \xi_\tau} \\[2mm]
&=\; -\tau\int_{\xi_\tau}^{\infty}\partial F(y) \;+\; (1 - \tau)\int_{-\infty}^{\xi_\tau}\partial F(y) \\[2mm]
&=\; -\tau(1 - F(\xi_\tau)) \;+\; (1 - \tau)F(\xi_\tau) \\[2mm]
&=\; F(\xi_\tau) \;-\; \tau \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (2.15)
\end{aligned}
$$

---

[5]In the second step, we use twice a variant of the Leibniz rule (see Chen (2001) or Pagan and Ullah (1999)): $\partial \int_{-\infty}^{g(x)} \psi(x,y)dy \,/\, \partial x \;=\; \int_{-\infty}^{g(x)}(\partial \psi/\partial x)dy \;+\; (\partial g/\partial x)\psi(x, g(x))$

Setting (2.15) to zero, it can be seen that the (convex) expected loss function is in fact minimized if and only if ξ_τ fulfils $F(\xi_\tau) = \tau$.

2.1.3 Regression quantiles

Having read about the alternative method to determine empirical quantiles in the last subsection, one might ask why the new formulation should be used. It seems to be more complicated (we have not talked about computational issues yet, but it is in fact a bit burdensome), so where is the surplus? The answer is clear and simple: In contrast to the classical sorting and ranking approach, the optimization procedure can be extended to regression settings.

Consider a classical linear regression model:

$$y_i = x_i'\beta + u_i \qquad i = 1, \ldots, n \qquad (2.16)$$

If one assumes that the expected value of the error term conditional on the regressors is zero ($E(u_i|x_i) = 0$), then the *conditional mean* of y_i with respect to x_i is

$$E(y_i|x_i) = x_i'\beta \qquad (2.17)$$

The parameter vector β can be estimated by the well-known method of least squares:

$$\widehat{\beta} = \operatorname*{argmin}_{\beta \in \mathbb{R}^K} \sum_i (y_i - x_i'\beta)^2 \qquad (2.18)$$

A solution to (2.18) is given by $\widehat{\beta} = (X'X)^{-1}X'y$, see e.g. Greene (2002).

Let us now assume that $y_i = x_i'\beta_\tau + u_{i,\tau}$ and that not the expected value, but the τ-th quantile of the error term conditional on the regressors is zero ($Q_\tau(u_{i,\tau}|x_i) = 0$).[6] Then it is ready to see that the τ-th *conditional quantile* of y_i with respect to x_i can be written

---

[6]We write $Q_\tau(u_{i,\tau}|x_i)$ instead of $Q_{u_{i,\tau}}(\tau|x_i)$ or $Q_{u_i(\tau)}(\tau|x_i)$ to underline the analogy to equation (2.17).

as

$$Q_\tau(y_i|x_i) = x_i'\beta_\tau \tag{2.19}$$

Assembling equations (2.11), (2.13) and (2.19), it should come to no surprise that for any τ in the interval $(0, 1)$, the parameter vector β_τ can be estimated by

$$\widehat{\beta}_\tau = \underset{\beta_\tau \in \mathbb{R}^K}{\operatorname{argmin}} \left\{ \sum_{i \in \{i|y_i \geq x_i'\beta_\tau\}} \tau\, |y_i - x_i'\beta_\tau| \quad + \sum_{i \in \{i|y_i < x_i'\beta_\tau\}} (1-\tau)\, |y_i - x_i'\beta_\tau| \right\}$$

$$= \underset{\beta_\tau \in \mathbb{R}^K}{\operatorname{argmin}} \sum_i \rho_\tau(y_i - x_i'\beta_\tau) \tag{2.20}$$

In letters: all observations above the estimated hyperplane given by $X\widehat{\beta}_\tau$ (again to be precise, the absolute difference between y_i and $x_i'\widehat{\beta}_\tau$) are weighted with τ, all observations below the estimated hyperplane are weighted with $(1-\tau)$. Again, the special case of the conditional median ($\tau = 0.5$) is well known and can be calculated by[7]

$$\widehat{\beta}_{0.5} = \underset{\beta_{0.5} \in \mathbb{R}^K}{\operatorname{argmin}} \sum_i |y_i - x_i'\beta_{0.5}| \tag{2.21}$$

Yu, Lu, and Stander (2001) presented alternative definitions of regression quantiles (they introduced four formulations based on the conditional distribution function, the check function, a regression model, and an asymmetric Laplace density function) and showed the equivalence of the different approaches. To our surprise, they cut out their nice result in the journal version (see Yu, Lu, and Stander (2003)).

2.1.4 Computation

In contrast to the least squares case, equation (2.20) cannot be solved explicitly since the check function is not differentiable at the origin. However, after a slight modification, it

---

[7]In the literature, a "bewildering variety of names" (Bassett and Koenker (1978)) have been proposed for the conditional median regression, as for example least absolute error (LAE), least absolute deviation (LAD), least absolute residuals (LAR), [least] mean absolute deviation (MAD), minimum sum [of] absolute error (MSAE) or just ℓ_1 regression.

can be shown to have a linear programming representation (see e.g. Buchinsky (1996, 1998b), Koenker and Portnoy (1999) or Cizek (2003)).

If one rewrites y_i as a function of only positive elements

$$y_i = \sum_{k=1}^{K} x_{ik}\beta_{k,\tau} + u_{i,\tau} = \sum_{k=1}^{K} x_{ik}(\beta_{k,\tau}^1 - \beta_{k,\tau}^2) + (\epsilon_{i,\tau} - \nu_{i,\tau}) \tag{2.22}$$

with $\beta_{k,\tau}^1 \geq 0$, $\beta_{k,\tau}^2 \geq 0$, $k = 1, \ldots, K$, and $\epsilon_{i,\tau} \geq 0$, $\nu_{i,\tau} \geq 0$, $i = 1, \ldots, n$, then the solution to (2.20) is reduced to the solution of the following problem:

$$\min_{\beta_{k,\tau}^1,\beta_{k,\tau}^2,\epsilon_{i,\tau},\nu_{i,\tau}} \sum_{i=1}^{n} \tau\epsilon_{i,\tau} + (1-\tau)\nu_{i,\tau} \tag{2.23}$$

subject to: $y_i = \sum_{k=1}^{K} x_{ik}(\beta_{k,\tau}^1 - \beta_{k,\tau}^2) + (\epsilon_{i,\tau} - \nu_{i,\tau}), \quad \beta_{k,\tau}^1, \beta_{k,\tau}^2, \epsilon_{i,\tau}, \nu_{i,\tau} \geq 0(\forall i, k)$

Finally, by setting $A = (X, -X, I, -I)$, $z = (\beta_\tau^{1\prime}, \beta_\tau^{2\prime}, \epsilon_\tau{}', \nu_\tau{}')$, and $c = (0', 0', \tau\iota', (1-\tau)\iota')'$, problem (2.23) can be written as the *primal problem* of linear programming:

$$\min_{z} c'z \qquad \text{subject to:} \qquad Az = y \quad (z \geq 0) \tag{2.24}$$

The according *dual problem* is given by

$$\max_{w} w'y \qquad \text{subject to:} \qquad w'A \leq c' \tag{2.25}$$

with the dual variable $w \in [\tau - 1, \tau]^n$. If the design matrix X is of full column rank, both the primal and the dual problem have feasible solutions with equal optimal values ($\min c'z = \max w'y$), see Buchinsky (1998b).

Barrodale and Roberts (1973, 1974) proposed a modified simplex algorithm for the efficient estimation of the conditional median.[8] Koenker and d'Orey (1987, 1994) generalized the approach to allow for the computation of conditional quantiles. Their algorithm turned out to be competitive to least squares estimation in calculation time for small to

---

[8]A similar algorithm was also developed by Bartels and Conn (1980a, 1980b).

medium numbers of observations (say, n up to 1000), but less favorable for larger data sets (it takes for example up to 50 times longer than least squares for $n = 50000$).

To eliminate this "inconvenience", Portnoy and Koenker (1997) developed an interior point algorithm for the computation of regression quantiles. The substitution of the (non-differentiable) objective function through a (differentiable) log barrier formulation, combined with an adequate preprocessing method, enables the "Laplacian tortoise" (quantile regression) to keep up with the "Gaussian hare" (least squares estimation) in calculation time even for very large data sets. See also Koenker and Portnoy (1999) and Koenker (2000).

Several software packages directly contain quantile regression routines in their core package, as for example SHAZAM, EASYREG, BLOSSOM (see Cade and Richards (2001)), XPLORE (see Cizek (2003)) or STATA (see Stata Corporation (2003)[9]). Furthermore, a tremendous number of adaptations for other programs can be found in the internet.[10] Finally, we should mention THE R PROJECT FOR STATISTICAL COMPUTING (see Venables and Smith (2003)) which is free of charge and includes a quantile regression package written by Roger Koenker (see Koenker (2004)). So, in our opinion, any econometrician willing to incorporate quantile regression into his "methodological canon" should have no problem (and is strongly encouraged) to do so.

2.2 Properties

In the last section, we have presented the basic fundamentals of the quantile regression approach. Now, we want to study several important properties of the method. We start by giving some rather general comments to clarify a few open questions from section one. Subsequently, a couple of significant characteristics are discussed and contrasted against other econometrical methods. Finally, we provide several detailed examples

---

[9]For more details on the STATA algorithms see also Rogers (1992, 1993) and Gould (1992, 1997).

[10]See for example *http://www.stat.psu.edu/~dhunter/qrmatlab/* for an implementation in MATLAB.

to elucidate the characteristics of the results obtained from the application of quantile regression in practice. We also give some hints on the graphical presentation of these results.

2.2.1 General comments

As already pointed out, in a linear quantile regression model, for any τ in the interval $(0, 1)$, the estimated conditional quantile of y with respect to a regressor matrix X can be calculated as

$$\widehat{Q}_\tau(y|X) = X\widehat{\beta}_\tau \qquad \text{with} \qquad \widehat{\beta}_\tau = \operatorname*{argmin}_{\beta_\tau \in \mathbb{R}^K} \sum_i \rho_\tau(y_i - x_i'\beta_\tau) \qquad (2.26)$$

So, by varying the value of τ, the quantile regression method enables us to evaluate the *entire* conditional distribution of the regressand. This stands in sharp contrast to the least squares approach which provides us only with a single value, namely the conditional mean.[11] It is evident that the flexible structure of the QR model is able to detect some forms of heteroscedasticity in the data by analyzing several quantiles. Of course, as in mean regression, the application of an appropriate weighting scheme creates opportunities for improved efficiency, see section 2.4.

This analysis of several quantiles leads us to the question of "how many different solutions to equation (2.26) can be found for a given problem". As we have already seen, the answer is easy for the ordinary sample quantiles. With n observations, the objective function in (2.13) has exactly n breakpoints where the primal solution flips from one basis to another. Furthermore, with respect to τ, the n solutions are equally spaced on the interval $[0, 1]$ with each distinct order statistic occupying an interval of length exactly $1/n$.

The situation is a bit more complicated for the regression quantiles. Both the num-

---

[11]Koenker's (2000) comment "There is more to econometric life than can be captured by the philosophy of the Gaussian location shift" nicely paraphrases this limitation.

ber J and the locations of the distinct τ_j's depend in a complicated way on the design configuration as well as the observed response (see e.g. Bassett and Koenker (1982)). While Koenker and d'Orey (1987) only provided a rule-of-thumb ($2n < J < 3n$), Portnoy (1991b) showed that the number of breakpoints J is of order $\mathcal{O}(n \log n)$ under some mild conditions. Fortunately, we do not have to conduct the whole calculation procedure described in the last section for each quantile. In contrast, by means of parametric programming, it is possible to evaluate the *entire* quantile regression process in roughly $n \log n$ simplex pivots by subsequently jumping from one vertex to its adjacent one (see e.g. Koenker and Portnoy (1999)). An implementation of this procedure is included in Koenker's (2004) quantile regression package for THE R PROJECT.

It is self-evident that, as long as they are not parallel (which is the "boring" case), the estimated quantile hyperplanes of y conditional on X cross each other at some place. This leads to the rather absurd result that the estimated conditional value of y is higher at a lower quantile and vice versa. In practice, this fact is much less a problem than it seems to be. Usually, the crossing only occurs at the remote region of the design space (if not even beyond), where *any* statistical assertion should be treated very cautiously. Even better, Bassett and Koenker (1982) showed that $\widehat{Q}_\tau(y|X)$ is *always* non-decreasing in τ at the centroid of the design (where all regressors are at their mean). Anyone not satisfied with this might refer to He (1997) who proposed a restricted version of regression quantiles that do not cross.

Coming back to the estimation of a single conditional quantile, Koenker and Portnoy (1999) stated that in a model with K regressors, there are exactly K residuals with value zero if there is no degeneracy.[12] In this case, the proportion of negative residuals N^- is approximately τ:

$$\frac{N^-}{n} \leq \tau \leq \frac{N^- + K}{n} \tag{2.27}$$

---

[12] Degeneracy can occur if the y_i's are discrete and leads to more than K zero residuals.

and the proportion of positive residuals N^+ is roughly $(1 - \tau)$:

$$\frac{N^+}{n} \leq 1 - \tau \leq \frac{N^+ + K}{n} \tag{2.28}$$

In least squares regression, it is common to calculate the so-called goodness-of-fit measure

$$R^2 = \frac{\min \sum_i (x_i'\widehat{\beta} - \bar{y})^2}{\min \sum_i (y_i - \bar{y})^2} = 1 - \frac{\min \sum_i (y_i - x_i'\widehat{\beta})^2}{\min \sum_i (y_i - \bar{y})^2} \tag{2.29}$$

Koenker and Machado (1999) proposed a similar measure for quantile regression models[13]

$$R^1(\tau) = \frac{\min \sum_i \rho_\tau (x_i'\widehat{\beta}_\tau - Q_\tau(y))}{\min \sum_i \rho_\tau (y_i - Q_\tau(y))} = 1 - \frac{\min \sum_i (y_i - x_i'\widehat{\beta})}{\min \sum_i \rho_\tau (y_i - Q_\tau(y))} \tag{2.30}$$

where $Q_\tau(y)$ denotes the unconditional τ-th quantile of y. Like R^2, the value of $R^1(\tau)$ lies between 0 and 1. Unlike R^2 which is a global measure of goodness of fit, $R^1(\tau)$ measures the relative success of the corresponding quantile regression model and can thus be interpreted as a local goodness of fit value for a particular quantile.

To conclude our general comments, we briefly want to address the sometimes encountered faulty notion that something like quantile regression could be achieved by simply segmenting y into subsets according to its unconditional distribution and then doing least squares fitting on these subsets. Hallock, Madalozzo, and Reck (2003) provided a nice example showing the "disastrous" results of this truncation on the dependent variable (due to sample selection bias elucidated by Heckman (1979)). Even though the concrete fit of a conditional quantile is determined by only K points (see below), the decision of *which* K points are chosen depends on the entire sample for any quantile. In contrast, segmenting the sample into subsets defined according to the regressors is of course a valid option. Such local fitting underlies all non-parametric quantile regression approaches, see section 2.4.

---

[13]Koenker and Machado's (1999) formulation is a bit more general, but we wanted to stress the analogy to the classical version of equation (2.29).

2.2.2 Equivariance, robustness, efficiency, and interpretation

Already in their original paper, Koenker and Bassett (1978) showed the following basic equivariance properties of the estimated quantile regression coefficients:

$$\widehat{\beta}_\tau(\lambda y, X) = \lambda \widehat{\beta}_\tau(y, X) \qquad\qquad \lambda \in [0, \infty) \qquad\qquad (2.31)$$

$$\widehat{\beta}_\tau(-\lambda y, X) = \lambda \widehat{\beta}_{1-\tau}(y, X) \qquad\qquad \lambda \in [0, \infty) \qquad\qquad (2.32)$$

$$\widehat{\beta}_\tau(y + X\gamma, X) = \widehat{\beta}_\tau(y, X) + \gamma \qquad\qquad \gamma \in \mathbb{R}^k \qquad\qquad (2.33)$$

$$\widehat{\beta}_\tau(y, XA) = A^{-1} \widehat{\beta}_\tau(y, X) \qquad\qquad A \text{ nonsingular} \qquad\qquad (2.34)$$

Equations (2.31) and (2.32) state that $\widehat{\beta}_\tau$ is scale equivariant. That is, if the regressand y is rescaled by a factor λ, then $\widehat{\beta}_\tau$ is rescaled by the same factor. Property (2.33) is called location, shift or regression equivariance. It means that if $\widehat{\beta}_\tau$ is the solution to (y, X), then $\widehat{\beta}_\tau + \gamma$ is the solution to (y^*, X) with $y^* = y + X\gamma$. Equation (2.34) is called equivariance to reparameterization of design and means that the transformation of $\widehat{\beta}_\tau$ is given by the inverse transformation of X.

Properties (2.31) to (2.34) are shared by the least squares estimator (this is not universally true for other regression estimators). However, regression quantiles enjoy another equivariance property which is much stronger than those already discussed. From equation (2.4), it follows that for any non-decreasing function $h(\cdot)$ on $\mathbb{R}$:

$$\widehat{Q}_\tau\big(h(y)|X\big) = h\big(\widehat{Q}_\tau(y|X)\big) \qquad\qquad (2.35)$$

In words, the conditional quantiles are equivariant to monotone transformations of the response variable. Of course, unless $h(\cdot)$ is affine, the conditional mean does not share this property:

$$E\big(h(y)|X\big) \neq h\big(E(y|X)\big) \qquad\qquad (2.36)$$

Equation (2.35) can be very useful under certain conditions. If we have for example

built a model for the logarithm of the regressand[14], we are perfectly justified in interpreting $\exp(x'\widehat{\beta}_\tau)$ as an appropriate estimate of the conditional τ-th quantile of y given X, while this interpretation is difficult to be justified formally for the conditional mean (see Koenker and Portnoy (1999)). Property (2.35) is also valuable for the analysis of a censored response variable, see section 2.4.

Another important property of regression quantiles is their robustness against outliers of the regressand.[15] This means that having fit a conditional quantile hyperplane, any observation above the plane can be made arbitrarily large (up to $+\infty$) and any observation below the plane can be made arbitrarily small (up to $-\infty$) *without altering* the fitted solution. This characteristic of quantile regression is also useful for the analysis of censored response variables (see section 2.4) and can be stated formally by (compare Koenker and Portnoy (1999))

$$\widehat{\beta}_\tau\Big(y, X\Big) = \widehat{\beta}_\tau\Big(X\widehat{\beta}_\tau(y, X) + D\big(y - X\widehat{\beta}_\tau(y, X)\big), X\Big) \qquad (2.37)$$

where D is a diagonal matrix with non-negative elements d_i. In contrast, the quantile regression approach is not robust against contamination of the conditioning covariates. In section 2.4, we will briefly describe a proposal by Rousseeuw and Hubert (1999) to "robustify" quantile regression also against outlying values of the regressors.

Our next point in this subsection concerns efficiency considerations. It is well known that for a normally distributed random variable, the sample median is "worse" than the sample mean in the sense that its (asymptotic) variance is about 50% larger (see e.g. Koenker (2000)). However, for a wide range of non-Gaussian distributions, this proportion is reversed with in some cases disastrous results for the mean. Koenker and Bassett (1978) extended the concept to a regression setting by stating that the conditional median is more efficient than the least squares estimator for any distribution for which the median is more efficient than the mean. So, they concluded that "it seems

---

[14]Of course, we could also have applied the more general Box-Cox-transformation $h(y) = (y^\lambda - 1)/\lambda$

[15]Interestingly, this feature was the main focus of Koenker and Bassett's (1978) original paper.

reasonable to pay a small premium in the form of sacrificed efficiency at the Gaussian distribution, in order to achieve a substantial improvement over least squares in the event of a non-Gaussian situation." Furthermore, as already noted, the use of an appropriate weighting procedure can additionally improve the efficiency of quantile regression estimators, see Newey and Powell (1990) and section 2.4.

It is usual in least squares regression to interpret the regression coefficients β_k as (ceteris paribus) partial derivatives of the expected value of y:

$$\beta_k = \frac{\partial E(y|X)}{\partial x_k} \tag{2.38}$$

Of course, if there is more than one coefficient associated with a particular covariate (e.g. the regressor itself plus its squared value), the partial derivative consists of an appropriate combination of the according coefficients.

The interpretation of the quantile regression model is analogous to (2.38), now the coefficient $\beta_{\tau,k}$ answers the question of "how does the τ-th conditional quantile of y react to a (ceteris paribus) change of x_k":

$$\beta_{\tau,k} = \frac{\partial Q_\tau(y|X)}{\partial x_k} \tag{2.39}$$

Some caution is required if we are interested in the effect on a single observation (e.g. an individual person). If the x_k of this subject changes, of course also the conditional quantile at which the subject lies, can (and probably will) change. One possible remedy for this identification problem might be the use of longitudinal data to explore in more detail the dynamics of response.

As we have already noted, the quantile regression method is invariant to any monotone transformation of the regressand, so if we have for example estimated a model for the logarithm of the response variable ($Q_\tau(\log y) = x'\beta_\tau$), we can trouble-free write

$$\frac{\partial Q_\tau(y|X)}{\partial x_k} = e^{x'\beta_\tau} \beta_{\tau,k} \tag{2.40}$$

Table 2.3: Simple regression example. The table shows the results of least squares and several quantile regression estimates of model (2.41). The estimated standard errors given in brackets have been calculated by bootstrapping with 1000 replications, see next section. The goodness-of-fit measures were introduced in equations (2.29) and (2.30).

y	LS	q1	q10	q25	q50	q75	q90	q99
constant	3.937	2.971	3.331	3.579	3.990	4.232	4.524	5.051
	(0.063)	(0.266)	(0.066)	(0.056)	(0.061)	(0.060)	(0.063)	(0.106)
coef of x	2.158	-0.250	0.896	1.609	2.061	2.855	3.468	4.714
	(0.109)	(0.518)	(0.177)	(0.130)	(0.131)	(0.161)	(0.150)	(0.364)
R^2, R^1	.2818	.0028	.0369	.1038	.1737	.2273	.2893	.3413

In contrast, an analogue interpretation is problematic in least squares regression models. However, this does not hinder many practitioners from regardlessly (and possibly wrongly) applying it.

2.2.3 Illustration

2.2.3.1 A simple regression example

Having presented the basic features and some important properties of quantile regression, we now want to further elucidate the theoretical results with the help of some practical examples. We start with a very simple (bivariate) model by creating a regressor x which consists of 1000 independently and uniformly distributed observations on the interval $(0, 1)$. The response variable y is subsequently generated by

$$y_i = 4 + 2x_i + u_i \quad \text{with} \quad u_i \sim \mathcal{N}\left(0, (x_i + 0.5)^2\right) \tag{2.41}$$

It can easily be seen that the classical assumption of independence between the error term and the regressor is clearly violated. Table 2.3 (second column) presents the results of a least squares regression of y on x. The estimated constant and the coefficient of x are nicely in line with the imposed model parameters, so at first sight, there is no indication that the least squares result hides any information.
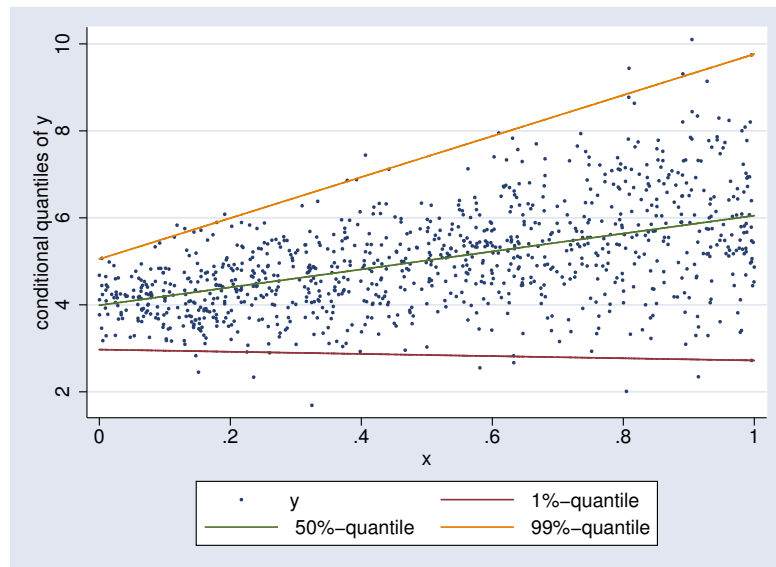
Figure 2.2: Conditional quantiles. The figure pictures every observation of model (2.41) as a blue dot. Furthermore, three estimated conditional quantile regression hyperplanes (which are, of course, in this simple regression model only straight lines) are superimposed.

However, table 2.3 also contains the outcomes from several quantile regression estimates for some selected values of τ ($\tau \in \{0.01, 0.1, 0.25, 0.5, 0.75, 0.9, 0.99\}$). While the conditional median (q50) roughly coincides with the mean regression, both the constant and the coefficient of x are considerably different for other conditional quantiles. Of course, recalling the specification from (2.41), this should be not too surprising. Figure 2.2 tries to clarify the underlying intuition. Every observation is plotted as a single dot, which clearly reveals the increasing conditional variance of y for increasing values of x. The three estimated quantile regression lines perfectly mirror the heteroscedastic structure of the error term.[16]

Instead of only analyzing some selected conditional quantiles, it is also possible to consider the whole range between 0 and 1. For practical reasons, we confine ourselves to 99 values ($\tau \in \{0.01, \ldots, 0.99\}$). Figure 2.3 plots the constant and the coefficient of x for these 99 values of τ. It can be seen that in our example, both the constant and the

---

[16]The interested reader might try to verify equations (2.27) and (2.28) which in this case say that roughly 10 points should lie below the conditional 1%-quantile line and roughly 10 points above the 99%-quantile line.
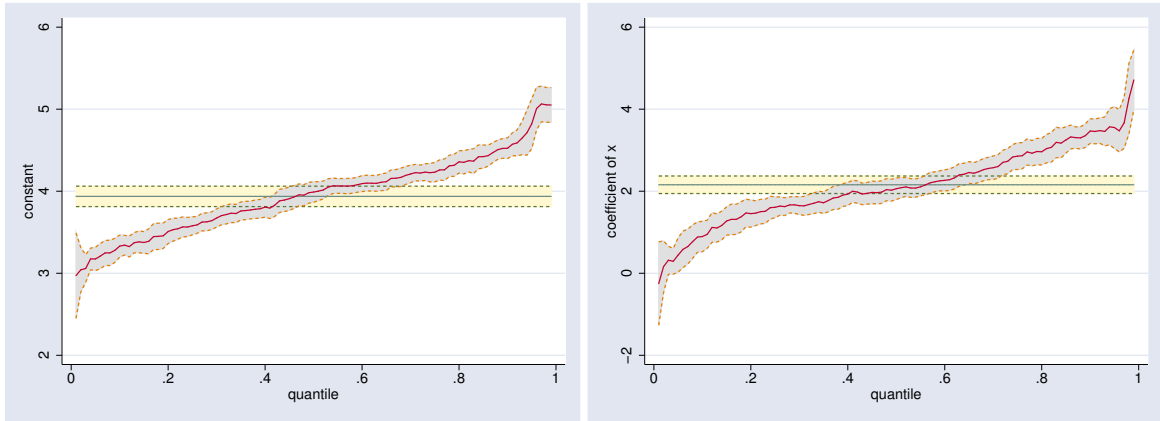
Figure 2.3: Quantile regression coefficients. The figure presents the constant and the coefficient of x from model (2.41) for 99 different quantiles. The respective values are connected as a red solid line along with an estimated 95%-confidence band shaded in gray (see next section). The least squares value is included as a horizontal solid blue line.

coefficient of x have higher values at higher quantiles. This will be different in other applications, see chapters three, four, and five. Figure 2.3 also contains the estimated least squares coefficients to demonstrate the additional information obtained by using the quantile regression approach.

The application of the quantile regression procedure paves the way to another and, as we see it, very meaningful analytical instrument. Assume that we are interested in the conditional distribution of y for a *specific* value of the regressor x. The implementation is easy and straightforward: First, we calculate 99 quantiles of y conditional on the desired value x^*:

$$\widehat{Q}_\tau(y|x^*) = \widehat{\alpha}_\tau + \widehat{\beta}_\tau x^* \tag{2.42}$$

These 99 values constitute a rough estimation of the empirical quantile function (and the empirical cumulative distribution function as its inverse). The resulting (discrete) empirical probability density function thus consists of 99 spikes of equal height. So we can easily apply a kernel density estimation on these 99 spikes to get an approximation of the estimated density of y conditional on x^*.[17] In figure 2.4, we present the result for four different values of x^*. As in our later applications, we chose the unconditional $2\%-$,

---

[17]Consider that we have n observations $Y_1, Y_2, \ldots, Y_n$. Then a so-called kernel density estimate is defined
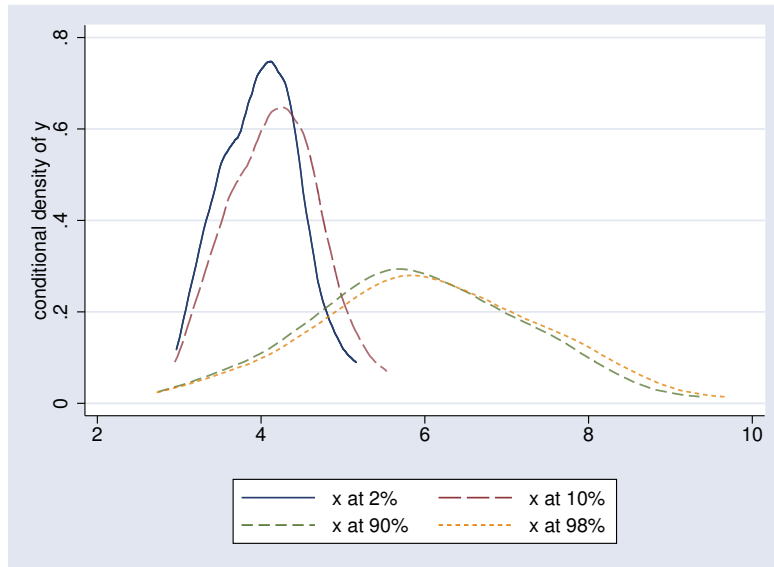
Figure 2.4: Conditional densities. The figure shows the estimated density of y from model (2.41) conditional on four different values of x (to be precise: the unconditional 2%- (solid blue line), 10%- (long-dashed red line), 90%- (dashed green line) and 98%-quantile (short-dashed orange line) of the regressor).

10%−, 90%− and 98%-quantile of x^* (which are of course just 0.02, 0.1, 0.9, and 0.98 in this example) to examine the different impacts on y. Figure 2.4 nicely shows that not only the conditional location but also the entire conditional shape of the distribution of y behaves differently for different values of x^*.

Of course, one could also think of calculating not only 99 but *all* different conditional quantiles (which are, as we have already seen, of order $\mathcal{O}(n \log n)$). In this case, the different heights of the spikes of the empirical probability density function would have to be accordingly included into the kernel density estimation process.

In our opinion, the specified method provides a simple, yet powerful tool for the post-regressional analysis of any model. Surprisingly, very little can be found on our proposal in the literature (see Koenker (2001), who presents a similar example, as an exception).

as

$$\widehat{f}(y) = (nh)^{-1} \sum_i K\left(\frac{y - Y_i}{h}\right) \tag{2.43}$$

One has to choose the kernel function K and the bandwidth parameter h, where usually the latter is more important, see e.g. Pagan and Ullah (1999). We chose the Epanechnikov kernel and the "optimal" bandwidth minimizing the mean integrated square error, see e.g. Silverman (1986).

Even worse, some of the few papers with related approaches do not adequately deal with the multiple regression setting, which will be clarified in the following example.

2.2.3.2 A multiple regression example

Our second example extends the concept to a model with two regressors. We start by generating 1000 independently and uniformly distributed observations on the interval $(20, 80)$. The values are stored as our first regressor x_1 (to put some life into it, x_1 could be seen as the age of a person). Subsequently, we create a second regressor x_2 (which could stand for the income of the respective person) by

$$x_2 = 100x_1 + \varepsilon \quad \text{with} \quad \varepsilon \sim \mathcal{N}\left(0, 500^2\right) \tag{2.44}$$

It is evident (and intended) that the two regressors are highly collinear (in our sample, the correlation coefficient was nearly 0.96). Finally, we generate our response variable y (which could for example be the person's expenses for traveling) by

$$y = 1000 - 10x_1 + 0.2x_2 + \nu \quad \text{with} \quad \nu \sim \mathcal{N}\left(0, (200 - x_1)^2\right) \tag{2.45}$$

Table 2.4 presents the outcomes of a least squares regression of model (2.45) along with several quantile regression estimates. It can be seen that the least squares result is close to the imposed coefficients. While the quantile regression coefficients of x_1 vary significantly, those of x_2 have rather the same value for all quantiles. This is perfectly in line with our specification, as the (variance of the) error term ν depends on x_1 but not on x_2. Figure 2.5 plots for three examples of τ the estimated conditional values of y against x_1 and x_2, respectively. As we now have two regressors, the interconnection of the points is of course no longer a simple straight line as in figure 2.2. Figure 2.6 visualizes the quantile regression coefficients of x_1 and x_2 for the whole range of τ. It can be seen that also in the multiple regression model, the use of quantile regression enables us to obtain a more complete picture of the underlying relationships between

Table 2.4: Multiple regression example. The table shows the results of least squares and several quantile regression estimates of model (2.45). As in table 2.3, the estimated standard errors given in brackets have been calculated by bootstrapping.

y	LS	q1	q10	q25	q50	q75	q90	q99
constant	1002.4	566.6	745.4	868.0	1013.3	1135.4	1265.0	1460.4
	(10.31)	(41.40)	(20.96)	(22.26)	(12.04)	(14.41)	(18.95)	(44.36)
coef of x_1	-10.362	-5.370	-7.015	-8.974	-11.116	-11.736	-14.246	-14.347
	(0.698)	(1.758)	(0.999)	(1.188)	(0.596)	(0.854)	(1.029)	(1.716)
coef of x_2	0.203	0.195	0.195	0.203	0.210	0.203	0.215	0.198
	(0.007)	(0.020)	(0.011)	(0.012)	(0.006)	(0.008)	(0.010)	(0.015)
R^2, R^1	.7778	.5796	.5659	.5641	.5594	.5394	.5136	.5201
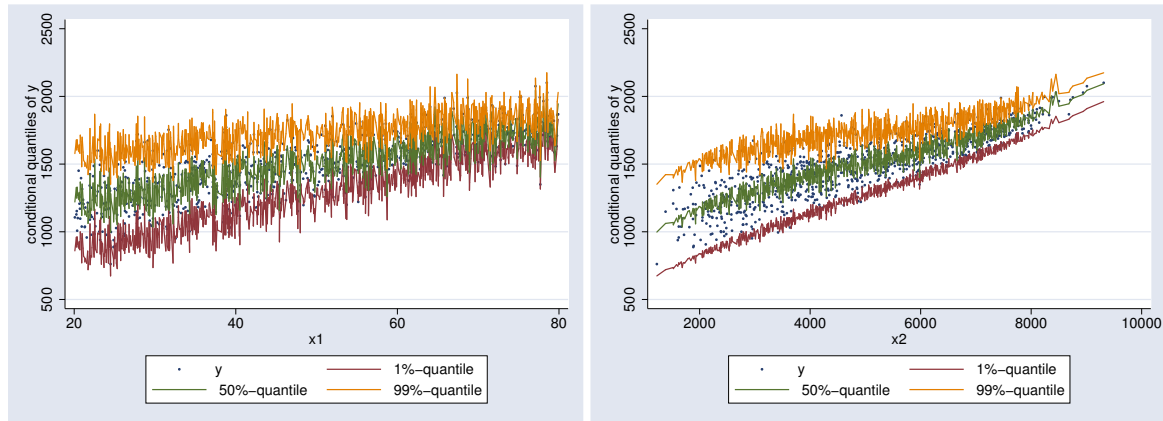
Figure 2.5: Conditional quantiles. The figure pictures for every observation of model (2.44) three conditional estimated quantiles of y against x_1 (left side) and x_2 (right side).

the regressors and the response variable.

Obviously, also in the multiple regression model, the computation of conditional densities of y would be of interest for the analysis of a given model. However, the situation is slightly more complicated than in the simple regression setting. As we know, in our example the estimated conditional quantiles of y are given by:

$$\widehat{Q}_\tau(y|x_1, x_2) = \widehat{\alpha}_\tau + \widehat{\beta}_\tau x_1 + \widehat{\gamma}_\tau x_2 \tag{2.46}$$

So if we are interested in the conditional distribution of y for a specific value of x_1 (say
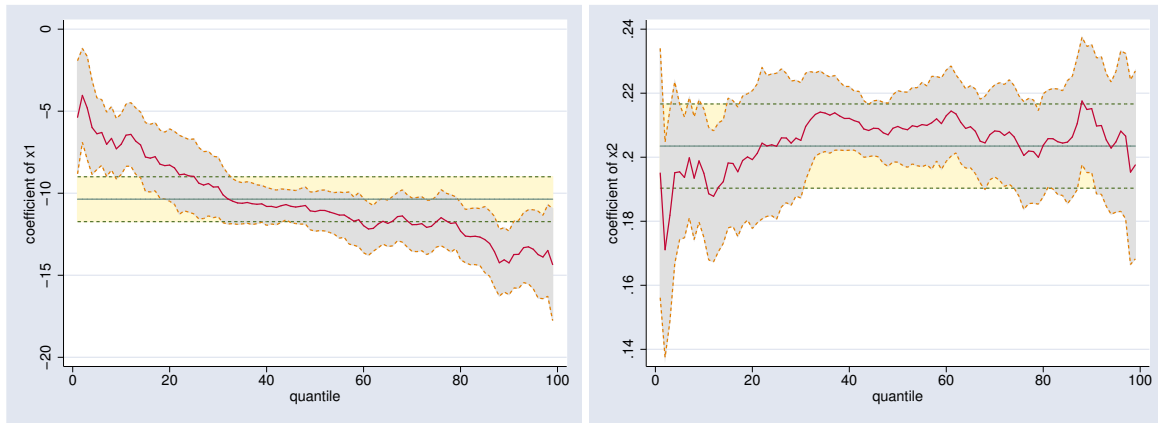
Figure 2.6: Quantile regression coefficients. The figure presents the coefficients of x_1 and x_2 from model (2.44) for 99 different quantiles along with the LS result.
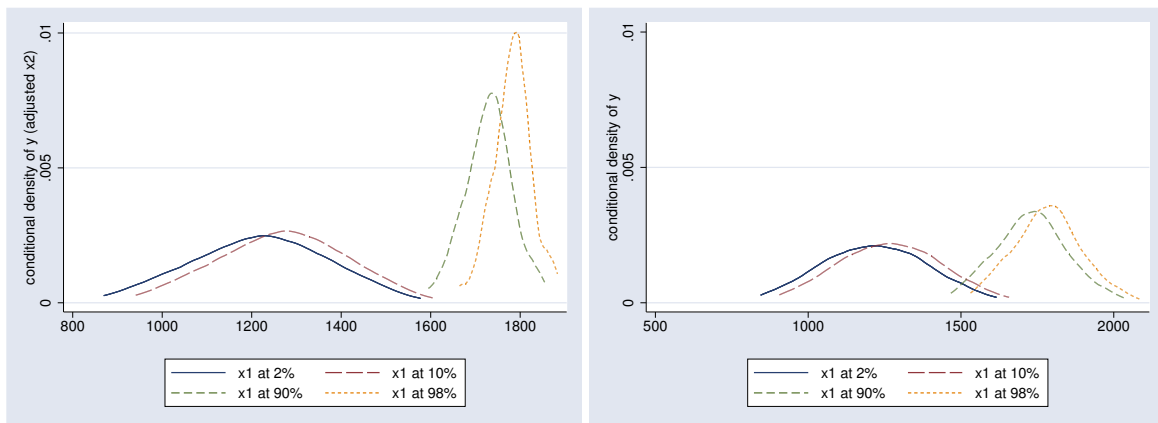
Figure 2.7: Densities of y conditional on x_1. The left side is based on an auxiliary regression of x_2 on x_1 for model (2.44). The right side is based on a simple regression of y on x_1.

x_1^*), we also need some number of x_2 to plug into equation (2.46). But which value to choose? The answer is given by the so-called Frisch-Waugh-theorem (see e.g. Greene (2002)). In our example, the theorem tells us that we can first conduct an auxiliary regression of x_2 on x_1 and subsequently plug in the obtained values into equation (2.46). The result is given on the left side of figure 2.7. To verify our approach, we also included the outcome of a simple regression of y on x_1 on the right side of figure 2.7. It can clearly be seen that both graphs are in line, while the smaller variance on the left side nicely reflects the additional information given in the full model compared to the simple regression.
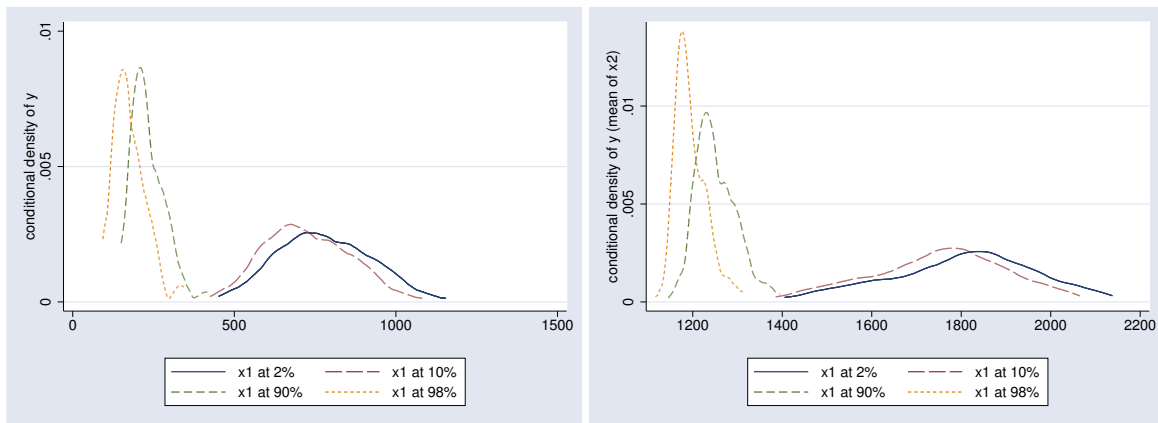
Figure 2.8: Two invalid implementations of conditional densities. The left side is calculated only from the constant and x_1, the right side includes the mean value of x_2.

To complete our discussion, figure 2.8 contains two *erroneous* implementations of the density estimation. The left side completely ignores x_2 and evaluates the densities of y only from the constant and x_1. Obviously, this is no meaningful specification and leads to completely false results. More interesting is the right side of figure 2.8 where we included the mean of x_2 in the estimation procedure of the density of y conditional on x_1. It can clearly be seen that also in this case, the procedure is far away from yielding correct outcomes.

However (and surprisingly to us), the latter approach can be found frequently (with respect to few papers dealing with the subject at all) in the literature. For example, Mata and Machado (1996) presented an estimation of the empirical quantile function (which could be used to calculate the conditional density of y as presented above). Unfortunately, they did not account for any possible collinearity among the regressors.

2.3   Asymptotics and Inference

According to Koenker and Hallock (2001), "it is a basic principle of sound econometrics that every serious estimate deserves a reliable assessment of precision". As the finite sample distribution of the estimated quantile regression coefficients $\widehat{\beta}_\tau$ is not very

tractable in most applications[18], we resort to asymptotic theory for inferential state-ments. We will see that the (estimation of the) already introduced sparsity function plays a crucial role for this purpose. We also discuss how resampling methods can be used to obtain information on the accuracy of the quantile regression process. Some test procedures are also addressed. Finally, we briefly outline the results of several Monte Carlo experiments comparing some of the different approaches.

2.3.1 Asymptotic normality

2.3.1.1 Sample quantiles

Before we extend our analysis to regression quantiles, we reconsider the ordinary uni-variate sample quantiles from equation (2.13):

$$\widehat{Q}_Y(\tau) = \underset{\xi_\tau \in \mathbb{R}}{\operatorname{argmin}} \sum_i \rho_\tau(Y_i - \xi_\tau) \tag{2.48}$$

The so-called uniform law of large numbers (also known as Glivenko-Cantelli-theo-rem[19]) tells us that the empirical distribution function of a one dimensional random variable converges uniformly to the true distribution function in probability (see e.g. Vapnik (1998)). This provides us with the desired consistency of the sample quantiles.

What size has the rate of convergence? Koenker and Portnoy (1999) show that if the continuous density of Y is bounded away from 0 and ∞ near ξ_τ[20], then

$$\sqrt{n}\left(\widehat{Q}_Y(\tau) - \xi_\tau\right) \longrightarrow \mathcal{N}\left(0, \omega^2\right) \quad \text{with} \quad \omega^2 = \frac{\tau(1-\tau)}{f_Y^2(\xi_\tau)} \tag{2.49}$$

---

[18]Koenker and Portnoy (1999) show that with i.i.d. errors $u_{i,\tau}$ (having strictly positive density f at $F^{-1}(\tau)$), the density of $\widehat{\beta}_\tau$ takes the rather unwieldy form

$$g(b) = \sum_{h \in \mathcal{H}} P\{\xi_h(b) \in [\tau - 1, \tau]^p\}|X(h)| \prod_{i \in h} f(x_i'(b - \beta_\tau) + F^{-1}(\tau)) \tag{2.47}$$

where $\xi_h(b) = \sum_i \psi_\tau(y_i - x_i b)x_i' X(h)^{-1}$ and $\psi_\tau(u) = \tau - I(u < 0)$

[19]The name dignifies the work of Glivenko (1933) and Cantelli (1933).
[20]Smirnov (1952) discussed cases in which this assumption fails. See also Knight (2000).

It can be seen that the numerator of ω^2 seems to make $\widehat{Q}_Y(\tau)$ more precise in the tails, but this effect is typically dominated by the density of Y (which is usually small if τ is near 0 or 1) in the denominator. If we remember the sparsity function from equation (2.8), we see that the asymptotic variance of the estimated sample quantile can also be expressed as $\omega^2 = \tau(1-\tau)s_Y^2(\tau)$. So it is evident that the *sparsity* of the data at a specific quantile determines the preciseness of the estimated value.

If F_Y is actually flat in a neighborhood of ξ_τ, we can only say that the sum of the probabilities that $\widehat{Q}_Y(\tau)$ falls near the lower *or* upper bound of the flat interval tends to 1 as $n \to \infty$. Ellis (1998) argued against this "instability", but the comment of Portnoy and Mizera in the same article rebutted his line of argumentation.

Equation (2.49) can be extended to the limiting form of the joint distribution of several quantiles. If ζ is a vector of m sample quantiles $\left(\zeta = (\xi_{\tau_1}, \ldots, \xi_{\tau_m})\right)$ and $\widehat{\zeta}$ denotes its estimate, it can be shown that (see again Koenker and Portnoy (1999))

$$\sqrt{n}\left(\widehat{\zeta} - \zeta\right) \longrightarrow \mathcal{N}(0, \Omega) \quad \text{with} \quad \Omega = \left(\omega_{ij}\right) = \frac{\min(\tau_i, \tau_j) - \tau_i \tau_j}{f(F^{-1}(\tau_i))f(F^{-1}(\tau_j))} \tag{2.50}$$

2.3.1.2   Linear regression quantiles

If we move on to the asymptotic theory of estimating conditional linear quantile functions, the situation gets a bit more complicated. Again, we start with considering the question under which circumstances the estimators converge in probability to their true value. El Bantli and Hallin (1999) provides us with necessary and sufficient conditions for the desired consistency of the estimated regression coefficients. First, they make the following two assumptions:[21]

1. There exists a vector $d > 0$ such that $\lim_{n \to \infty} \inf \inf_{||u||=1} n^{-1} \sum_i I(|x_i'u| < d) = 0$

2. There exists a matrix $D > 0$ such that $\lim_{n \to \infty} \sup \sup_{||u||=1} n^{-1} \sum_i (x_i'u)^2 \leq D$

---

[21]The first condition ensures that the x_i's are not concentrated on a subspace and is needed for identifiability, the second one controls the rate of growth of the x_i's.

Subsequently, they show that for fixed $\varepsilon > 0$,

$$\lim_{n \to \infty} \sqrt{n} \left(n^{-1} \sum_i F_i(x_i'\beta_\tau - \varepsilon) - \tau \right) \to \infty \quad \text{and} \quad \lim_{n \to \infty} \sqrt{n} \left(\tau - n^{-1} \sum_i F_i(x_i'\beta_\tau + \varepsilon) \right) \to \infty$$

are necessary and sufficient for $\widehat{\beta}_\tau \to \beta_\tau$. Other formulations of consistency with slightly different imposed conditions can be found, inter alia, in Oberhofer (1982), Koenker and Bassett (1984), Bassett and Koenker (1986), Zhao, Rao, and Chen (1993), Chen, Wu, and Zhao (1995), and Mizera and Wellner (1998).

In analogy to equation (2.49), it can be shown that (under some mild regularity conditions) the rate of convergence of the estimated regression coefficients is also $\mathcal{O}(1/\sqrt{n})$:

$$\sqrt{n} \left(\widehat{\beta}_\tau - \beta_\tau \right) \longrightarrow \mathcal{N}\left(0, \tau(1 - \tau) H_\tau^{-1} J H_\tau^{-1} \right) = \mathcal{N}\left(0, \Lambda_\tau \right) \tag{2.51}$$

$$\text{with} \quad J = \lim_{n \to \infty} n^{-1} \sum_i x_i x_i'$$

$$\text{and} \quad H_\tau = \lim_{n \to \infty} n^{-1} \sum_i x_i x_i' f_i(Q_\tau(y_i|x_i))$$

A detailed proof of (2.51) can be found in Kim and White (2003). Their argument uses the so-called stochastic equicontinuity of the gradient of the objective function. This means that the discontinuous first order conditions are made asymptotically stochastically uniformly continuous to allow for taking the usual Taylor expansion to obtain the asymptotic distribution. See Huber (1967), Bickel (1975), and Jureckova (1977) for important contributions on stochastic equicontinuity and monotonicity of the gradient. The asymptotic results of quantile regression estimators under i.i.d. errors were derived by Ruppert and Carroll (1980).

Other authors based their argument on the convexity of the limiting objective function, see e.g. Pollard (1991), Niemiro (1992), Hjort and Pollard (1993), Bai (1995), Geyer (1996), and Knight (1998, 1999). Further information on the asymptotic properties of least absolute deviation and quantile regression under various conditions are provided in, among others, Bassett and Koenker (1978), Koenker and Portnoy (1987), Phillips

(1991), Portnoy (1991a), Andrews (1994), Welsh, Carroll, and Ruppert (1994), Koenker (1994), Herce (1996), Zhou and Portnoy (1996), Fitzenberger (1997b), Zhou and Portnoy (1998), Van Der Vaart (1998), and Zhao (2001).

If the errors are assumed to be i.i.d., equation (2.51) simplifies to

$$\sqrt{n}\left(\widehat{\beta}_\tau - \beta_\tau\right) \longrightarrow \mathcal{N}\left(0, \frac{\tau(1-\tau)}{f^2(F^{-1}(\tau))}J^{-1}\right) \quad \text{with} \quad J = \lim_{n\to\infty} n^{-1}\sum_i x_i x_i' \qquad (2.52)$$

We see that in this case all density values f_i are identical, so the "sandwich" variance from equation (2.51) collapses to a much simpler expression. One could argue that the assumption of i.i.d. errors is very restrictive and often violated in empirical applications. Furthermore, if the assumption is true, the information gain from using quantile regression is limited, as all conditional quantile planes are parallel in this case.[22] However, the analysis of the special case has its justification for at least two reasons: (i) as we will see in the next subsection, the estimation of the variance in (2.51) is not necessarily straightforward and some of the proposed methods are only feasible in the i.i.d. case; (ii) equation (2.52) can be regarded as a kind of benchmark model and thus be utilized for the construction of tests, which is also shown below.

In analogy to equation (2.50), we can also specify the joint asymptotic distribution of several regression coefficient vectors (calculated at different quantiles). Let ζ be a vector of m K-variate quantile regression estimators ($\zeta = (\beta_{\tau_1}', \ldots, \beta_{\tau_m}')'$) and $\widehat{\zeta} = (\widehat{\beta}_{\tau_1}', \ldots, \widehat{\beta}_{\tau_m}')'$ its estimate. Then the joint asymptotic distribution of these m estimated regression coefficient vectors is given as (see Koenker and Portnoy (1999))

$$\sqrt{n}\left(\widehat{\zeta} - \zeta\right) \longrightarrow \mathcal{N}\left(0, \Omega\right) \qquad (2.53)$$

$$\text{with} \quad \Omega = \left(\omega_{ij}\right) = \left(\min(\tau_i, \tau_j) - \tau_i\tau_j\right) H_{\tau_1}^{-1} J H_{\tau_2}^{-1}$$

$$\text{and} \quad J = \lim_{n\to\infty} n^{-1}\sum_i x_i x_i' \quad \text{and} \quad H_\tau = \lim_{n\to\infty} n^{-1}\sum_i x_i x_i' f_i(Q_\tau(y_i|x_i))$$

[22]To quote Chamberlain (1994): "This is hardly an attractive foundation for inference, given our focus on how the slopes differ for different quantiles."

2.3.2 Sparsity estimation

2.3.2.1 Independently and identically distributed errors

In the last subsection, we have seen in (2.52) that under the assumption of i.i.d. errors, the asymptotic normality of an estimated regression quantile coefficient is given as

$$\sqrt{n}\left(\widehat{\beta}_\tau - \beta_\tau\right) \longrightarrow \mathcal{N}\left(0, \frac{\tau(1-\tau)}{f^2(F^{-1}(\tau))}J^{-1}\right) \quad \text{with} \quad J = \lim_{n\to\infty} n^{-1}\sum_i x_i x_i' \qquad (2.54)$$

It is easy to see that for any inferential assertion, we somehow have to estimate the (square of the) reciprocal of the density $f(F^{-1}(\tau))$, already introduced as sparsity $s(\tau)$.[23] As the sparsity function is the derivative of the quantile function, it is natural to just use a simple difference quotient of the empirical quantile function to estimate the sparsity[24]

$$\widehat{s}(\tau) = \frac{\widehat{F}^{-1}(\tau + h_n) - \widehat{F}^{-1}(\tau - h_n)}{2h_n} \qquad (2.55)$$

where $\widehat{F}^{-1}$ is an estimate of the quantile function and h_n is a bandwidth parameter which tends to zero as $n \to \infty$. This procedure was originally proposed by Siddiqui (1960) for constructing confidence intervals for univariate sample quantiles (see also Koenker (1994)). Obviously, the next question arising from (2.55) is how to optimally choose the bandwidth parameter h_n. Bofinger (1975) showed that

$$h_n = n^{-1/5}\left(\frac{4.5(s(\tau))^2}{(s''(\tau))^2}\right)^{1/5} \qquad (2.56)$$

minimizes the mean squared error (MSE) under mild regularity conditions on F. Of course, if we knew $s(\tau)$ and $s''(\tau)$ we wouldn't need h_n, but as $s(\tau)/s''(\tau)$ is not very sensitive to F, we can choose h_n for some typical distributional shape (see Sheather

---

[23]The estimation of the sparsity can to some extent be compared with the estimation of the standard deviation σ in the least squares case.

[24]As the centered difference formula in (2.55) has $\mathcal{O}(h^2)$ truncation error, it is preferred over a simple forward (or backward) difference formula which has $\mathcal{O}(h)$ truncation error, see Chen (2001).
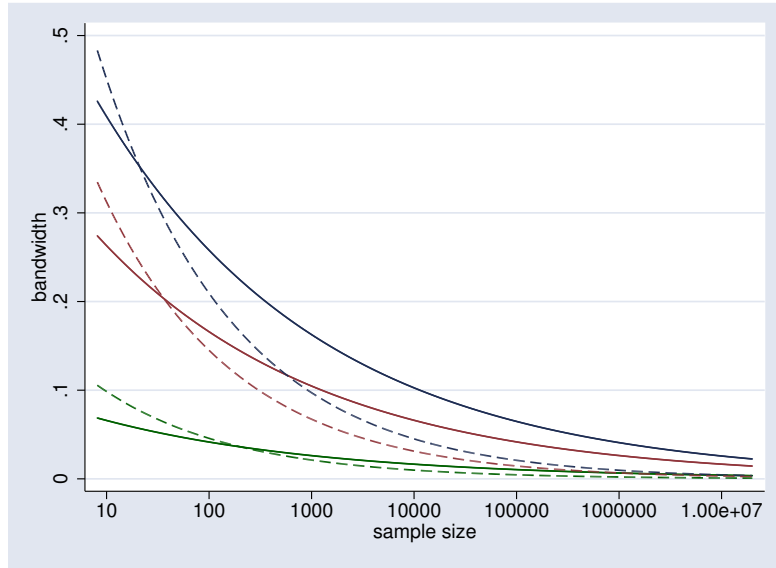
Figure 2.9: Bandwidth parameter h_n. The three solid lines show the Bofinger bandwidth from equation (2.58) depending on n. From top to bottom, the blue line represents $\tau = 0.5$, the red one shows $\tau = 0.75$ and green indicates $\tau = 0.95$. The dashed lines show the Hall and Sheather bandwidth from equation (2.60) for the same values of τ and $\alpha = 0.05$.

and Maritz (1983)). In general,

$$\frac{s(\tau)}{s''(\tau)} = \frac{f^2}{2(f'/f)^2 + \left((f'/f)^2 - f''/f\right)} \tag{2.57}$$

so if we plug in the Gaussian distribution ($f = \phi$), we see that $(f'/f)(F^{-1}(\tau)) = \Phi^{-1}(\tau)$ and thus Bofinger's bandwidth parameter is given as

$$h_n = n^{-1/5} \left(\frac{4.5(\phi(\Phi^{-1}(\tau)))^4}{(2(\Phi^{-1}(\tau))^2 + 1)^2} \right)^{1/5} \tag{2.58}$$

The three solid lines in figure 2.9 show the resulting values of (2.58) for $\tau = 0.5$, $\tau = 0.75$ and $\tau = 0.95$.[25] As the assumed Gaussian distribution is symmetric, it is clear that the h_n's are identical at τ and $1 - \tau$. Figure 2.9 indicates that the bandwidth gets smaller for larger sample sizes and values of τ further away from 0.5.

---

[25]In figure 1 of Koenker (1994), which shows the same than figure 2.9, the bandwidth curves have some odd kinks at small n. We believe that this must be a computational inaccuracy since neither (2.58) nor our results suggest such kinks.
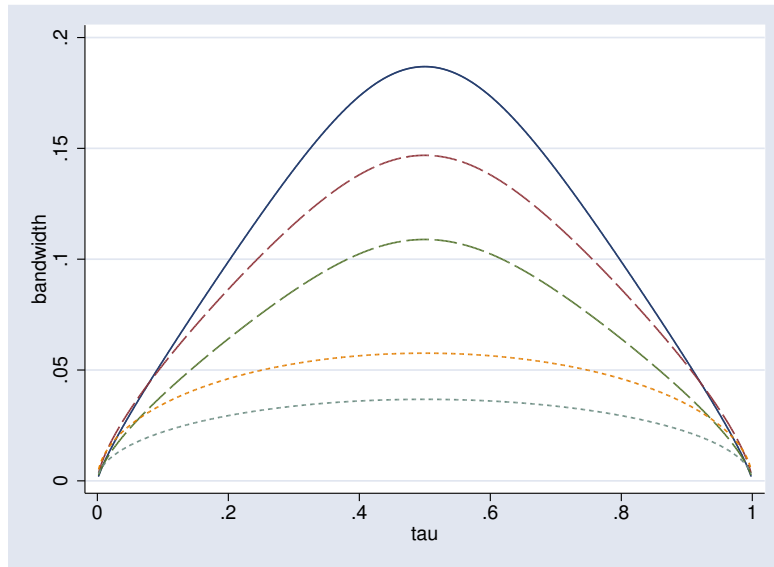
Figure 2.10: Bandwidth parameter h_n. The blue solid line shows the Bofinger bandwidth for $n = 500$ depending on the value of τ. The red ($\alpha = 0.01$) and green ($\alpha = 0.1$) long-dashed lines present the Hall and Sheather bandwidth, the yellow ($\alpha = 0.01$) and grey ($\alpha = 0.1$) short-dashed lines indicate the Chamberlain bandwidth.

Hall and Sheather (1988) questioned Bofinger's rule and suggested instead

$$h_n = n^{-1/3} z_\alpha^{2/3} \left(\frac{1.5 s(\tau)}{s''(\tau)} \right)^{1/3} \tag{2.59}$$

where $z_\alpha = \Phi^{-1}(1 - \frac{\alpha}{2})$. The parameter α denotes the desired size of the test. Since the Hall and Sheather rule is explicitly designed for confidence interval construction, rather than simply optimizing MSE-performance for the sparsity estimate itself, Koenker (1994) argued that it seems to be more reasonable for inferential purposes. If we again plug in the Gaussian distribution, we obtain

$$h_n = n^{-1/3} z_\alpha^{2/3} \left(\frac{1.5 (\phi(\Phi^{-1}(\tau)))^2}{2(\Phi^{-1}(\tau))^2 + 1} \right)^{1/3} \tag{2.60}$$

Figure 2.9 also contains three examples of the Hall and Sheather bandwidth, each calculated for $\alpha = 0.05$. It can be seen that the values are bigger than the Bofinger sandwich for small sample sizes but (considerably) smaller for medium to big data sets.

A third and rather simple alternative to the bandwidth estimation was proposed by

Buchinsky (1991) and is derived from the asymptotics of binomially distributed quantiles (see also Chen (2001)). It is named Chamberlain bandwidth and also contains a significance level parameter α:

$$h_n = z_\alpha \sqrt{\frac{\tau(1-\tau)}{n}} \tag{2.61}$$

Figure 2.10 pictures the different bandwidth choice rules for the whole range of τ with $n = 500$. We imposed two different confidence levels for the Hall and Sheater bandwidth and the Chamberlain bandwidth, respectively.

Having presented three possibilities to determine the bandwidth parameter h_n, we now consider the question of how to calculate the estimated empirical quantile function $\widehat{Q}(\tau) = \widehat{F}^{-1}(\tau)$ in equation (2.55). A simple approach is to just take the residuals from a quantile regression fit at an arbitrarily chosen quantile (denoted with τ^*):[26]

$$\widehat{u}_i = y_i - x_i'\widehat{\beta}_{\tau^*} \qquad i = 1, \dots, n \tag{2.62}$$

Sorting the resulting values, we obtain the corresponding order statistics $\widehat{u}_{(i)} : i = 1, \dots, n$ and can estimate an empirical quantile function by

$$\widehat{F}^{-1}(\tau) = \widehat{u}_{(i)} \qquad \tau \in [\tfrac{i-1}{n}, \tfrac{i}{n}) \tag{2.63}$$

Of course, now the estimated empirical quantile function refers to u_{τ^*} instead of the regressand. But since the sparsity is defined as a function of the difference of two $\widehat{F}^{-1}$'s and we have assumed i.i.d. errors, both approaches are valid. If one prefers a piecewise linear version of (2.63), one can use

$$\widetilde{F}^{-1}(\tau) = \begin{cases} \widehat{u}_{(1)} & \text{if } \tau \in [0, \tfrac{1}{2n}) \\ \lambda\widehat{u}_{(j+1)} + (1-\lambda)\widehat{u}_{(j)} & \text{if } \tau \in [\tfrac{2j-1}{2n}, \tfrac{2j+1}{2n}) \qquad j = 1, \dots, n-1 \\ \widehat{u}_{(n)} & \text{if } \tau \in [\tfrac{2n-1}{2n}, 1] \end{cases} \tag{2.64}$$

---

[26]Unfortunately, Koenker (1994) is unclear at this point. Bassett and Koenker (1982) are more specific, they use not only the median but also (alternatively) a least squares preliminary estimation to obtain the residuals.

with $\lambda = \tau n - j + 1/2$. The fact that the K residuals equal to zero may be problematic if K/n is large relative to h_n can easily be circumvented by ignoring the zero residuals. This procedure can be compared to the usual degrees of freedom correction in least squares regression.

A perhaps more appealing approach to obtain the estimated empirical quantile function was proposed by Bassett and Koenker (1982). As we have already seen in the last subsection, the sample path of the conditional quantiles of y is non-decreasing in τ at the mean of x, so we can just use (see also Bassett and Koenker (1986))

$$\widehat{F}^{-1}(\tau) = \bar{x}'_i \widehat{\beta}_\tau \tag{2.65}$$

To give an impression of some typical sparsity values, figure 2.11 shows a very simple example. We drew 500 sample observations from a standard normal distribution, determined the empirical quantile function (e.g. by just setting $x = \iota$ in (2.65)), and applied the three bandwidth choice rules. It can be seen that the Bofinger bandwidth yields the smoothest result (our plot shows the estimates for $\tau \in [0.08, 0.92]$).

Other techniques for the sparsity estimation are also possible: Welsh (1988) used a kernel approach and gave greater weight to values with narrower bandwidth. In Koenker and Bassett (1982a), the sparsity function is estimated by twice differentiating a smoothed version of $\widehat{R}(\tau)$, which denotes the minimum value achieved by the objective function at each regression quantile. Chamberlain (1994) obtained the sparsity from an estimated confidence band constructed from appropriate order statistics.

2.3.2.2 Non-i.i.d. errors

In equation (2.51), we have seen that in a non-i.i.d. setting, the asymptotic variance of the estimated quantile regression coefficients has the slightly unpleasant property of
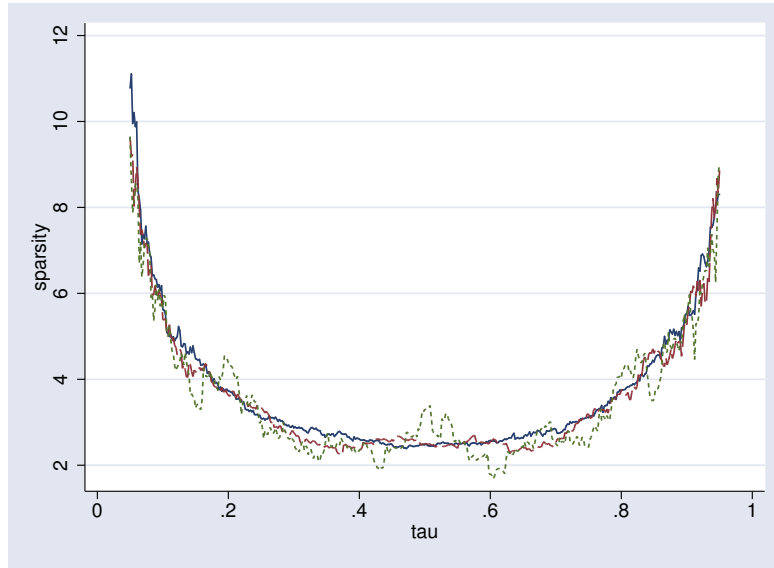
Figure 2.11: Sparsity estimation. The figure shows the estimated sparsity for a random sample of 500 standard normally distributed values with the Bofinger (solid blue line), the Hall and Sheather (long-dashed red line) and the Chamberlain (short-dashed green line) bandwidth.

different densities at each observation:

$$\sqrt{n}\left(\widehat{\beta}_\tau - \beta_\tau\right) \longrightarrow \mathcal{N}\left(0, \tau(1-\tau)H_\tau^{-1}JH_\tau^{-1}\right) = \mathcal{N}\left(0, \Lambda_\tau\right) \qquad (2.66)$$

$$\text{with} \quad J = \lim_{n\to\infty} n^{-1}\sum_i x_i x_i' \quad \text{and} \quad H_\tau = \lim_{n\to\infty} n^{-1}\sum_i x_i x_i' f_i(Q_\tau(y_i|x_i))$$

So, in this case we somehow have to estimate the entire matrix H_τ. The first proposal (see Hendricks and Koenker (1992)) is to estimate the conditional density $f_i(Q_\tau(y_i|x_i))$ at each observation by

$$\widehat{f}_i(Q_\tau(y_i|x_i)) = \frac{2h_n}{x_i'(\widehat{\beta}_{\tau+h_n} - \widehat{\beta}_{\tau-h_n})} \qquad (2.67)$$

with the same bandwidth parameter h_n (e.g. Bofinger, Hall and Sheather, or Chamberlain bandwidth) as already discussed. Subsequently, the obtained values are substituted into H_n, and we have an estimator for the non-i.i.d. model. A possible drawback of this procedure is the fact that the denominator of (2.67) is not necessarily always positive. As we have already seen, conditional quantile planes can cross, although usually only at the periphery of the design space. To correct for this "inconvenience", we simply

replace $\widehat{f}_i(Q_\tau(y_i|x_i))$ by its positive part and add a small tolerance parameter $\varepsilon > 0$ to avoid dividing by zero:

$$\widehat{f}_i^+(Q_\tau(y_i|x_i)) = \max\left(0, \frac{2h_n}{x_i'(\widehat{\beta}_{\tau+h_n} - \widehat{\beta}_{\tau-h_n}) - \varepsilon}\right) \tag{2.68}$$

Another approach to estimate H_τ was proposed by Powell (1986, 1991) and is given as (see also Buchinsky (1998b))

$$\widehat{H}_\tau = (nh_n)^{-1} \sum_i K\left(\frac{\widehat{u}_{i,\tau}}{h_n}\right) x_i x_i' \tag{2.69}$$

where $\widehat{u}_{i,\tau} = y_i - x_i'\widehat{\beta}_\tau$ are again the residuals and h_n is a bandwidth parameter with $\lim_{n\to\infty} h_n = 0$ and $\lim_{n\to\infty} \sqrt{n}h_n = \infty$. As in the last section, K denotes the kernel function. So, we again have to choose an appropriate kernel and an optimal bandwidth, where the latter selection is more important. Buchinsky (1998b) suggests to use a cross-validation method for the choice of h_n.

2.3.3 Bootstrapping

As we have seen in the last subsection, the estimation of the (asymptotic) covariance matrix can be a bit burdensome, in particular if one assumes the more realistic case of non-i.i.d. errors. So it comes to no surprise that several authors tried to circumvent the laborious estimation of the sparsity and proposed alternatives to obtain inferential information. A large part of these proposals is based on resampling methods known as *bootstrapping* (see Efron (1979) as an important initial contribution and e.g. Efron and Tibshirani (1993) for a comprehensive introduction).[27]

As in other applications, there are several possible implementations of the bootstrap procedure. The first one is called residual bootstrap and was suggested by Efron (1982)

---

[27]Under a generous perspective, some parallels can be found between quantile regression and bootstrapping: Both approaches were "invented" in the late 70's, both provided new insights unattainable by conventional methods, and both received a "career boost" in the 90's due to improved availability of computational power.

for a nonlinear median regression problem. An adaptation for general quantile regression settings can be found e.g. in Hahn (1995). The idea is to fit a quantile regression model, obtaining the residuals

$$\widehat{u}_{i,\tau} = y_i - x_i'\widehat{\beta}_\tau \qquad i = 1, \ldots, n \tag{2.70}$$

Subsequently, a bootstrap sample $u_{i,\tau}^*, \ldots, u_{n,\tau}^*$ is drawn (of course with replacement) from the estimated empirical distribution. With $y_i^* = x_i\widehat{\beta}_\tau + u_{i,\tau}^*$, a bootstrapped regression coefficient is given as

$$\widehat{\beta}_\tau^* = \operatorname*{argmin}_{\beta_\tau \in \mathbb{R}^K} \sum_i \rho_\tau(y_i^* - x_i'\beta_\tau) \tag{2.71}$$

Repeating this process B times yields $\widehat{\beta}_{\tau,1}^*, \ldots, \widehat{\beta}_{\tau,B}^*$, and we can consistently estimate the asymptotic variance of $\widehat{\beta}_\tau$ by (compare Bickel and Freedman (1981) and Freedman (1981))[28]

$$\widehat{\Lambda}_\tau = \frac{n}{B} \sum_{b=1}^B \left(\widehat{\beta}_{\tau,b}^* - \widehat{\beta}_\tau\right)\left(\widehat{\beta}_{\tau,b}^* - \widehat{\beta}_\tau\right)' \tag{2.72}$$

De Angelis, Hall, and Young (1993) showed that the error of the approximation is of order $\mathcal{O}(n^{-1/4})$ as $n \to \infty$. They also presented a smoothed version of the empirical distribution function of the residuals with improved error of only $\mathcal{O}(n^{-2/5})$. A major drawback of the residual bootstrap is the obvious fact that it is only valid under the assumption of i.i.d. errors. As we have already explained, this is a strong limitation which confines the practical utility of the method.

If the error term is independently but not necessarily identically distributed, the so-called design matrix bootstrap provides a viable alternative to the residual bootstrap. The idea is to draw (x_i^*, y_i^*)-pairs from the joint empirical distribution of the sample, of course again with replacement. Subsequently, from each (x_i^*, y_i^*)-pair, a bootstrapped regression coefficient $\widehat{\beta}_\tau^*$ is calculated, and the resulting B coefficient vectors are again

---

[28]Fitzenberger (1997b) argued that it might be preferable to replace $\widehat{\beta}_\tau$ by the bootstrap mean $\bar{\beta}_\tau^*$ in (2.72).

inserted into equation (2.72) to obtain the estimated asymptotic covariance matrix. Horowitz (1998) suggested some refinements of the design matrix bootstrap based on smoothing of the quantile regression objective function.

Having calculated a set of bootstrapped regression coefficients, the so-called percentile method provides an alternative to equation (2.72) by just sorting the $\widehat{\beta}_\tau^*$'s and defining a confidence interval for β_τ with the $B\alpha/2$-th element as lower bound and the $B(1-\alpha)/2$-th element as upper bound (see e.g. Buchinsky (1996)). Andrews and Buchinsky (2000, 2001, 2002) provide an extensive discussion on the optimal number of bootstrap repetitions. Several authors (see Bickel and Freedman (1981), Buchinsky (1994), Sakov and Bickel (2000), Abrevaya (2001)) showed that a smaller size of the bootstrap sample m than the original sample n can be advantageous under some cirumstances. Fitzenberger (1997b) analyzed the case of autocorrelated errors and proposed a moving blocks bootstrap. Parzen, Wei, and Ying (1994) proposed a resampling method based on the subgradient condition (see also Bilias, Chen, and Ying (2000)). Hahn (1997) investigated the large sample property of the bootstrapped quantile regression estimator under a Bayesian setting.

Some further applications of bootstrap methods on quantile regression problems can, among others, be found in Knight (2002), He and Hu (2002), Kocherginsky, He, and Mu (2003), and Machado and Parente (2003).

2.3.4   Testing procedures

In the last two subsections, we have seen several ways to obtain the (asymptotic) covariance matrix of the estimated regression coefficients. In this subsection, we briefly present some testing approaches based on the previous results.

2.3.4.1 Wald tests

Consider a linear quantile regression model

$$y_i = x_i'\beta_\tau + u_{i,\tau} \qquad i = 1, \ldots, n \tag{2.73}$$

with the linear hypothesis

$$H_0 : R\beta_\tau = r \tag{2.74}$$

Koenker and Bassett (1982a, 1982b) showed that the test statistic

$$W_\tau = n(R\widehat{\beta}_\tau - r)'[R\widehat{\Lambda}_\tau^{-1}R']^{-1}(R\widehat{\beta}_\tau - r) \tag{2.75}$$

is asymptotically chi-square distributed under the null hypothesis with rank(R) degrees of freedom. $\widehat{\Lambda}_\tau$ denotes the estimated variance-covariance-matrix of the estimated regression coefficient vector $\widehat{\beta}_\tau$ from equation (2.51) and can be determined by one of the already presented methods.

An even more general Wald statistic is given by

$$W = n(R\widehat{\zeta} - r)'[R\widehat{\Omega}^{-1}R']^{-1}(R\widehat{\zeta} - r) \tag{2.76}$$

where $\widehat{\Omega}^{-1}$ is the estimated asymptotic joint matrix from equation (2.53) and the null hypothesis ($H_0 : R\zeta = r$) refers to $\zeta = (\beta_{\tau_1}', \ldots, \beta_{\tau_m}')'$. This formulation accommodates a wide variety of testing situations, see Koenker and Portnoy (1999).

2.3.4.2 Likelihood ratio tests

Let

$$\widehat{V}_\tau = \operatorname*{argmin}_{\beta_\tau \in \mathbb{R}^K} \sum_i \rho_\tau(y_i - x_i'\beta_\tau) \tag{2.77}$$

denote the value of the objective function at the unrestricted minimizer $\widehat{\beta}_\tau$ and

$$\widetilde{V}_\tau = \underset{\beta_\tau \in \mathbb{R}^K | R\beta_\tau = r}{\operatorname{argmin}} \sum_i \rho_\tau(y_i - x_i'\beta_\tau) \tag{2.78}$$

refer to the value of the objective function under the restricted estimator $\widetilde{\beta}_\tau$. Koenker and Machado (1999) showed that under the i.i.d. error assumption, the test statistic

$$L_\tau = \frac{2(\widetilde{V}_\tau - \widehat{V}_\tau)}{\tau(1-\tau)s(\tau)} \tag{2.79}$$

is asymptotically chi-square distributed under the null hypothesis with rank(R) degrees of freedom. The estimation of the sparsity function $s(\tau)$ has already been discussed. Koenker and Machado (1999) also proposed a second LR statistic

$$L_\tau = \frac{2n\sigma(\tau)}{\tau(1-\tau)s(\tau)} \log\left(\widetilde{V}_\tau/\widehat{V}_\tau\right) \tag{2.80}$$

with similar properties, assumed that $\sigma(\tau) = E\rho_\tau(u)$.

2.3.4.3 Rank tests

The classsical theory of rank tests employs the so-called rankscore function (see Hájek and Sidák (1967))

$$\widehat{a}_{ni}(\tau) = \begin{cases} 1 & \text{if } \tau < (R_i - 1)/n \\ R_i - \tau n & \text{if } (R_i - 1)/n \leq \tau \leq R_i/n \\ 0 & \text{if } R_i/n < \tau \end{cases} \tag{2.81}$$

where R_i represents the i-th observation y_i in $(y_1, \ldots, y_n)$. The integration of $\widehat{a}_{ni}(\tau)$ with respect to various score generating functions ψ creates vectors of rank-like statistics which are suitable for testing. For instance, integrating $\widehat{a}_{ni}(\tau)$ using the Lebesgue measure ($\psi(\tau) = \tau$) yields the so-called Wilcoxon scores

$$s_i = \int_0^1 \widehat{a}_{ni}(\tau)d\tau = \frac{R_i - 1/2}{n} \qquad i = 1, \ldots, n \tag{2.82}$$

while e.g. using $\psi(\tau) = \text{sgn}(\tau - 1/2)$ generates the so-called sign scores $s_i = \widehat{a}_{ni}(1/2)$. Gutenbrunner and Jureckova (1992) extended the theory to regression models by showing that the rankscores in (2.81) are a solution to the dual problem of the quantile regression minimization problem from equation (2.25). Gutenbrunner, Jureckova, Koenker, and Portnoy (1993) designed a test of the null hypothesis $H_0 : \beta_{\tau(2)} = 0$ based on the regression rankscore process for the model

$$y = X_1 \beta_{\tau(1)} + X_2 \beta_{\tau(2)} + u_\tau \tag{2.83}$$

First, the $\widehat{a}_{ni}(\tau)$'s are computed at the restricted model $y = X_1 \beta_{\tau(1)} + u_\tau$, leading to the corresponding rankscores vector with elements $s_i = \int \widehat{a}_{ni}(\tau) d\psi(\tau)$. Next, a vector

$$S_\tau = n^{-1/2} X_2' s \tag{2.84}$$

is formed, which converges in distribution to $\mathcal{N}(0, A^2(\psi)Q)$ under the null, where $A^2(\psi) = \int_0^1 \psi^2(\tau) d\tau$ and $Q = \lim_{n \to \infty} (X_2 - \widehat{X_2})'(X_2 - \widehat{X_2})/n$ with $\widehat{X_2} = X_1(X_1'X_1)^{-1}X_1'X_2$. Finally, the test statistic

$$T_\tau = \frac{S_n' Q^{-1} S_n}{A^2(\psi)} \tag{2.85}$$

can be calculated, which is asymptotically chi-square distributed under the null hypothesis with rank(X_2) degrees of freedom. An important feature of the test statistic T_τ is that it requires no estimation of nuisance parameters (as e.g. the sparsity function), since the functional $A(\psi)$ depends only on the score function, but not on the distribution of the error term (see e.g. Koenker (1994)). Huskova (1994) showed that inverted rank tests can be used to estimate confidence intervals for quantile regression parameters. Koenker (1997) and Hallin and Jureckova (1999) analyzed the efficiency of T_τ tests. Further references on rank tests in a quantile regression context are, among others, Koul and Saleh (1995), Hasan and Koenker (1997), and Mukherjee (1999).

2.3.4.4 Further approaches

In the literature, several further testing procedures under different conditions were developed. Buchinsky (1998b) proposed a test using the GMM framework, Melly (2001) described a test for symmetry based on Newey and Powell (1987), and Koenker and Xiao (2002a) used a martingale transformation to develop new tests. Chernozhukov (2002b) offered an alternative resampling test, Park (2002) compared different procedures, and He and Zhu (2003) developed a lack-of-fit test. Komunjer (2003) and Whang (2003) are two recent examples of contributions on the issue of specification testing.

2.3.5 Monte Carlo results

To conclude this section, we briefly mention several Monte Carlo studies that compared some of the presented approaches. Please refer to the cited articles for a more extensive discussion.

Buchinsky (1995a) examined six different estimation procedures of the asymptotic covariance matrix: an order statistic estimator (from Chamberlain (1994)), a design matrix bootstrap, a residual bootstrap, a sigma bootstrap (where the sparsity is bootstrapped), and two kernel approaches (one for the general model, one under i.i.d. assumption). He concluded that the design matrix bootstrap yields the best results, while some of the estimators with i.i.d. assumption performed poorly when the assumption was not satisfied (which is, after all, not too astonishing).

Koenker and Hallock (2000) compared the Hall and Sheather sparsity estimation, the sandwich formula proposed by Powell (see above), another sandwich proposed by Hasan and Koenker (1997), a rank inversion confidence interval, and the percentile estimation version of the bootstrap with 20, 200, and 600 repetitions. Their main conclusion was that *all* methods performed rather good. If forced to choose, one may want to take the bootstrap with 600 replications. In a similar study, Koenker and Portnoy (1999) in-

cluded the Parzen-Wei-Ying approach and a so-called Heqf-bootstrap. While the sparsity approach failed if the i.i.d. assumption was not satisfied (again not very surprising), they recommended the rank-inversion method due to faster computation time than any of the bootstrap procedures.

Chen (2001) also advocated the use of the rank-inverse test. Further studies can be found, among others, in Fitzenberger (1997b), Koenker and Machado (1999), Koenker and Xiao (2002a), and Kocherginsky, He, and Mu (2003).

2.4 Extensions and Applications

Since Koenker and Bassett's (1978) invention of quantile regression, a multiplicity of authors have extended the theoretical framework as well as presented empirical applications of the method. In this section, we briefly want to mention some of the (in our opinion) most important contributions to both branches of the literature. As, of course, many papers contain theoretical extensions *and* empirical elements, our classification can be seen as somewhat arbitrary.

2.4.1 Theoretical contributions

2.4.1.1 Weighted quantile regression

As we have already noted, the application of an appropriate weighting scheme creates opportunities for improved efficiency of estimation. Newey and Powell (1990) showed that the following estimator attains a semiparametric efficiency bound:

$$\widehat{\beta}_\tau \quad = \quad \operatorname*{argmin}_{\beta_\tau \in \mathbb{R}^K} \sum_i f_i(Q_\tau(y_i|x_i))\Big(\tau - \frac{1}{2} + \frac{1}{2}\mathrm{sgn}(y_i - x_i'\beta_\tau)\Big)(y_i - x_i'\beta_\tau) \qquad (2.86)$$

Similar results are given in Koenker and Zhao (1994). For other contributions on weighted quantile regression see e.g. Machado and Santos Silva (2001) and Zhao (2001).

2.4.1.2 Censored quantile regression

In section 2.2, we have seen that the quantile regression procedure is equivariant to monotone transformations of the response variable *and* robust against outliers of the regressand. From each of both properties, it follows immediately that the use of a censored dependent variable (e.g. top-coded income data) does not at all influence the results for conditional quantiles below the censoring threshold. Of course, this is not true for the conditional mean.

In order to enable the estimation of all conditional quantiles, Powell (1984, 1986) proposed the following estimator for a response variable with top coding value $\overline{y}$:

$$\widehat{\beta}_\tau \quad = \quad \operatorname*{argmin}_{\beta_\tau \in \mathbb{R}^K} \sum_i \rho_\tau \Big(y_i - \min\{\overline{y}, x_i'\beta_\tau\} \Big) \tag{2.87}$$

Fitzenberger (1997a) extended equation (2.87) to the case of observation specific censoring from the left ($\underline{y_i}$) and right ($\overline{y_i}$):

$$\widehat{\beta}_\tau \quad = \quad \operatorname*{argmin}_{\beta_\tau \in \mathbb{R}^K} \sum_i \rho_\tau \Big(y_i - \max\{\underline{y_i}, \min\{\overline{y_i}, x_i'\beta_\tau\}\} \Big) \tag{2.88}$$

A slight drawback of equations (2.87) and (2.88) is the fact that the estimation problem has no longer a strict linear programming representation. However, several procedures for censored quantile regression have been proposed, see, inter alia, Buchinsky and Hahn (1998), Chay and Honoré (1998), Chen and Khan (1998, 1999, 2000), Khan and Powell (1999), Yang (1999), Galfalvy, He, and Simpson (2000), Arias (2001), Chay and Powell (2001), Chernozhukov and Hong (2002), Honoré, Khan, and Powell (2002), and Portnoy (2003).

2.4.1.3 Regresssion depth

We have seen in section 2.2 that the quantile regression method is robust against outliers of the response variable, but not to extreme values of the regressors. Rousseeuw

and Hubert (1999) introduced an approach called "deepest regression" which is also robust against such so-called leverage points. However, the procedure is less intuitive, more computational demanding and less efficient than quantile regression (see e.g. He and Portnoy (1998)). Furthermore, additional problems can arise beyond a bivariate setting (see e.g. Koenker (2000)), so we will not further consider regression depth.

2.4.1.4   Autoregression, ARMA, and ARCH

A variety of authors have analyzed the characteristics to be taken into account if applying quantile regression to time series. Bloomfield and Steiger (1983) considered an autoregressive model under stationary conditions. Hasan and Koenker (1997) provided a test for the unit root hypothesis. Davis and Dunsmuir (1997) analyzed regression models with ARMA errors. Koenker and Zhao (1996) proposed a specification for ARCH models.

For more on quantile regression for time series see, among others, Weiss (1991), Koul and Saleh (1995), Herce (1996), Hallin and Jureckova (1999), Galbraith, Zernov, and Zinde-Walsh (2001), Ling and McAleer (2001), El Bantli and Hallin (2002), Koenker and Xiao (2002b), Furno (2003), and Ferreira (2003).

2.4.1.5   Two-stage and IV quantile regression

Amemiya (1982) and Powell (1983) proposed a two-stage estimation procedure for the conditional median. Some recent contributions to the implementation of quantile regression for endogenous response variables can be found in Chernozhukov and Hansen (2001), Arias, Hallock, and Sosa-Escudero (2001), Chernozhukov and Hansen (2002), Abadie, Angrist, and Imbens (2002), Lucchetti (2002), Blundell and Powell (2003a), Blundell and Powell (2003b), Kim and Muller (2003), Chernozhukov and Hong (2003), Honoré and Hu (2004), Lee (2004), and Ma and Koenker (2004).

2.4.1.6 Nonlinear quantile regression

So far, we have confined our analysis to the linear-in-parameters quantile regression model, since it offers a flexible approach suitable for many applications. Furthermore, a linear model can always be regarded as the (best) linear approximation of possibly underlying nonlinear relationships. However, it is of course also possible to define a *nonlinear* quantile regression estimator:

$$\widehat{\beta}_\tau \quad = \quad \operatorname*{argmin}_{\beta_\tau \in \mathbb{R}^K} \sum_i \rho_\tau \Big(y_i - g(x_i, \beta_\tau) \Big) \tag{2.89}$$

Koenker and Park (1996) described a general approach for the computation of (2.89). Other analyses of nonlinear quantile regression include Weiss (1991), Jureckova and Prochazka (1994), Koenker, Ng, and Portnoy (1994), Welsh, Carroll, and Ruppert (1994), Fitzenberger (1997a), Hunter and Lange (1998), Mukherjee (1999), Taylor (1999), Craig and Ng (2000), Engle and Manganelli (2002), and Lee (2003).

2.4.1.7 Nonparametric quantile regression

Several authors have addressed the estimation of nonparametric quantile regression models. Chaudhuri (1991) proposed a locally polynomial approach, while Hendricks and Koenker (1992) introduced the computation of quantile regression splines. For more on nonparametric methods see, inter alia, Koenker, Ng, and Portnoy (1994), Abberger and Heiler (1997), Yu and Jones (1998), Khan (2001), Nahm (2001), Horowitz and Lee (2002), Yu and Lu (2002), Yu (2002), Chen and Khan (2003), Gannoun, Saracco, and Yu (2003), and Otsu (2003).

2.4.1.8 Multivariate quantile regression

Having read so far, one might ask why we have confined ourselves to the estimation of univariate conditional quantiles. However, the definition of multivariate (condi-

tional) quantiles is not straight-forward. Some proposals have been given by Chaudhuri (1996), Koltchinskii (1997), Chakraborty and Chaudhuri (1998), De Gooijer, Gannoun, and Zerom (2002), and Chakraborty (2003).

2.4.1.9 Further extensions

Manski (1985) introduced least absolute deviation estimation for binary response variables. Subsequent articles on quantile regression for discrete regressands include Horowitz (1992), Kordas (2002a), Kordas (2002b), Blundell and Powell (2003b), Kordas and Lehrer (2003), Bilias and Haliassos (2004), and Lee and Yang (2004).

Machado and Santos Silva (2002) proposed a jittering procedure to estimate count data. Other contributions for duration models can be found in Ying, Jung, and Wei (1995), Chaudhuri, Doksum, and Samarov (1997), Koenker and Bilias (2001), and Koenker and Geling (2001).

Chay (1995) extended the focus of quantile regression onto the estimation for panel data. Chen and Khan (1998), He and Kim (2002), He, Fu, and Fung (2003), and Koenker (2003) also analyzed longitudinal data sets.

Chernozhukov (2002a) provided a comprehensive theory for estimated conditional extremes and near-extremes. For more on extremal quantile regression see also Portnoy and Koenker (1989), Portnoy and Jureckova (1999), and Chernozhukov (2000).

Several articles have introduced a linear Bahadur representation for the quantile regression estimator, which can be helpful for inferential assertation. See for example Koenker and Portnoy (1987), Chaudhuri (1991), Hendricks and Koenker (1992), Gutenbrunner and Jureckova (1992), and He and Shao (1996).

A regression method that is somewhat related to quantile regression is the so-called asymmetric least squares estimation procedure. A variety of names have been proposed for the computed regression outputs: expectiles, percentiles, graviles, heftiles, loadiles,

or projectiles. See, among others, Aigner, Amemiya, and Poirier (1976), Newey and Powell (1987), and Efron (1991).

Hahn (1997) examined the quantile regression process under a Bayesian setting. Subseqent references include Jureckova and Klebanov (1999), Yu and Moyeed (2001), and Tsionas (2003).

Buchinsky (2001) and Fitzenberger (2003) addressed sample selection issues, Chaudhuri and Loh (2002) and Bergez, Cardot, and Garcia (2003) introduced quantile regression trees, Morillo (1999b) used a recursive estimation procedure, and Hastie and Loader (1993) proposed a kernel smoothing of the estimator.

2.4.2 Empirical examples

2.4.2.1 Labor and educational economics

Inspired by Chamberlain's (1994) address to the World Congress of the Econometric Society and Buchinsky's (1991, 1994, 1995b, 1997, 1998a, 2001) pioneering work, a variety of subsequent authors has applied quantile regression to labor economics and educational questions. Without claiming to be exhaustive, we could mention Poterba and Rueben (1994), Chay and Honoré (1998), Eide and Showalter (1998b), Eide and Showalter (1998a), Kahn (1998), Schultz and Mwabu (1998), Maloney and Ribeiro (1999), Morillo (1999a), Machado and Mata (2000), Pereira and Martins (2000), Fertig (2001), Fitzenberger, Hujer, MaCurdy, and Schnabel (2001), García, Hernández, and López-Nicoláz (2001), González and Miles (2001), Hartog, Pereira, and Viera (2001), Levin (2001), Machado and Mata (2001), Montenegro (2001), Nielsen and Rosholm (2001), Ribeiro (2001), Tannuri-Pianto (2001), Bassett, Tam, and Knight (2002), Ding and Lehrer (2002), Fitzenberger and Wunderlich (2002), Fitzenberger and Kurz (2002), Fersterer and Winter-Ebmer (2002), Guimaraes (2002), Madalozzo (2002), Skans (2002), Tam, Bassett, and Sukhatme (2002), Tannuri-Pianto and Pianto (2002), Wang (2002),

Albrecht, Björklund, and Vroman (2003), Bedard (2003), Eide, Goldhaber, and Showalter (2003), Falaris (2003), Hallock, Madalozzo, and Reck (2003), Maloney, Cunningham, and Bosch (2003), Melly (2003), Nguyen, Albrecht, Vroman, and Westbrook (2003), Rangvid (2003), Schunk (2003), Reck (2003), Angrist, Chernozhukov, and Fernández-Val (2004), Kee (2004), Wößmann (2004), Melly (2004), and Martins (2004).

2.4.2.2   Time series and financial data

Another important field of successfully implemented quantile regression methods is the analysis of time series and financial data. Refer for example to Koul and Saleh (1995), Abberger (1997), Chang and Weigend (1999), Hallin and Jureckova (1999), Mukherjee (1999), Taylor (1999), Taylor and Bunn (1999), Leggett and Craighead (2000), Taylor (2000), Bassett and Chen (2001), Chernozhukov and Umantsev (2001), Galbraith, Zernov, and Zinde-Walsh (2001), Barnes and Hughes (2002), Chen and Chen (2002), Cai (2002), Engle and Manganelli (2002), Komunjer (2002), Park (2002), and Bassett, Koenker, and Kordas (2004).

2.4.2.3   Medicine and health economics

Also in medical and health economics applications, quantile regression has proved to be a viable alternative to other competing estimation methods. Examples are Cole (1988), Cole and Green (1992), Gasser, Ziegler, Seifert, Prader, Molinari, and Largo (1994), Royston and Altman (1994), Manning, Blumberg, and Moulton (1995), Lipsitz, Fitzmaurice, Molenberghs, and Zhao (1997), Wright and Royston (1997), Royston and Wright (1998), Heagerty and Pepe (1999), Royston and Wright (2000), Jones and Nicolás (2002), Variyam, Blaylock, and Smallworld (2002), Sun (2003), Yee (2004), and Kan and Tsai (2004).

2.4.2.4 Environmental applications

To demonstrate the universality of quantile regression, we further want to mention some applications with biological and geological background. Please refer to Cade, Terrell, and Schroeder (1999), Haire, Bock, Cade, and Bennett (2000), Haire, Bock, and Cade (2000), Green and Kozek (2001), Knight and Ackerly (2002), Sankarasubramanian and Lall (2003), and Cade (2003).

2.4.2.5 Economic growth and welfare

Finally, back in economics, several articles have used quantile regression for the analysis of growth and welfare. See for example Mello and Novo (2002), Cunningham (2003), Mello and Perrelli (2003), Gomanee, Girma, and Morrissey (2003), and Barreto and Hughes (2004).

2.5 Concluding Remarks

Instead of repeating all features, properties, and advantages of quantile regression from the last four sections, we want to conclude this chapter by notionally turning back time for a moment to the year 1975. At that time, roughly three years *before* the "invention" of quantile regression, Robert V. Hogg (1975) analyzed the relationship between the salary and the number of years in rank of 96 American statistics professors. He quickly remarked that the classical homoscedastic assumption of the error term seemed to be violated in the data set.

So, he proposed the following method for the estimation of percentile regression lines: First, he divided the data set into two groups according to x (years of rank) with equal number of observations in both groups. Then, he plotted all 96 observations in a (x, y) coordinate system. Finally, he took a pencil and a ruler and moved the ruler as long as

The Case of 96 Professors' Salaries
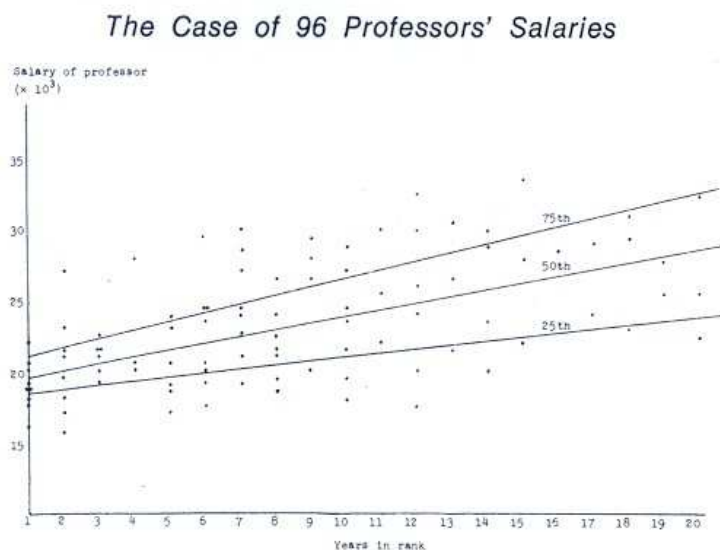
Figure 2.12: Hogg's estimator. The figure shows Hogg's (1975, page 58) estimate of percentile regression lines for the salary of 96 statistics professors, conditional on the number of years in rank (see text).

$n\tau$ points were below and $n(1-\tau)$ were above it *simultaneously in both groups* to draw the "estimated" τ-th percentile regression line. Figure 2.12 shows the result for three different values of τ.

Without intending to denigrate Hogg's approach, in our opinion a comparison between his method and the quantile regression procedure as described in the last four sections should clearly clarify the ground-breaking characteristics of Koenker and Bassett's (1978) proposal. The next three chapters contain the application of their method to three different empirical settings.

Chapter 3

Household Demand for Consumption Goods

Abstract

This chapter uses cross-section micro data from the GfK consumer panel for an econometric demand analysis of households in Germany. Contrary to most research which considered "average" behavior, we analyze consumer behavior for different "intensities" of consumption. Our analytical tool is quantile regression which allows us to describe the conditional distribution for any quantile including the (conditional) median representing "average" behavior. As an illustrative example, we use the demand for beer and wine. The results show quite distinct patterns regarding price and income effects for light and heavy consumption, respectively, which leads to a better characterization of household demand.

3.1   Introduction

Econometric demand analysis played an important role in the 70's and the 80's: Christensen, Jorgenson, and Lau (1975) proposed the "Translog demand system" and Deaton and Muellbauer (1980) introduced an alternative specification termed "Almost Ideal Demand System" (AIDS). Both approaches based their empirical analysis on the aggregate time series data from a sample of households continuing the work started by Richard Stone who established the "Linear Expenditure System" quite a while earlier (see Stone (1954)). Many studies failed in trying to verify the constraints like symmetry and Engel and Cournot aggregation established by (static) microeconomic theory. Econometric issues arose from the fact that both translog and AIDS were formulated in terms of (dynamic) share equations; the implied problems have been solved only marginally. See Ronning (1992) for an overview.

On the other hand, surprisingly little work has been done in consumption analysis on the basis of individual cross-section (or panel) household data. This fact is even more worth mentioning when comparing it with a huge bulk of microeconometric studies involved in qualitative choice behavior initiated by Daniel McFadden (1974) who studied the structure of travel demand. Exceptions are, among others, the microeconometric studies based on the British family expenditure survey (see for example Atkinson, Gomulka, and Stern (1990) and Blundell, Pashardes, and Weber (1993)). Deaton (1997) is the most recent example for this kind of research who studies the consumption pattern in underdeveloped countries.

All research so far has concentrated on "average" behavior, i.e. on the expected value of the conditional distribution. Our study is concerned with a more detailed description of characteristics of this (conditional) distribution; in particular we consider quantiles of consumption which means that we do not only consider average behavior but also behavior of "extreme" consumers thereby installing a new characterization of consumer

behavior.[1] For example, it might well happen for some good that extreme consumers' demand elasticity differs in sign from the one shown by the average consumer: the study by Manning, Blumberg, and Moulton (1995, page 125) on the demand for alcohol tries to find out "...whether light and heavy drinkers have different price responses". In our analysis, we try to give an explanation for this varying demand structure which to our best knowledge so far has only be noted - in a quite different context - by the paper just mentioned. We use data from the German GfK consumer panel[2] for the year 1995. Special attention is given to the kind of temporal aggregation employed in order to include all purchases within this period.

The chapter is organized as follows: in section two, we establish or rather report some basic results from empirical demand analysis needed when interpreting our own empirical results later on. Section three contains a short description of the data. In section four, we report on our results from which we derive some tentative more general statements regarding consumer behavior. Section five summarizes our findings and concludes.

3.2 Econometric Demand Analysis

The main concern of econometric demand analysis is with consumers' reactions to changes in prices and income[3] as described, for example, in Varian's textbook on mi-

---

[1]Former studies have tried to provide such information by evaluating the price elasticities at certain quantiles of the expenditures. See, for example, Blundell, Pashardes, and Weber (1993, table 3, part D and E) or - as a most recent example - Newey (2001, tables 3 to 5).

[2]GfK = Gesellschaft für Konsumforschung, Nuremberg/Germany.

[3]The table displays the possible cases. Normal goods are further separated into luxuries when income elasticity is greater than one and necessities if income elasticity is smaller than one. Two goods are called substitutes if the cross-price elasticity is positive and complements if it is negative.

		change of income μ	
		"normal good" $\partial q_j/\partial \mu > 0$	"inferior good" $\partial q_j/\partial \mu < 0$
change of price p_j	"ordinary good" $\partial q_j/\partial p_j < 0$	example: butter	example: margarine
	"Giffen good" $\partial q_j/\partial p_j > 0$		example: potatoes in Ireland, 19-th century

croeconomic theory (see Varian (1992, section 3.3)). Two main approaches have been
used:[4]

- analysis of Engel curves for certain goods or groups of goods, that is the relation
  between consumption and income for a certain good (group of goods). From this
  analysis the income elasticity may be deduced.

- estimation of demand systems (LES, Translog, AIDS, generalized Leontief) using
  information on prices and quantities for all goods (groups of goods). Only this
  approach allows an adequate examination of substitution patterns between goods.

In the following, we list some of the most important topics and problems arising from
econometric demand analysis:

(a) Aggregate data versus micro data It has already been pointed out in the intro-
duction that most of the work concerning estimation of demand systems has been
done on the basis of (monthly or yearly) aggregate data for some population. Typi-
cally, consumption shares for a moderate number of categories have been analyzed
over time. See for example the pioneering papers by Christensen, Jorgenson, and
Lau (1975) and Deaton and Muellbauer (1980). Most of the studies failed to verify
the restrictions postulated by (static) microeconomic theory.

About a decade ago, Richard Blundell and others posed the question: "What do we
learn about consumer demand patterns from micro data?" Their paper shows - at
least for the data set used - that estimation on the basis of micro data much bet-
ter fulfills the demand restrictions (Blundell, Pashardes, and Weber (1993, page
577)), a fact which to our best knowledge has not been appropriately noted in the
literature. On the other hand, each model based on micro data has to relate its
results to the macro level thereby facing the problem of aggregation. This aspect
has been treated, too, in Blundell, Pashardes, and Weber's (1993) paper.[5]

---

[4]See, for example, Ronning (1988) for an overview.

[5]Ronning and Zimmermann (1991) give an introduction to a series of papers in *Ifo-Studien* on the
relevance of microeconometric models for economic policy.

(b) Income versus expenditure (consumption versus purchase) Ideally, consumption for all goods and services should be included into the analysis. However, usually only a subset of goods and services has been considered, so that it is unclear how "income" has been distributed between this subset and the remaining goods. Therefore, the expenditures for this subset are used instead. Moreover, Keen (1986) has stressed the important distinction between consumption and purchase of a good: "Zero consumption" can only be defined via observed purchases. If an household buys a good infrequently, we speak of zero consumption, although it is not clear whether the good really is not consumed. See, for example, Ronning (1988, page 71) and Blundell, Pashardes, and Weber (1993, page 575).

(c) Price information Prices of most goods will not vary over individuals in a cross-section. This is an argument in favor of disregarding price effects in the analysis of Engel curves where typically cross-section data are used. On the other hand, it complicates the estimation of price effects on consumption, especially for goods with regulated prices (price maintenance agreements, for example). The situation is improved when panel data are available. If groups of goods are used in microeconometric research, then prices have to be aggregated (see, for example, Blundell, Pashardes, and Weber (1993)). The data set used in our analysis contains information for single households over one year indicating expenditures and quantities for each single purchase. However, prices have to be derived from these data. As we shall explain in section three, prices have been averaged over the whole year to make the econometric demand analysis possible since all other variables, in particular income, are only given on a yearly basis.

Manning, Blumberg, and Moulton (1995) have pointed out that consumers with identical income and facing the same price of a good may behave quite differently depending on the *intensity of consumption*. In their study on the demand for alcohol, they find remarkable differences between "heavy" and "light" drinkers with respect to the own-price elasticity and income elasticity. For example, the same good is considered

as "inferior" by some consumers and "normal" by others. Another example is given by Koenker and Hallock (2001), where different Engel curves are shown for "heavy" and "light" consumption implying varying income elasticities for these subgroups.[6] In both cases, the method of quantile regression first proposed by Koenker and Bassett (1978) is employed.

3.3 The Data

Our empirical analysis is based on data from the *ConsumerScan* household panel maintained by "Gesellschaft für Konsumforschung" (GfK) since 1957. The panel currently consists of about 12,000 households constantly reporting their purchases of *Fast Moving Consumer Goods*. A subset of this data set is available for scientific use from Zentrum für Umfragen, Methoden und Analysen (ZUMA) at Mannheim.[7] This file is confined to the year 1995.

The data set contains all 9,064 households continuously reporting their purchases in 1995. The products are divided into 81 categories and no brand names are given. For each individual purchasing act the following information is available: date, day of the week, subcategory of a certain good, type of retailer, product identification number, type of price (normal/special), total quantity, amount spent, time since last buying. Other specific characteristics of the products are given as well (for example, packaging). For some product categories only a subsample (4,426 respectively 4,638) of households has been reporting.

Table 3.1 provides some descriptive statistics. In particular it contains information about the total number of purchases from which we estimated the average purchasing frequency: For example, households buy about 49 times milk and about 21 times mineral water during the year. Additionally, the table reports the proportion of house-

---

[6]The paper uses the original data from Engel's study on food expenditures (see Engel (1857)).

[7]A detailed description of the provided data can be found in Papastefanou (2001).

Table 3.1: Descriptive statistics

ident.	product	purchasing acts	number of households	purchasing frequency (av.)	proportion of non-buyers (%)
04	detergents for dishes	34556	7780	4.44	14.17
08	milk	204339	*4185	48.83	5.45
12	pure coffee (roasted)	143194	8457	16.93	6.70
17	frozen food	230841	8175	28.24	9.81
22	fats	233124	*4617	50.49	0.45
33	beer	131245	7485	17.53	17.42
35	wine	27614	2954	9.35	67.41
46	lemonade	155447	7254	21.43	19.97
66	animal food	123133	3056	40.29	66.28
84	mineral water	174470	8414	20.74	7.17
91	pasta	53201	*4165	12.77	10.20
99	toilet paper	32435	*4039	8.03	8.74

Note: An asterisk indicates that only a subsample of the households has been reporting.

Table 3.2: Distribution of income

net income	number	percentage	accumulated
up to 499 DM	24	0.26	0.26
500 DM - 999 DM	189	2.09	2.35
1000 DM - 1249 DM	312	3.44	5.79
1250 DM - 1499 DM	366	4.04	9.83
1500 DM - 1999 DM	890	9.82	19.65
2000 DM - 2499 DM	1235	13.63	33.27
2500 DM - 2999 DM	1233	13.60	46.88
3000 DM - 3499 DM	1235	13.63	60.50
3500 DM - 3999 DM	907	10.01	70.51
4000 DM - 4499 DM	852	9.40	79.91
4500 DM - 4999 DM	484	5.34	85.25
5000 DM - 5499 DM	474	5.23	90.48
5500 DM and more	829	9.15	99.62
not reported	34	0.38	100.00
total	9064	100.00	

holds not buying from a certain product group. This ranges from 0.45% not buying fats to 67.41% not buying wine and thereby illustrating the fact of "zero consumption" as discussed in section 2.

Furthermore, some socioeconomic and demographic information is provided on a household basis. Regional information includes federal state and size of the community. Additional variables are age of head of household, number of children in different age groups (up to 6, 7-14 and 15-18 years), income, occupational and educational status. Moreover, information about equipment of the household is provided. For some of the households these variables are missing.

A major drawback of this data set is the fact that income and age are reported only in grouped form. Table 3.2 shows the number of missing values as well as the group-

ing for the variable household income which plays a central role in demand analysis. Finally, the GfK panel reports on attitudes of consumers. These attitudes concern e.g. nutritional, environmental and other aspects (like price consciousness) of daily life.

For our empirical analysis, some aggregation of the data is required. This results from the fact that we have detailed information on each purchase regarding quantity and amount spent from which we can deduce - for each purchase - the price per unit by computing for each product

$$\text{price} \quad = \quad \frac{\text{amount spent}}{\text{quantity}} \quad .$$

However, we have - of course - only one observation regarding income for each household. Therefore, we determine the yearly average price for each household by computing for each product

$$\text{average price} \quad = \quad \frac{\text{total amount spent by a household on this product within the year}}{\text{total quantity within the year}} \quad .$$

In the following, we call this derived average price simply "the price". In order to illustrate the effect of our aggregation approach, we show in figure 3.1 the distribution of quantity, amount spent and price (per purchase) for the case of beer. Note the peaks at 10 and 20 liters whereas the price distribution is rather well behaved. Figure 3.2 then displays the result of aggregation in two scatter diagrams (double linear scale and double log scale) for price and quantity where each point represents one household.[8]

3.4   Empirical Results

In the following, we present our estimation results. We start by showing the least-squares estimates as a sort of benchmark which then are contrasted with outcomes

---

[8]Due to size restrictions, only a randomly chosen 20% subsample of all households is presented.
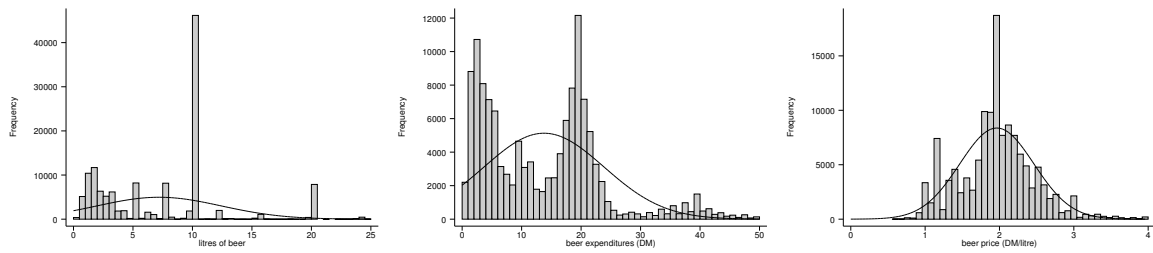
Figure 3.1: Beer consumption (quantity, expenditures and price)
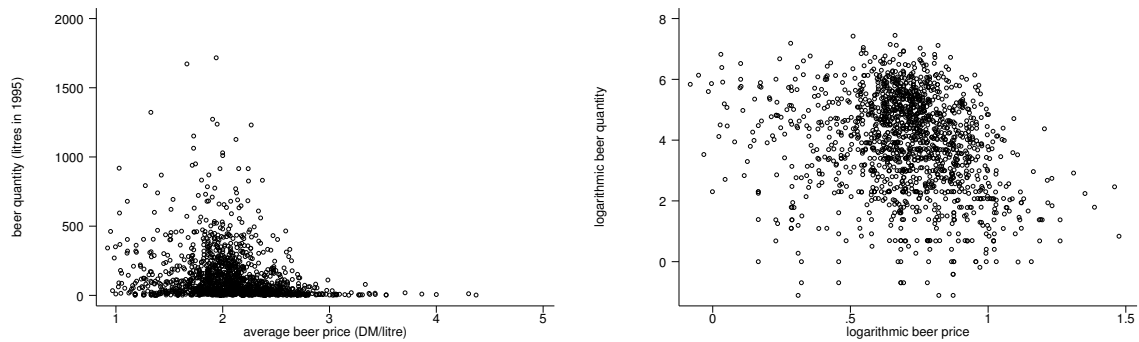
Figure 3.2: Aggregated beer data (levels and logs)

from quantile regression. In the last subsection, we will discuss our estimation results with regard to possible bias due to potential endogeneity of regressors.

We will concentrate on the consumption of beer and wine since alcohol consumption has a particularly clear interpretation of the "intensity" of consumption. We also would like to contrast our results with those from Manning, Blumberg, and Moulton (1995). Later on (see subsection 3.4.3) we add some results for other products trying to explain the varying demand patterns for different goods more generally.

We use the simplest specification possible relating quantity and price by a log-linear model. This has the advantage that coefficients can be interpreted as elasticities. Income is available only in grouped form. We therefore first exploit this information as well as that from other discrete or grouped explanatory variables (age, household size). However, in order to obtain at least rough estimates for the income elasticity, we construct an artificial continuous income variable. The same is done for age. Details of our data transformations are presented in subsection 3.4.2.

Table 3.3: Least-squares regression (beer)

log. beer quantity	coef.	std. err.	t	$P > |t|$	95% conf. interval	
constant	3.35284	.1889605	17.74	0.000	2.982424	3.723256
log. av. price	-1.753606	.0786581	-22.29	0.000	-1.907799	-1.599414
income $\in$ [1000,1249]	.0878681	.1583885	0.55	0.579	-.2226182	.3983544
income $\in$ [1250,1499]	.0880574	.1531222	0.58	0.565	-.2121056	.3882203
income $\in$ [1500,1999]	.2436030	.1359381	1.79	0.073	-.0228742	.5100802
income $\in$ [2000,2499]	.1308066	.1341529	0.98	0.330	-.1321710	.3937842
income $\in$ [2500,2999]	.3098026	.1353236	2.29	0.022	.0445299	.5750752
income $\in$ [3000,3499]	.3405131	.1359922	2.50	0.012	.0739298	.6070963
income $\in$ [3500,3999]	.3250298	.1392488	2.33	0.020	.0520627	.5979969
income $\in$ [4000,4499]	.3633615	.1398491	2.60	0.009	.0892177	.6375053
income $\in$ [4500,4999]	.3150890	.1476771	2.13	0.033	.0256000	.6045780
income $\in$ [5000,5499]	.2552472	.1475486	1.73	0.084	-.0339899	.5444842
income >= 5500	.2976291	.1412991	2.11	0.035	.0206429	.5746153
hhsize = two	1.062992	.0522729	20.34	0.000	.9605224	1.165462
hhsize = three	1.213293	.0622057	19.50	0.000	1.091352	1.335234
hhsize = four	1.320731	.0689660	19.15	0.000	1.185538	1.455924
hhsize = five	1.330039	.0938747	14.17	0.000	1.146018	1.514060
hhsize >= six	1.357907	.1490717	9.11	0.000	1.065685	1.650130
age $\in$ [25,29]	.3611248	.1597493	2.26	0.024	.0479710	.6742786
age $\in$ [30,34]	.4054011	.1567852	2.59	0.010	.0980577	.7127445
age $\in$ [35,39]	.5560247	.1573868	3.53	0.000	.2475020	.8645473
age $\in$ [40,44]	.6279981	.1572959	3.99	0.000	.3196535	.9363426
age $\in$ [45,49]	.7142163	.1578534	4.52	0.000	.4047788	1.023654
age $\in$ [50,54]	.8675696	.1564659	5.54	0.000	.5608521	1.174287
age $\in$ [55,59]	.8535417	.1536883	5.55	0.000	.5522691	1.154814
age $\in$ [60,65]	.7362940	.1544137	4.77	0.000	.4335994	1.038989
age $\in$ [65,69]	.5831481	.1545569	3.77	0.000	.2801727	.8861235
age >= 70	.4721374	.1534021	3.08	0.002	.1714259	.7728490

3.4.1 Least-squares estimation

Table 3.3 shows the results from least-squares estimation for the consumption of beer. In addition to the price of this good, the impacts of income (grouped), age of head of household (grouped) and household size are considered by defining three sets of dummies. For each categorical variable, the first category is omitted (household size one, income lower than 1000 DM and age less than 25). We note a rather pronounced price elasticity of -1.75 which is comparable in size to the estimated elasticities of alcohol in Blundell, Pashardes, and Weber (1993, table 3) and Manning, Blumberg, and Moulton (1995, table 2). Household size matters much more than age or income when looking at the t-ratios. We note for later reference that income has an (albeit slight) significant effect at higher income classes.

Turning to the corresponding results for wine (see table 3.4), we obtain a quite different picture: the price elasticity is *positive* making wine a "Giffen" good which is at odds

Table 3.4: Least-squares regression (wine)

log. wine quantity	coef.	std. err.	t	$P > \|t\|$	95% conf. interval	
constant	.8621149	.2742967	3.14	0.002	.3242790	1.399951
log. av. price	.1011763	.0542206	1.87	0.062	-.0051385	.2074911
income ∈ [1000,1249]	.1044016	.224500	0.47	0.642	-.3357938	.5445971
income ∈ [1250,1499]	-.1895866	.2251126	-0.84	0.400	-.6309831	.2518100
income ∈ [1500,1999]	.3246291	.1906123	1.70	0.089	-.0491199	.6983782
income ∈ [2000,2499]	.3041583	.1862344	1.63	0.103	-.0610067	.6693233
income ∈ [2500,2999]	.3381186	.1900110	1.78	0.075	-.0344516	.7106887
income ∈ [3000,3499]	.5028804	.1919761	2.62	0.009	.1264571	.8793038
income ∈ [3500,3999]	.4895946	.1978534	2.47	0.013	.1016473	.8775419
income ∈ [4000,4499]	.5577149	.1964025	2.84	0.005	.1726125	.9428172
income ∈ [4500,4999]	.6424704	.2096058	3.07	0.002	.2314791	1.053462
income ∈ [5000,5499]	.8378914	.2125869	3.94	0.000	.4210548	1.254728
income >= 5500	.9920589	.1991890	4.98	0.000	.6014927	1.382625
hhsize = two	.2183280	.0779064	2.80	0.005	.0655705	.3710855
hhsize = three	.0995142	.0956803	1.04	0.298	-.0880939	.2871223
hhsize = four	.2701201	.1071512	2.52	0.012	.0600201	.4802202
hhsize = five	.1343788	.1470194	0.91	0.361	-.1538942	.4226518
hhsize >= six	.2052321	.2434585	0.84	0.399	-.2721368	.6826010
age ∈ [25,29]	.3174385	.2453108	1.29	0.196	-.1635625	.7984394
age ∈ [30,34]	.4703598	.2413946	1.95	0.051	-.0029624	.943682
age ∈ [35,39]	.4929683	.2423548	2.03	0.042	.0177633	.9681732
age ∈ [40,44]	.6611684	.2403771	2.75	0.006	.1898414	1.132495
age ∈ [45,49]	.6735898	.2424984	2.78	0.006	.1981033	1.149076
age ∈ [50,54]	.8447658	.2408852	3.51	0.000	.3724424	1.317089
age ∈ [55,59]	.7958071	.2366303	3.36	0.001	.3318268	1.259787
age ∈ [60,64]	.6219815	.2366123	2.63	0.009	.1580364	1.085927
age ∈ [65,69]	.6377288	.2378547	2.68	0.007	.1713476	1.104110
age >= 70	.7447125	.2352835	3.17	0.002	.2833730	1.206052

with a priori expectations. However, the estimate is not significantly different from zero. Household size is almost insignificant contrary to the results for beer. The income effect is nearly monotone, i.e. coefficients are greater for larger incomes. Age, too, has an impact on wine consumption.

3.4.2 Results from quantile regression

We now turn to the results from quantile regression which analyzes the conditional distribution to a greater extent. For this, we compute quantile regressions for 99 different quantiles ($\tau = 0.01, \ldots, 0.99$) by the methods discussed in chapter two. Again, we include the explanatory variables price, household size, income and age as explanatory variables. Table 3.5 contains the regression results for some selected quantiles.

A more comprehensive way of presenting the results is in form of graphics: Figure 3.3 displays the estimated price elasticities for all 99 quantiles. The 95% confidence

Table 3.5: Quantile regression results (beer)

log. beer quantity	least sq.	5% quant.	25% quant.	50% quant.	75% quant.	95% quant.
constant	3,353***	0,174	2,263***	3,91***	4,601***	5,235***
log. av. price	-1,754***	-0,908***	-1,895***	-2,196***	-1,889***	-1,41***
income ∈ [1000,1249]	0,088	0,394	0,231	-0,307	-0,114	0,589***
income ∈ [1250,1499]	0,088	0,029	0,164	-0,107	0,056	0,632***
income ∈ [1500,1999]	0,244*	0,122	0,457**	0,148	0,145	0,606***
income ∈ [2000,2499]	0,131	0,046	0,289*	0,023	0,012	0,431**
income ∈ [2500,2999]	0,310**	0,410	0,497***	0,233	0,117	0,479***
income ∈ [3000,3499]	0,341**	0,234	0,539***	0,252	0,162	0,496***
income ∈ [3500,3999]	0,325**	0,612*	0,460***	0,215	0,127	0,473***
income ∈ [4000,4499]	0,363***	0,452	0,559***	0,276	0,204	0,455***
income ∈ [4500,4999]	0,315**	0,439	0,556**	0,183	0,153	0,391**
income ∈ [5000,5499]	0,255*	0,418	0,469***	0,160	0,034	0,418**
income >= 5500	0,298**	0,443	0,438***	0,226	0,038	0,375**
hhsize = two	1,063***	0,895***	1,097***	1,184***	1,125***	0,689***
hhsize = three	1,213***	1,258***	1,311***	1,289***	1,203***	0,799***
hhsize = four	1,321***	1,528***	1,495***	1,399***	1,250***	0,816***
hhsize = five	1,330***	1,494***	1,357***	1,409***	1,320***	0,911***
hhsize >= six	1,358***	1,363***	1,343***	1,408***	1,507***	1,412***
age ∈ [25,29]	0,361**	0,663*	0,454	0,183	0,283	0,374
age ∈ [30,34]	0,405***	0,521**	0,471*	0,302	0,446***	0,514*
age ∈ [35,39]	0,556***	0,578**	0,593**	0,505**	0,662***	0,601**
age ∈ [40,44]	0,628***	0,474*	0,646**	0,584***	0,739***	0,720**
age ∈ [45,49]	0,714***	0,469*	0,764***	0,688***	0,818***	0,888***
age ∈ [50,54]	0,868***	0,953***	1,038***	0,861***	0,809***	0,832***
age ∈ [55,59]	0,854***	0,97***	0,951***	0,783***	0,843***	0,820***
age ∈ [60,64]	0,736***	0,915***	0,864***	0,618***	0,790***	0,741**
age ∈ [65,69]	0,583***	0,572**	0,718***	0,589***	0,562***	0,501*
age >= 70	0,472***	0,458*	0,518*	0,360*	0,513***	0,587**

Note: An asterisk indicates that the coefficient is significantly different from zero at the 90%-level (** at the 95%-level, *** at the 99%-level). The t-values have been calculated by bootstrapping.
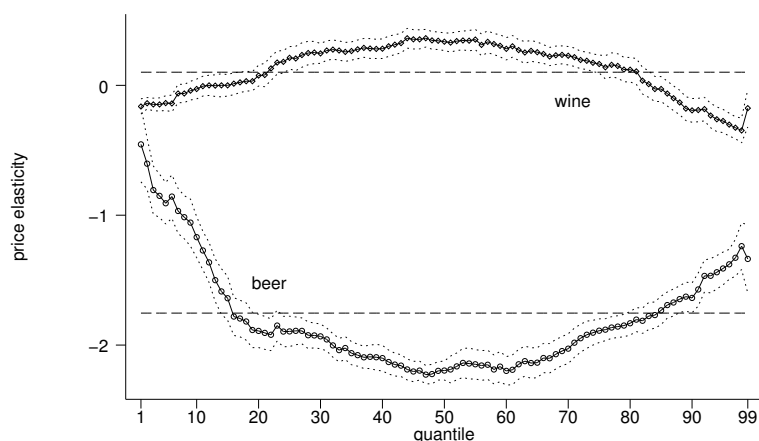
Figure 3.3: Comparison of price elasticity coefficients

bands from bootstrapped estimation errors are also shown as dotted lines. The same figure shows additionally the corresponding results for wine. We note at first sight the positive price elasticity of wine and a large negative elasticity for beer at the median (50% quantile) which is roughly comparable to the least-squares procedure given in subsection 3.4.1 and presented in this figure by horizontal dashed lines. Note that the confidence band regarding elasticities for wine is strictly positive for quantiles around 50%, whereas the corresponding least-squares estimate was insignificant.

Taking a closer look reveals interesting findings: For beer, the price elasticity coefficient shows a pronounced U-shaped form, starting at values between -0.4 and -0.9 for small quantiles, peaking at -2.23 (47% quantile) and coming back to values around -1.3 for the largest quantiles. In other words, those consumers either purchasing a very little *or* a very high amount of beer are much less price sensitive than "average" consumers. These findings could be explained as follows: People with small beer consumption do not care much about price , whereas some of the heavy consumers may be partly addicted to alcohol and therefore as well less price sensitive. The presented results coincide in some way with the findings of Manning, Blumberg, and Moulton (1995). They analyzed the relationship between alcohol consumption (not only beer) and regional average price by quantile regression and also reported a U-shaped price elasticity.
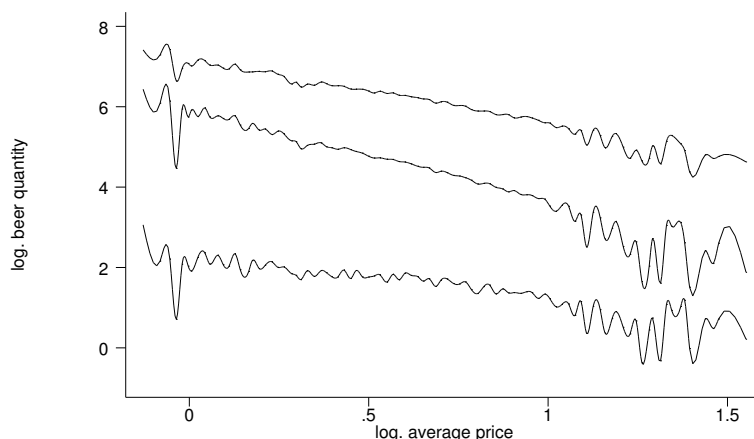
Figure 3.4: Estimated conditional 5%-, 50%- and 95%-quantiles

The corresponding results for wine in figure 3.3 however remind us that the U-shaped pattern is not typical for all alcoholic beverages: Besides the fact already noted of a *positive* price elasticity, the quantile estimation outcome for wine shows a *inversed* U-shape. The elasticity is negative for quantiles smaller than 14% and bigger than 83%, but reaches values greater than +0.35 for the quantiles around 50%. This results could perhaps be explained by the much stronger dispersion of the average price paid for wine and the fact that wine is much more related to social status which may, for example, lead so-called "yuppies" (young urban professionals) or "dinks" (double income, no kids) to buy the more expensive wine whenever available.

A different view of these results is depicted in figure 3.4 which shows estimated conditional 5%-, 50%- and 95%-quantiles of beer consumption with respect to price. The curves in this figure have been obtained by evaluating all $(x, \widehat{Q}_y(\tau))$-pairs (one for each household). Subsequently, the results have been connected via a median spline function, respectively. One can easily recognize the different slopes for different consumption intensities. (The fluctuations at the edges are caused by variation of the other regressors included in the model.)

Since the results so far (which have not been presented besides the price elasticities

Table 3.6: Pseudo continuous variables

hhsize	%	value	income	%	value	age	%	value
one	25.59	1	<= 999	2.36	700	<= 24	1.48	22
two	35.20	2	1000-1249	3.46	1125	25-29	6.37	27
three	18.44	3	1250-1499	4.05	1375	30-34	9.30	32
four	14.95	4	1500-1999	9.86	1750	35-39	9.72	37
five	4.46	5	2000-2499	13.68	2250	40-44	9.57	42
>= six	1.35	6	2500-2999	13.65	2750	45-49	8.29	47
			3000-3499	13.68	3250	50-54	8.64	52
			3500-3999	10.04	3750	55-59	11.32	57
			4000-4499	9.44	4250	60-64	10.12	62
			4500-4999	5.36	4750	65-69	10.68	67
			5000-5499	5.25	5250	>= 70	14.53	77
			>= 5500	9.18	6000			

Table 3.7: Regression results for beer (t-values in brackets)

log. beer quantity	least sq.	5% quant.	25% quant.	50% quant.	75% quant.	95% quant.
constant	-.6420831	-4.088360	-1.797473	-1.183578	.9832792	3.921834
	[-1.88]	[-4.43]	[-3.75]	[-3.19]	[2.41]	[9.04]
log. av. price	-1.902398	-.9192157	-2.105968	-2.367971	-2.122683	-1.529505
	[-23.83]	[-4.87]	[-15.67]	[-23.29]	[-27.72]	[-11.14]
log. income	.3875034	.4626996	.4378815	.5056042	.3160614	.0799454
	[9.39]	[3.45]	[6.89]	[9.83]	[5.94]	[1.28]
household size	.2660790	.2936658	.2946095	.2693489	.2452102	.1659373
	[15.23]	[7.50]	[11.50]	[12.06]	[10.52]	[7.53]
age	.0852777	.0679076	.0820750	.0866123	.0915853	.0818739
	[11.07]	[4.51]	[6.63]	[10.38]	[11.50]	[8.85]
squared age	-.0007800	-.0006507	-.0007563	-.0007851	-.0008163	-.0007303
	[-10.51]	[-4.78]	[-6.25]	[-9.72]	[-10.05]	[-8.09]

in figure 3.3) have not allowed us to estimate income elasticity, we now convert the grouped data back to artificial continuous variables. The details are given in table 3.6. Most importantly, income classes are now related to a certain income value thereby only approximating the variation between groups and disregarding the variation within groups. For example we assign the income of DM 2,750 to all households from the income interval DM 2,500 to 3,000. Moreover, we take the logarithms of these values in the estimations presented below. For age, a similar procedure is adopted which allows us also to include age squared. For the household size, we use just the integers as regressor variables. The estimation results are presented - for some selected quantiles - in tables 3.7 and 3.8. As one can see in figure 3.5, the coefficients for the price elasticity have not changed much compared to the first model. This may serve as an indicator for the suitability of our data manipulations. Moreover, figure 3.6 displays the estimated income elasticities for all 99 quantiles for both wine and beer.

Figure 3.6 shows that income elasticities behave quite differently for beer and wine.

Table 3.8: Regression results for wine (t-values in brackets)

log. wine quantity	least sq.	5% quant.	25% quant.	50% quant.	75% quant.	95% quant.
constant	-3.727660	-1.067146	-3.421165	-4.261983	-3.949232	-3.981429
	[-7.50]	[-2.48]	[-4.15]	[-8.07]	[-8.79]	[-5.32]
log. av. price	.1030630	-.1747103	.1888987	.3705933	.1859693	-.3212998
	[1.92]	[-3.14]	[1.80]	[5.43]	[2.69]	[-4.53]
log. income	.5559961	.1218493	.4245471	.5511245	.6407591	.8672602
	[9.07]	[2.67]	[3.78]	[9.20]	[9.76]	[10.76]
household size	.0405598	.0409576	.0561165	.0373011	.0455761	-.018845
	[1.51]	[1.15]	[1.32]	[1.18]	[1.88]	[-0.37]
age	.0454465	.0068891	.0367241	.0553589	.0597191	.0620024
	[3.96]	[0.61]	[2.14]	[4.72]	[4.10]	[2.97]
squared age	-.0003653	-.0000657	-.0003397	-.0004768	-.0004792	-.0004107
	[-3.31]	[-0.61]	[-1.99]	[-4.40]	[-3.48]	[-1.98]
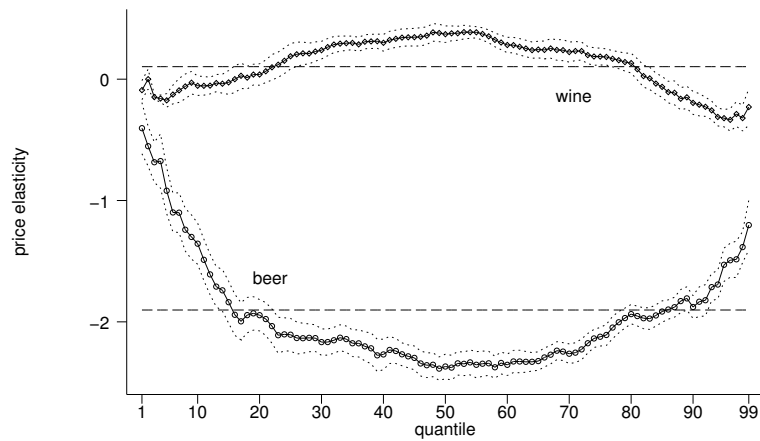
Figure 3.5: New price elasticities
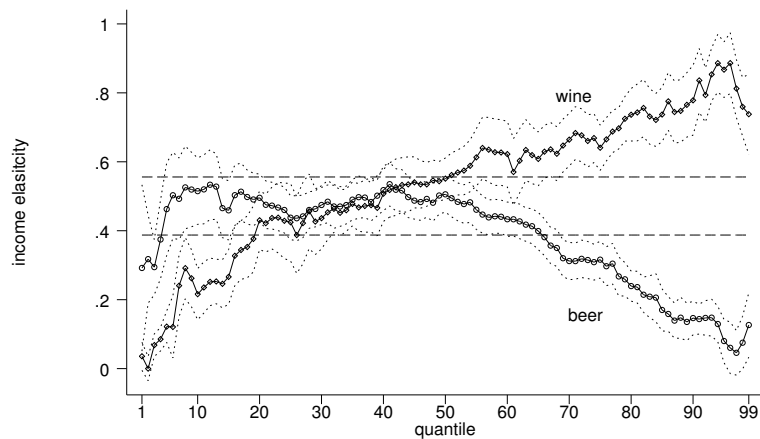
Figure 3.6: Income elasticities
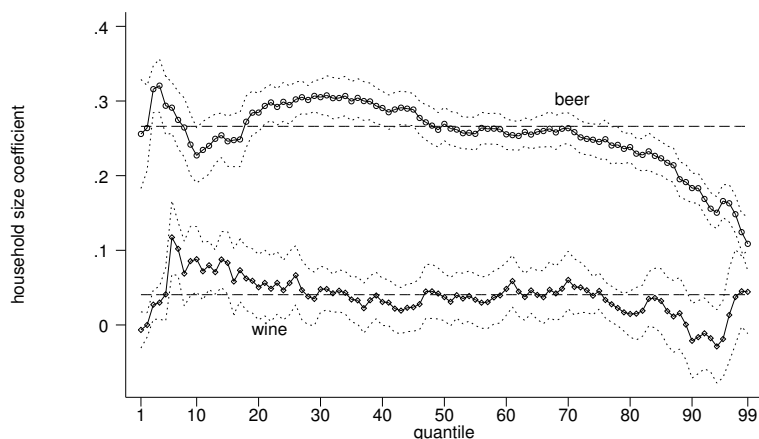
Figure 3.7: Effect of household size

First, we note that the least-squares coefficients are 0.3875 for beer and 0.5560 for wine, respectively. In other words, the positive effect of an income rise on the expected consumption is a bit higher for wine than for beer. Looking at the quantile values, a different picture can be stated. For beer, the elasticity is roughly constant for the quantiles lower than the median, but diminishes at the right tail of the distribution. As far as the impact on wine consumption is concerned, quite the opposite can be observed. The coefficient is negligible small at low quantiles and rises up to 0.8862 at the 96%-quantile. In conclusion, those households consuming only a small amount of beer are more income sensitive than those purchasing a higher quantity, while this relationship is reversed for wine. Again, this could be explained by the association with social status in the case of wine, whereas beer is considered as a every-day good.

Finally, the influence of the household size is depicted in figure 3.7. It can be seen that the quantile coefficients do not differ much from the least-squares results for wine. The same applies to beer, only at large quantiles some smaller values can be observed.
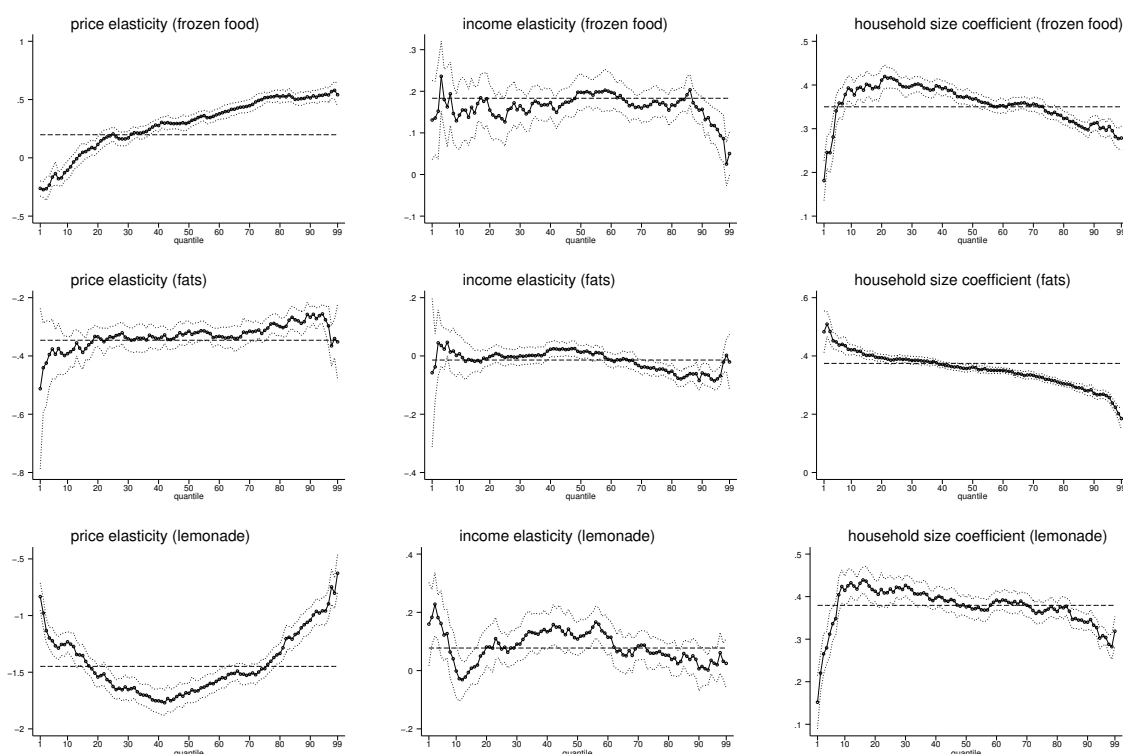
Figure 3.8: Price elasticity, income elasticity and household effect (from left to right) for frozen food, fats and lemonade (from up to down)

3.4.3   Results for other goods

In this subsection, we add some results for other groups of goods. Again, the estimated coefficients are presented in a graphical manner. The method of quantile regression here, too, enables us to reveal more differentiated and detailed results than from standard least-squares estimation. Figures 3.8 and 3.9 shows results for coefficients regarding price, income and household size for the following categories: frozen food, fats, lemonade, animal food and mineral water. The different vertical scales should be noted.

For the groups considered, the following facts can be stated:

- Only for frozen food (first row), the price elasticity moves from negative to positive values for the larger proportion of quantiles. All other goods show normal price reactions over all quantiles.
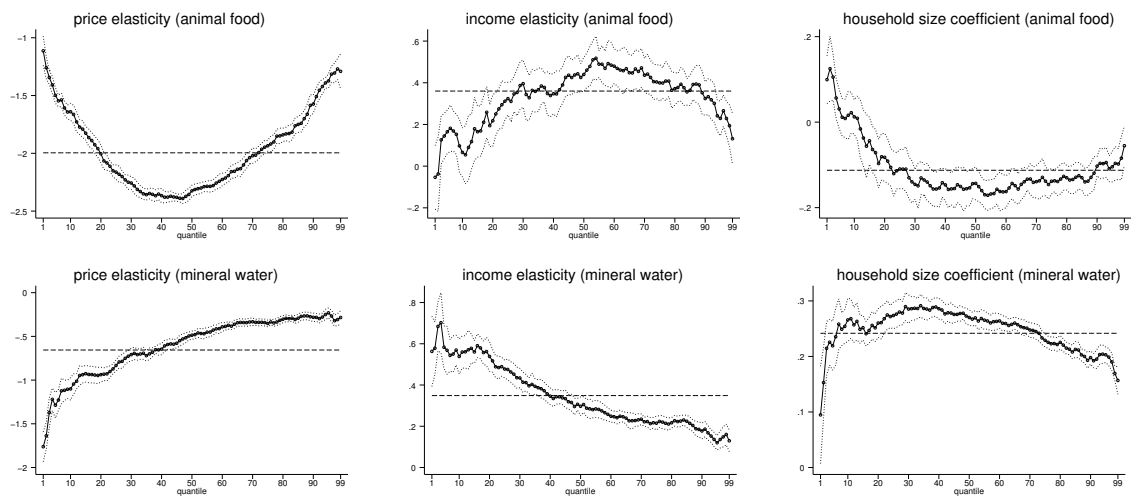
Figure 3.9: Results for animal food and mineral water (compare figure 3.8)

- None of the goods shows a monotonically *decreasing* graph of price elasticities. A monotonically increasing pattern is given for frozen food and mineral water, whereas lemonade and animal food show an U-shape as in the case of beer.

- Income elasticities are - with the exception of fats - always positive and much smaller in size than price elasticities. The income elasticity of 0.8 for wine (see figure 3.6) is by far the greatest value observed in our data set, and the monotonically increasing graph for this good seems to be an exception. Mostly, the income elasticities show now a pronounced pattern. None of the goods is a "luxury" one.

- Household size should have an impact on substitutional processes which cannot be observed from our results. For example, households with children will switch from expensive food to less expensive food. Therefore, negative coefficients should be possible. However, in all cases considered, the impact of household size is positive with the exception of animal food where a negative sign arises for "moderate" and "heavy" consumption. This may be a good example for such substitutional processes: If the household has children and also has animals, then for those households spending a lot of money for animals, this will result in a decrease of consumption of other goods.

3.4.4 Instrumental variables estimation

Our discussion of estimation results so far has not raised the question whether our estimates are biased due to non-exogeneity of regressors. Since we use the income of the household and not total expenditures for the goods considered, we maintain that this explanatory variable should be of no concern regarding biased estimation.[9] However, prices of single purchases are expected to be endogenous indeed since prices will have impact on the decision to buy. Whether this is still a problem for the aggregated prices (see Section 3.3) is an open question. Leaving this aside, the use of unit prices by construction creates negative correlation between this explanatory variable and the error term thereby asking for an instrumental variable approach.[10]

There is not much sample information which could be used. Alternatively the following instrumental variables were employed:

- We split the data set for the full year into two half-year data sets and used (average) prices for the first half as instruments in the quantile regressions based on the second half of the year.

- We used the (discrete) variables describing attitudes of consumers (suspicion towards new products, type of nutrition, price sensitivity, profession, . . .) as instruments.

In both cases the approach of Arias, Hallock, and Sosa-Escudero (2001, Section 5.1) was used. Price elasticities derived from these estimation results are shown in figures 3.10 and 3.11 which should be compared with figure 3.3. It can be observed that in both approaches the price elasticities roughly maintain their shape pointing to the robustness of our results.

---

[9]See Blundell, Pashardes, and Weber (1993, Section C) for a discussion on the treatment of total expenditure being endogenous.

[10]This has been pointed out by one of the editors.
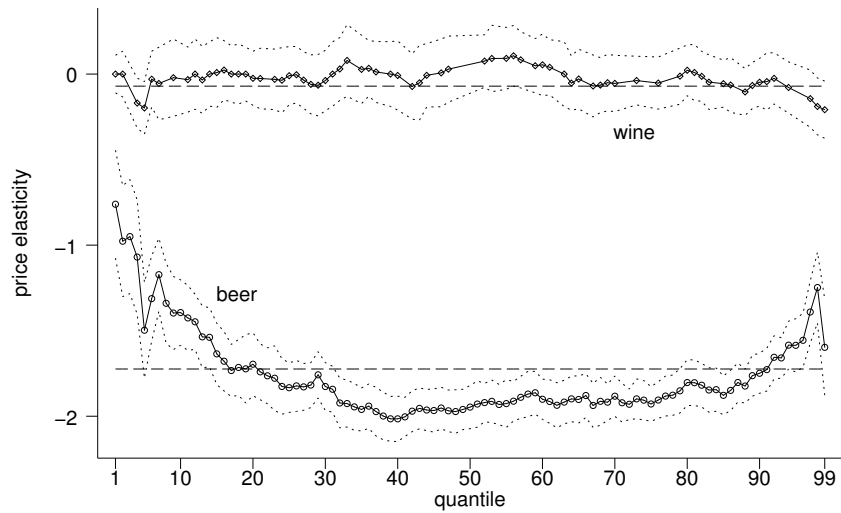
Figure 3.10: Price elasticities from IV estimation (instruments derived from split of the sample)
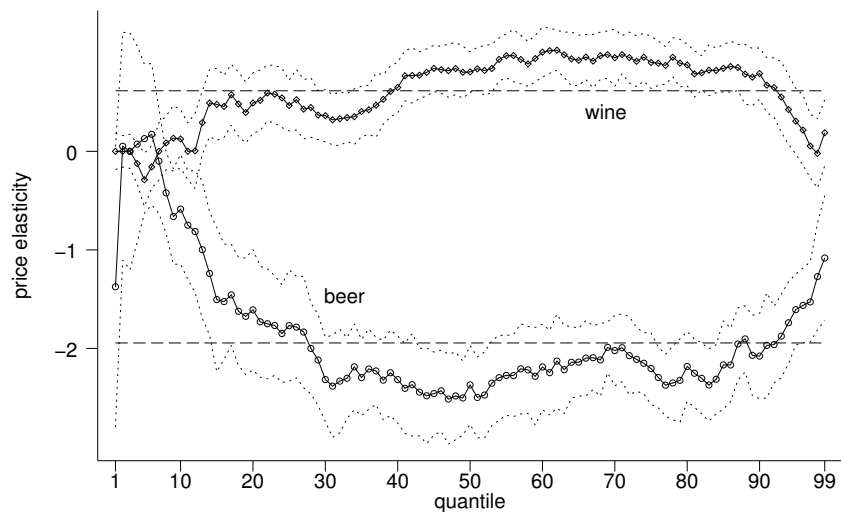
Figure 3.11: Price elasticities from IV estimation (attitudes used as instruments)

3.5 Concluding Remarks

Our study presents empirical results obtained from quantile regressions which show that there is much heterogeneity around the "average" consumer regarding reactions to prices and income. Some typical patterns have been observed for different groups of goods which we try to characterize by different attitudes towards consumption of these goods. For example, beer consumption shows the greatest (negative) price elasticity for "moderate" drinkers, whereas both "light" and "heavy" drinkers are less price sensitive. On the other hand, for wine consumption, price elasticity is *positive* for moderate drinkers and negative for those with very large and very small demand, although the price reactions are much smaller than for beer. We have argued that this particular response pattern may be caused, at least partly, by the fact that wine is much more associated with (higher) social status. Contrary to this, beer is more likely an everyday consumption good (see Table 3.1 for the proportion of non-buyers), and therefore consumers of beer are rather price-sensitive. For "heavy drinkers" however, the problem of addiction makes them less price-sensitive. For quite another reason, "light" beer drinkers also care less about price. Since by construction the unit prices derived from aggregation of the single purchases are correlated with the error term, we have also used instrumental variable estimation employing two different sets of instruments. The results for wine and beer show very similar price responses as obtained in the standard setting and can thus be seen as a justification of our approach.

Chapter 4

Coexceedances in Financial Markets

Abstract

This chapter introduces a new model to analyze financial contagion based on a modified coexceedance measure. We use the quantile regression framework to examine the occurrences and the degrees of coexceedances. Contagion is defined as the crisis-specific coexceedance not explained by the covariates conditional on certain quantiles. Our approach can identify the extent of contagion and also reveal any linear and non-linear linkages between contagion and its determinants. Estimation results for daily stock index returns show that contagion exists and is predictable within and across regions. Furthermore, contagion depends on regional (world) market returns and its volatility and is stronger for extreme negative returns than for extreme positive returns. An analysis of the evolution of coexceedances additionally reveals clusters of extremes. Finally, the computation of conditional densities shows the impact of different influence factors on the entire conditional distribution of coexceedances.

4.1   Introduction

The analysis of the linkages of financial markets is an important topic and of special interest in times of market turmoil or financial crisis. International investors need to assess the benefits of portfolio diversification and policymakers are concerned about the stability of the financial system. In a highly interdependent financial system, a crisis in one country is likely to transmit to other countries which is associated with *contagion*. As a result of the transmission, the linkages between the markets become stronger which (i) diminish diversification benefits of investors and (ii) raise policymakers' concerns about financial stability. The strengthening of market linkages is also the rationale for many empirical tests: if the correlation between the market of the crisis origin and other markets increase, there is contagion. Otherwise, the correlation is constant which is called *interdependence* (see Forbes and Rigobon (2002)). Hence, it is essential to assess the linkages before, during and after a crisis to identify contagion. These linkages can be estimated with Pearson's correlation coefficient (see among others King and Wadhwani (1990), Lee and Kim (1993), Calvo and Reinhart (1996), Baig and Goldfajn (1999), Loretan and English (2000), and Forbes and Rigobon (2002)), with a regression model (Hamao, Masulis, and Ng (1990), Edwards (1998), Bekaert, Harvey, and Ng (2004)) or by a cointegration analysis (Cashin, Kumar, and McDermott (1995)).

A common approach to detect contagion in financial markets[1] is based on the correlation between markets in a period of turbulence compared to a normal situation (see Baig and Goldfajn (1999) and Forbes and Rigobon (2002)). We argue that tests for contagion based on the correlation coefficient are inadequate for the following reasons: First, the correlation coefficient is not an adequate measure to assess market linkages due to its sensitivity to heteroscedasticity (see Boyer, Gibson, and Loretan (1999), Forbes and Rigobon (2002), and Longin and Solnik (2001)), and second, the correlation coefficient is a linear measure which is inappropriate if contagion is an event that is characterized

---

[1]A list of different definitions of contagion is provided by the World Bank (*http://www1.worldbank.org/ contagion/definitions.html*). See also Pericoli and Sbracia (2003).

by nonlinear changes of market association (see Bae, Karolyi, and Stulz (2003)).[2]

Thus, we base our analysis of contagion on another measure to assess the joint movement of financial markets. We use joint exceedances (coexceedances) of two financial market returns below or above a certain threshold to test the hypothesis whether there are increased coexceedances among financial markets in particular periods of market turbulence (i.e. contagion) or not.

We contribute to the literature in various ways: first, we compute coexceedances for every point of time t without arbitrarily categorizing the coexceedances, second, we use the quantile regression (QR subsequently) model of Koenker and Bassett (1978) to analyze the behavior of extreme coexceedances for different regimes of coexceedances, and third, we use the model to detect contagion among financial markets by additionally analyzing certain crisis periods. The advantage of the quantile regression model in comparison to a multinomial logit model used by Bae, Karolyi, and Stulz (2003) is the possibility to analyze not only the occurrence but also the degree of coexceedances. In addition, coexceedances are estimated conditional on the dependence structure which yields more conservative results regarding the detection of contagion. Furthermore, conditional quantile estimates show the evolution of coexceedances over time, and conditional density estimates can reveal the existence of multiple equilibria.

The chapter is organized as follows: section two describes the computation of time-varying coexceedances and introduces the econometric framework for analyzing coexceedances, section three presents the data set used, and section four displays the empirical results. Finally, section five summarizes the main results and concludes.

---

[2]Embrechts, McNeil, and Straumann (2002) summarize the shortcomings of the correlation coefficient and discuss other types of dependence measures.

4.2   Coexceedances

4.2.1   Definition

The term *exceedance* has been introduced by Bae, Karolyi, and Stulz (2003). It is defined as the occurrence of an extreme return (i.e. a return value below (above) a prespecified threshold) of a financial market at a certain point of time t. Bae, Karolyi, and Stulz (2003) use the 5th (95th) quantile of the overall (unconditional) return distribution as the threshold that defines an exceedance. Positive and negative returns are treated separately. The joint occurrence of exceedances in two or more markets at the same point of time is defined as *coexceedance*. Finally, the *number of coexceedances* at time t is determined by the number of countries jointly exceeding their thresholds.

We propose a different approach that does not only specify the existence of coexceedances but also reveals information about their *degree*. In the bivariate case, the coexceedance of two return pairs r_1 and r_2 at time t is defined as follows:

$$Coex_t(r_1, r_2) \quad = \quad \begin{cases} \min(r_{1t}, r_{2t}) & \text{if} \quad r_{1t} > 0 \wedge r_{2t} > 0 \\[2mm] \max(r_{1t}, r_{2t}) & \text{if} \quad r_{1t} < 0 \wedge r_{2t} < 0 \\[2mm] 0 & \text{otherwise} \end{cases} \qquad (4.1)$$

The measure can be interpreted as the value of (extreme) movement that is shared by both markets. It can be shown that the coexceedance $Coex_t$ is related to lower and upper tail dependence (see e.g. Poon, Rockinger, and Tawn (2004) for a recent application) in the following way:

$$Prob(Coex_t \leq a) = Prob(\forall i : r_i \leq a) \qquad (4.2)$$

which is equal to lower tail dependence if the scalar a is sufficiently small or if $u \to 0$ for $a = F_{Coex}^{-1}(u)$.[3]

---

[3]Equation (4.2) can also be written as follows

$$Prob(Coex_t \leq a) = H(\forall i : r_i \leq a) = C(\forall i : F_{r_i}(a)) \qquad (4.3)$$

Since the definition above is based on raw (absolute) values of the market returns, a direct interpretation of the measure is not straightforward. To overcome this problem, both markets are standardized to have zero mean and a variance of one before calculating the coexceedance. This leads to the following intuitive interpretation: if the coexceedance at time t has a value of e.g. -2, both returns are at least two standard deviations below their mean at that point of time.

4.2.2 Estimation framework

Bae, Karolyi, and Stulz (2003) analyze the number of coexceedances (according to their definition) within a multinomial logistic regression framework with the dependent variable defined by the number of countries jointly exceeding their thresholds at the same time. They investigate the influence of exogenous variables like exchange rates, interest rates and volatilities on the number of observed coexceedances. Due to the use of a multinomial logistic regression, they are unable to make any statement on the degree of the examined coexceedances.

In contrast, we propose to estimate coexceedances (as defined in this chapter) within a quantile regression (QR) framework (see Koenker and Bassett (1978) and chapter two for details) in order to also account for the degree of the investigated coexceedances. An additional advantage of the QR model is that no distributional assumptions have to be made as for example in Bae, Karolyi, and Stulz (2003) or applications of Extreme Value Theory (see Longin and Solnik (2001) and Poon, Rockinger, and Tawn (2004)).

The use of the quantile regression model to analyze extreme coexceedances, i.e. extreme negative and positive coexceedances, enables us to consider any values of the lowest or highest coexceedances without prespecifying any distribution or threshold. A simple

---

where H denotes the joint distribution function of the returns r_i, C is the copula function which is equal to the Frechet upper bound (see Nelson (1999) for an introduction to copulas) and F_{r_i} is the marginal distribution function of return r_i.

linear quantile regression equation is given as follows:

$$Coex = X\beta(\tau) + \varepsilon(\tau) \qquad \text{with} \qquad Q_{\varepsilon(\tau)}(\tau|X) = 0 \qquad (4.4)$$

where $Coex$ denotes the $(n \times 1)$ vector of the coexceedances, X is a $(n \times k)$ matrix of k exogenous variables, $\beta(\tau)$ represents a $(k \times 1)$ parameter vector and $\varepsilon(\tau)$ stands for the $(n \times 1)$ error term. It is assumed that the τ-th quantile of the error term conditional on the regressors has value zero. From this specification, it follows that the τ-th conditional quantile of the coexceedances can be expressed as

$$Q_{Coex}(\tau|X) = X\beta(\tau) \qquad (4.5)$$

4.2.3   Analysis of contagion

In the literature, many definitions of *contagion* have been proposed. As an example, Pericoli and Sbracia (2003) present five different concepts. In accordance with most of the latest papers, we speak of contagion as a "structural break in the international propagation mechanism during a crisis period". In contrast, the case of a stable data-generating process with constant or increased variances is labelled *interdependence*.

Most of the early papers dealing with contagion rely on the analysis of the correlation coefficient during crisis periods. However, this approach is only appropriate in the absence of heteroscedasticity, simultaneous equations and omitted variables. Unfortunately, in most empirical applications at least one of these conditions is violated. There are many proposals to circumvent the problems, like a correction of the correlation coefficient (Boyer, Gibson, and Loretan (1999), Forbes and Rigobon (2002)) or covariance-based tests (Rigobon (2003)), but often new problems arise in these cases (see e.g. Billio and Pelizzon (2003)).

Our approach is different and combines the advantages of the Bae, Karolyi, and Stulz (2003) method with a greater flexibility also taking into account the degree of the ex-

amined contagion. Our starting point is the estimation of the following simple quantile regression model:

$$Q_{Coex_t}(\tau) = \beta_0(\tau) + \beta_1(\tau)D_t^{crisis} \tag{4.6}$$

$Q_{Coex_t}(\tau)$ again denotes the τ-th conditional quantile of the coexceedance at time t, $\beta_0(\tau)$ is the constant and $\beta_1(\tau)$ is a parameter estimating the effect of a dummy variable D_t^{crisis} which is one if t is in the crisis period and zero otherwise. It is assumed that the turmoil period is known by institutional information.[4]

How can the estimated parameters $\beta_0(\tau)$ and $\beta_1(\tau)$ be interpreted? The constant $\beta_0(\tau)$ simply states for any value of τ the respective (unconditional) quantile of the coexceedances during the tranquil period. If e.g. $\beta_0(\tau)$ takes the value -2 for $q = 0.05$, this means that in five percent of all cases (number of days), both markets have returns lower than their mean minus two standard deviations, respectively. In ninety-five percent of the cases (days), at least one of the returns is not below its mean minus two standard deviations.

The parameter $\beta_1(\tau)$ reveals information about the behavior of the coexceedance during the crisis period. If, for example, the coefficient $\beta_1(\tau)$ is significantly below zero for small quantiles (e.g. $\tau \in \{0.01, \ldots, 0.05\}$), we can not only argue that the (extreme) negative movement shared by both markets is significantly lower during the turmoil period, but we are also able to make statements about the severeness of the observed contagion. Since the quantile regression model accounts for different regimes of coexceedances, our approach is more conservative than existing definitions and measures of contagion: Contagion is not detected just because values are usually larger in the tails. Furthermore, by taking into consideration all quantiles, our approach can also reveal different structures of dependence[5] and uncover potential non-linearities (see Yu, Lu, and Stander (2003)).

The simple specification given by equation (4.6) neglects any other covariates that po-

---

[4]As we will show later, the concept can be expanded by endogenously determining the crisis period.
[5]Hu (2003) clarifies the separation between degree of dependence and structure of dependence.

tentially have an influence on the structure of the coexceedances. Among others, one could think of the return and the volatility of a regional or global factor, interest rates or exchange rates. As an example, an increased volatility during the crisis period might lead to larger extreme coexceedances, although the underlying data-generating process remains stable. To correct for these influences, we include them into our model and define contagion as the unexplained coexceedances during crisis times.

In other words, model (4.6) can be seen as a benchmark model that provides fundamental information about the (raw) behavior of the coexceedance during the crisis period. A model that does not neglect the influence of a regional or global market and also accounts for any potential persistence of coexceedances is given by

$$Q_{Coex_t}(\tau) = \beta_0(\tau) + \beta_1(\tau)D_t^{crisis} + \beta_2(\tau)r_{Mt} + \beta_3(\tau)\widehat{h}_{Mt} + \beta_4(\tau)Coex_{t-1} \qquad (4.7)$$

where r_{Mt} is the return of a global or regional market index, $\hat{h}_{Mt}$ is its estimated conditional variance and $Coex_{t-1}$ is the lagged coexceedance. If $\beta_1(\tau)$ is significantly smaller than zero even after controlling for the effect of a global or regional stock index, there is evidence of contagion. This concept can be seen in line with Bae, Karolyi, and Stulz (2003) who define contagion as "the fraction of (co-)exceedance events that is not explained by the covariates" and is consistent with the definition given by Bekaert, Harvey, and Ng (2004).

4.3 The Data

We use daily (close-to-close)[6] continuously compounded index returns of eleven Asian stock markets calculated in U.S. dollars[7]: China, Hong Kong, India, Indonesia, Japan, South Korea, Malaysia, Philippines, Singapore, Taiwan and Thailand. Furthermore,

---

[6]We are aware of the potential bias that is introduced by using this type of returns since trading hours are not synchronous.

[7]The data is provided by Morgan Stanley Capital International Inc. (MSCI) and can be retrieved under *www.mscidata.com*.

Table 4.1: Descriptive statistics for the eleven countries and four regional indices utilized in the analysis (1176 observations from 30/04/1997 to 31/10/2001)

market	median	mean	std dev	min	max	skewness	kurtosis	autocorr
China	-0.0013	-0.0014	0.0267	-0.1444	0.1274	0.1333	5.7582	0.1563
Hongkong	0.0000	-0.0004	0.0220	-0.1377	0.1601	0.1767	9.6380	0.0324
India	0.0000	-0.0003	0.0189	-0.0732	0.0782	-0.1556	4.7245	0.0933
Indonesia	-0.0011	-0.0019	0.0448	-0.4306	0.2381	-0.7614	16.3923	0.1307
Japan	-0.0009	-0.0003	0.0165	-0.0716	0.1227	0.4743	6.5383	0.0203
Korea	-0.0002	-0.0002	0.0359	-0.2167	0.2688	0.3418	9.5961	0.1056
Malaysia	-0.0008	-0.0008	0.0308	-0.3695	0.2568	-0.5733	32.5324	0.0868
Philippines	-0.0008	-0.0013	0.0216	-0.1036	0.2119	1.2706	16.0811	0.2181
Singapore	-0.0005	-0.0007	0.0206	-0.1003	0.1552	0.5025	8.7563	0.1483
Taiwan	-0.0009	-0.0008	0.0204	-0.1113	0.0739	-0.0056	5.2528	0.0391
Thailand	-0.0013	-0.0013	0.0294	-0.1489	0.1644	0.6388	7.1727	0.1683
EMF Asia	-0.0009	-0.0009	0.0161	-0.0753	0.0761	0.0421	5.0560	0.2178
EMF La.Am.	0.0004	-0.0003	0.0188	-0.1448	0.1307	-0.4029	11.7116	0.1246
Europe	0.0010	0.0004	0.0143	-0.0680	0.0619	-0.3509	4.7824	0.0589
USA	0.0000	0.0002	0.0127	-0.0697	0.0488	-0.2550	5.8229	-0.0032

four regional stock indices are analyzed: Emerging Markets Free Asia[8], Emerging Markets Free Latin America, Europe and the USA. The indices span a time-period of four and a half years from April, 30th 1997 until October, 31th 2001. The number of observations is $T = 1176$.

Table 4.1 presents several descriptive statistics for the fifteen time series. It can be seen that the mean is negative for Asian countries and Latin America, but positive for Europe and the United States. The skewness exhibits a different behavior among the analyzed markets, whereas all returns are (in some cases considerably) leptokurtic. Except for the USA, all autocorrelations are positive.

Table 4.2 shows the unconditional correlation structure between all variables for the whole time span. One can see that all values are positive, thereby reflecting regional and economic relationships. Table 4.3 lists the corresponding values during the crisis period under the assumption that (i) the Hongkong market is the origin of the crisis (October, 17th until November, 17th 1997) and, alternatively, (ii) the financial market of Thailand is the crisis-breeding element (July, 2nd until November, 17th 1997). In most cases, the correlation rises during the crisis in the Hongkong case, whereas the

---

[8]The MSCI Free indices reflect investable opportunities for global investors by taking into account local market restrictions on share ownership by foreigners.

Table 4.2: Unconditional pairwise correlations between all analyzed markets and regional indices

	CHN	HON	INA	IND	JAP	KOR	MAL	PHI	SIN	TAI	THA	ASI	LAT	EUR	USA
China	1.00														
Hongkong	0.60	1.00													
India	0.18	0.21	1.00												
Indonesia	0.26	0.35	0.10	1.00											
Japan	0.28	0.36	0.15	0.21	1.00										
Korea	0.25	0.29	0.19	0.17	0.26	1.00									
Malaysia	0.27	0.31	0.11	0.33	0.23	0.20	1.00								
Philippines	0.31	0.37	0.14	0.38	0.22	0.21	0.26	1.00							
Singapore	0.44	0.61	0.17	0.46	0.38	0.26	0.39	0.44	1.00						
Taiwan	0.23	0.26	0.12	0.18	0.19	0.20	0.18	0.18	0.28	1.00					
Thailand	0.31	0.38	0.17	0.38	0.25	0.31	0.37	0.40	0.48	0.23	1.00				
EMF Asia	0.49	0.54	0.45	0.47	0.38	0.63	0.56	0.45	0.58	0.60	0.57	1.00			
EMF LA	0.13	0.21	0.12	0.08	0.14	0.18	0.11	0.14	0.22	0.07	0.17	0.21	1.00		
Europe	0.19	0.37	0.15	0.12	0.18	0.21	0.13	0.13	0.30	0.14	0.20	0.28	0.47	1.00	
USA	0.04	0.13	0.05	0.01	0.06	0.10	0.02	0.07	0.13	0.03	0.05	0.08	0.58	0.43	1.00

Table 4.3: Comparision of the correlations during the tranquil and crisis period for Hongkong and Thailand as origins

	CHN	HON	INA	IND	JAP	KOR	MAL	PHI	SIN	TAI	THA	ASI	LAT	EUR	USA
HON	0.60	1.00	0.21	0.35	0.36	0.29	0.31	0.37	0.61	0.26	0.38	0.54	0.21	0.37	0.13
HON Crisis	0.81	1.00	0.10	0.63	0.47	0.18	0.58	0.67	0.79	0.11	0.01	0.50	0.12	0.83	0.05
THA	0.31	0.38	0.17	0.38	0.25	0.31	0.37	0.40	0.48	0.23	1.00	0.57	0.17	0.20	0.05
THA Crisis	-0.03	0.02	0.19	0.24	0.03	0.25	0.28	0.24	0.14	0.16	1.00	0.48	0.02	0.07	-0.08

Thailand values decrease, even leading to two negative outcomes. This result already shows that the definition of the crisis period can be crucial. We will show below that conditional quantile estimates can be used to detect the true crisis periods. Note that the time period for the Hongkong crisis is equal to the one used by Forbes and Rigobon (2002).

In order to clarify our approach, table 4.4 shows the Hongkong and Malaysian returns

Table 4.4: Descriptive statistics for Hongkong, Malaysia and the resulting coexceedances between the two markets. The first four rows indicate the raw values, the last four states the results obtained by standardization of the two time series.

Variable	obs	median	mean	std dev	min	max	skewness	kurtosis	autocorr
Hongkong	1175	0.0000	-0.0004	0.0220	-0.1377	0.1601	0.1767	9.6380	0.0324
Malaysia	1175	-0.0008	-0.0008	0.0308	-0.3695	0.2568	-0.5733	32.5324	0.0868
coexceedance	1175	0.0000	-0.0007	0.0123	-0.0803	0.0888	-0.0388	13.2605	0.0977
coexceedance $\neq 0$	668	-0.0005	-0.0012	0.0163	-0.0803	0.0888	0.0686	7.5787	0.0960
stdrhong	1175	0.0176	0.0000	1.0000	-6.2518	7.3030	0.1767	9.6380	0.0324
stdrmal	1175	0.0001	0.0000	1.0000	-11.9608	8.3595	-0.5733	32.5324	0.0868
coexstd	1175	0.0000	-0.0104	0.4782	-2.5785	4.0573	0.4899	15.3251	0.1162
coexstd $\neq 0$	706	0.0162	-0.0173	0.6169	-2.5785	4.0573	0.4135	9.2316	0.1423
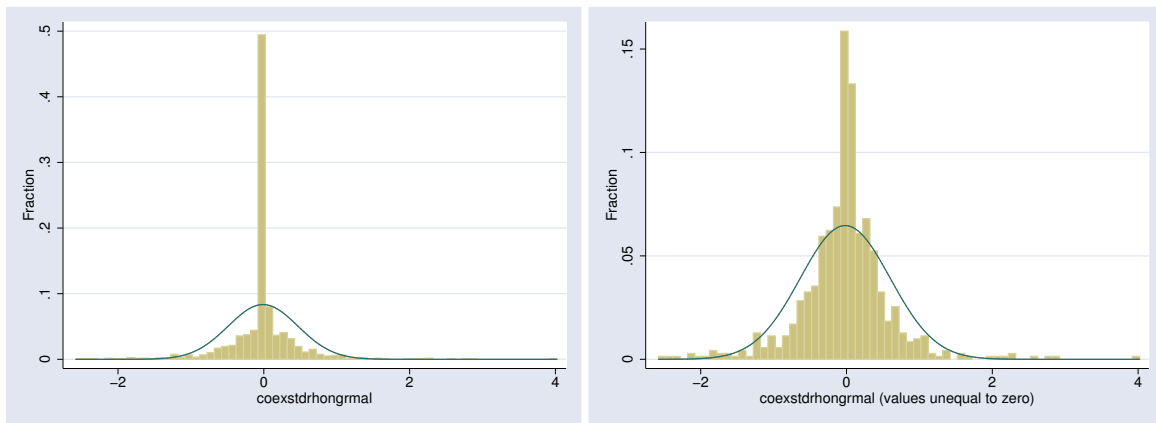
Figure 4.1: Coexceedance between Hongkong and Malaysia. The left figure presents the histogram of the coexceedance obtained from standardized values (as comparison, we superimposed a normal distribution). The right graph shows the distribution with all zero coexceedances excluded.

as well as the calculated coexceedances in exemplary fashion. The upper part of the table contains the unstandardized values, the lower part their standardized analogues. As illustrated in section two, we advocate the use of the standardized values in order to ease the interpretation of the results. The last four columns show that the characteristics of the time series are not changed by the standardization. It can be seen that the coexceedances are somewhat less leptokurtic than the original returns, especially if only the values different from zero are considered. The autocorrelations are slightly higher than those of the original returns. Figure 4.1 plots the histogram of the computed coexceedance. Both graphs (with and without values equal to zero) reveal that the distribution is clearly non-normal.

The question whether joint negative shocks are more common or more pronounced than joint positive shocks can reveal important information and is often analyzed with the correlation coefficient (see e.g. Ang and Chen (2002)). Such an analysis can also be performed with the coexceedance measure by analyzing the percentages of positive and negative coexceedances as well as the skewness of the coexceedances. Results are shown in tables 4.5 and 4.6, respectively. Table 4.5 indicates that joint negative shocks are less frequent than joint positive shocks (the difference seems to be bigger for Hongkong than for Thailand). Table 4.6 shows that joint negative shocks are not

Table 4.5: Percentages of Coexceedances. The table shows the proportions (in percent) of negative and positive coexceedance of Hongkong and Thailand with all other markets, respectively.

	CHN	HON	INA	IND	JAP	KOR	MAL	PHI	SIN	TAI	THA	LAT	EUR	USA
negative (HON)	34.1	xxxx	24.3	25.9	28.2	28.2	28.1	26.6	30.4	27.0	28.3	24.8	28.3	25.8
positive (HON)	38.3	xxxx	31.0	31.3	30.7	32.3	32.1	31.9	35.0	30.9	32.8	31.8	34.4	28.4
negative (THA)	29.9	28.3	25.2	28.2	29.3	30.3	29.9	29.5	31.0	27.8	xxxx	26.2	28.0	25.7
positive (THA)	30.7	32.8	28.6	30.7	28.4	31.0	30.5	31.5	32.1	28.3	xxxx	29.9	30.7	25.0

Table 4.6: Skewness of Coexceedances. The table presents the skewness of the computed coexceedance of Hongkong and Thailand with all other markets, respectively. In each case, the second row takes into account only coexceedances different from zero.

	CHN	HON	INA	IND	JAP	KOR	MAL	PHI	SIN	TAI	THA	LAT	EUR	USA
Hongkong	-0.07	xxxx	-0.51	1.09	-0.29	-0.00	0.49	0.96	0.53	0.16	0.78	-0.65	-0.54	-0.97
HON (coexstd $\neq 0$)	-0.04	xxxx	-0.36	0.81	-0.18	-0.00	0.41	0.74	0.45	0.11	0.62	-0.47	-0.41	-0.70
Thailand	0.45	0.78	0.51	1.42	0.20	0.70	0.20	1.18	0.99	0.08	xxxx	-0.13	-0.56	-0.07
THA (coexstd $\neq 0$)	0.38	0.62	0.43	1.10	0.22	0.59	0.22	0.95	0.81	0.09	xxxx	-0.05	-0.38	0.04

generally more pronounced than joint positive shocks. Results are mixed within Asia, whereas the skewness tends to be negative for coexceedances across regions. These findings are counter to the outcomes in the literature (see e.g. Ang and Chen (2002) and Longin and Solnik (2001)) and may partly be explained by the different properties of the coexceedance measure compared to the correlation coefficient.

4.4 Empirical Results

In this section, we present the empirical results of models (4.6) and (4.7) introduced in section two. First, an analysis of coexceedances of market indices belonging to the same region (in our case the Asian market) is conducted and second, effects across regions are analyzed. For the coexceedances across regions, we also consider the evolution of the conditional quantile estimates and present a (to our knowledge) new concept of conditional density estimates.

Table 4.7: Estimation Results for Hongkong and Malaysia. The table shows the coefficients and estimated t-values (in brackets) for several quantiles along with the least squares outcome. The first part presents the benchmark model only containing a constant and a crisis dummy. The second model also includes a regional index, its volatility and the lagged coexceedance as regressors.

Model	q2	q4	q6	q8	q10	LS	q90	q92	q94	q96	q98
Benchmark Model											
Pseudo-R^2	0.053	0.039	0.030	0.023	0.020	0.004	0.005	0.008	0.010	0.013	0.021
Constant	-1.205***	-0.792***	-0.621***	-0.529***	-0.396***	-0.006	0.374***	0.449***	0.576***	0.712***	1.032***
	(14.33)	(9.74)	(12.52)	(11.94)	(9.89)	(0.44)	(11.78)	(11.77)	(11.00)	(12.83)	(11.67)
Dummy	-1.374***	-1.787***	-1.341**	-1.433**	-1.454**	-0.222**	0.626	0.613	0.485	1.830**	1.510*
	(3.13)	(3.72)	(2.28)	(2.26)	(2.23)	(2.16)	(0.82)	(0.74)	(0.55)	(2.09)	(1.82)
Full Model											
Pseudo-R^2	0.369	0.344	0.312	0.283	0.261	0.363	0.234	0.259	0.286	0.329	0.412
Constant	-0.865***	-0.618***	-0.541***	-0.475***	-0.414***	-0.008	0.374***	0.426***	0.510***	0.648***	0.822***
	(10.93)	(15.22)	(17.96)	(15.52)	(13.32)	(0.73)	(15.70)	(12.58)	(9.92)	(12.17)	(16.96)
Dummy	-0.776*	-0.929**	-0.610	-0.677	-0.712	-0.114	-0.015	0.312	0.226	0.305	-0.004
	(1.90)	(2.53)	(1.47)	(1.57)	(1.52)	(1.35)	(0.04)	(0.81)	(0.67)	(1.03)	(0.01)
EMF Asia	0.365***	0.319***	0.301***	0.304***	0.278***	0.285***	0.269***	0.297***	0.334***	0.366***	0.376***
	(9.33)	(12.13)	(12.04)	(11.44)	(9.71)	(25.15)	(12.26)	(11.99)	(9.38)	(8.33)	(8.58)
Egarch	-0.143*	-0.189***	-0.146***	-0.134***	-0.118***	0.013	0.114***	0.109***	0.123***	0.160***	0.221***
	(1.93)	(4.32)	(4.47)	(4.62)	(4.46)	(1.16)	(4.39)	(3.27)	(2.87)	(2.90)	(4.92)
Coex$_{t-1}$	-0.189	-0.035	-0.028	-0.021	-0.037	0.014	-0.012	0.001	0.080	0.081	0.076
	(1.14)	(0.28)	(0.30)	(0.25)	(0.51)	(0.59)	(0.30)	(0.01)	(1.22)	(0.86)	(0.60)

* indicates that the coefficient is significantly different from zero at the 90%-level (** at the 95% level, *** at the 99% level); The t-values have been calculated by bootstrapping with 2000 replications.

4.4.1 Contagion within regions

Table 4.7 lists the estimation results for the coexceedance between Hongkong and Malaysia for the crisis period assuming that the Hongkong market is the origin of the crisis. The benchmark model displays the regression of the coexceedance on a constant and a crisis dummy (one during October 17th until November 17th 1997). Looking at the latter, it can be stated that the coefficient is highly significant in the negative tail implying larger coexceedances during the crisis period. The estimates of the full model mainly show that the coefficient of the crisis dummy becomes smaller in absolute values and less significant by including the regional market return EMF-Asia, its volatility and the lagged coexceedance. Nevertheless, the dummy remains slightly significant for some of the low quantiles, so we can see evidence of "some" contagion. The regional market return and its volatility have a significant influence on the coexceedance at all reported quantiles and capture a significant portion of the shocks in the lower quantiles.

Figures 4.2, 4.3 and 4.4 plot the regression results (coefficients) of the full model for

99 different quantiles ($\tau \in \{0.01, \ldots, 0.99\}$).[9] The shaded areas represent the 95%-confidence intervals calculated by bootstrapping with 2000 replications.[10] It can be seen that the crisis dummy has negative values in the whole left tail, which are, however, mostly insignificant. The figures also show the considerable influence of the regional market return and its volatility and the rather negligible influence of the lagged coexceedance.

Since all figures also include the least-squares estimates (represented by a solid horizontal line), the additional information that is provided by the quantile regression model in general and in this application in particular is evident: the coefficient estimates are not constant among the quantiles which indicates that the distribution of the error term is not independent from the covariates. In other words, not only the location but also the scale and the shape of the response distribution are affected by the regressors. It is important to stress that the pseudo-R^2 is not comparable with its least-square analogue as it is a local and not a global measure of goodness of fit.[11]

Having presented the outcomes for Hongkong and Malaysia, now the other Asian countries are included into the consideration. Furthermore, Thailand is taken as an alternative source of potential contagion. Since we are mainly interested in the effect of the crisis dummy and to simplify and clarify the analysis, tables 4.8 and 4.9 provide a summary of the crisis dummy coefficients for all analyzed countries.

Table 4.8 shows that for low quantiles, the crisis dummy is significantly negative in the benchmark model (upper part) in most cases. Turning to the full model (lower part), the picture is more ambivalent: for some countries, the coefficient remains significantly negative thus indicating contagion, for other countries its effect is now captured by the other covariates thus signalling interdependence.

---

[9]It has to be noticed that due to the construction of the coexceedances (allocation of value zero for opposite returns) no relevant outcomes are to be expected for the "middle" quantiles (roughly between 30% and 70%).

[10]This number of replications is large enough to guarantee a small variability of the estimated covariance matrix (Buchinsky 1998b).

[11]The measure is calculated as $1 - \hat{V}(\tau)/\tilde{V}(\tau)$ with $\hat{V}(\tau)$ and $\tilde{V}(\tau)$ referring to the unrestricted and restricted quantile regression minimization problems (see Koenker and Machado (1999)).
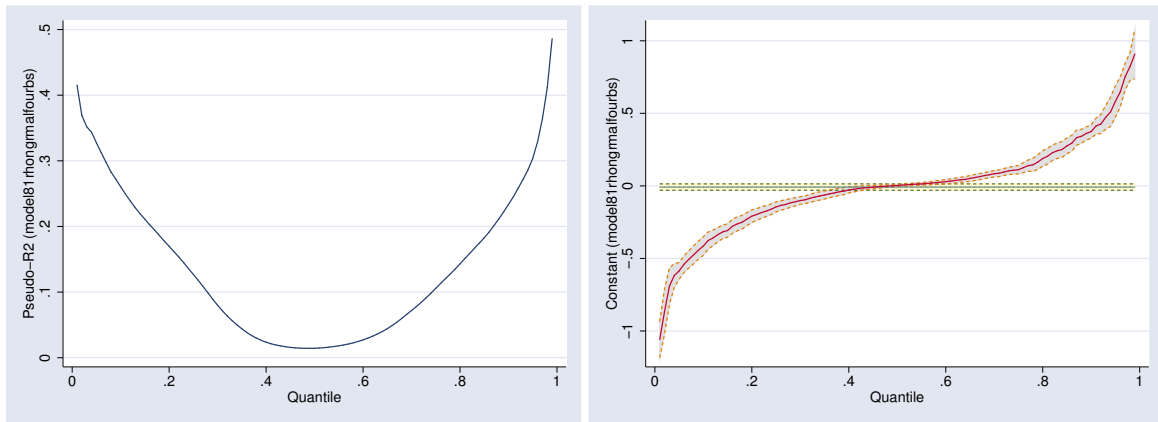
Figure 4.2: Pseudo-R^2 and Constant for the coexceedance between Hongkong and Malaysia. The left figure shows the pseudo-R^2 (see text) for 99 different quantiles ($\tau \in \{0.01, \ldots, 0.99\}$) for the full regression model (see section two). The right graph pictures the constant ($\beta_0(\tau)$) of the model for the same quantiles. The respective values are connected as a solid red line along with the estimated confidence band shaded in grey. The least squares value is included as a horizontal blue solid line.

Figure 4.3: Crisis Dummy and Market Return. The two figures present the coefficients $\beta_1(\tau)$ (crisis dummy) and $\beta_2(\tau)$ (regional market index EMF-Asia) for 99 different quantiles ($\tau \in \{0.01, \ldots, 0.99\}$). Again, the respective values are connected as a solid red line (along with the estimated confidence band shaded in grey) with the least squares result added as a horizontal blue solid line.
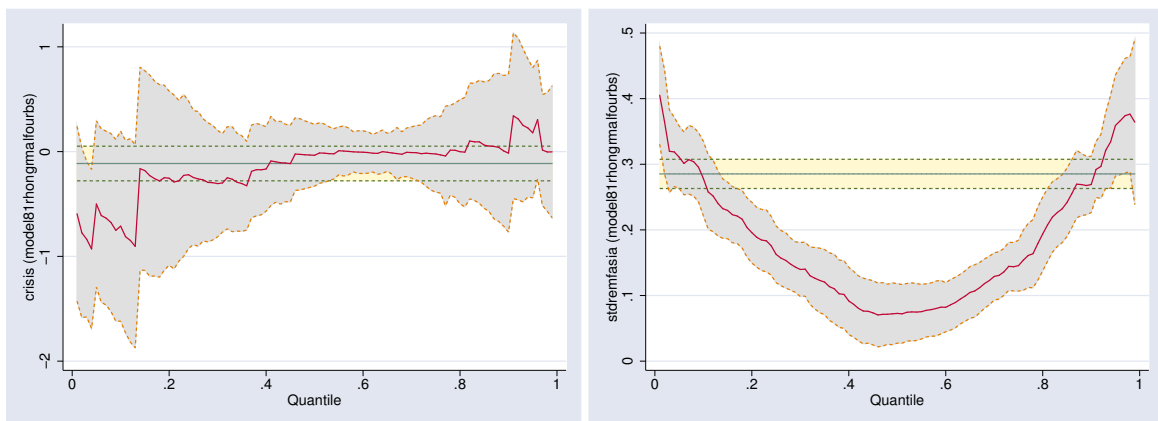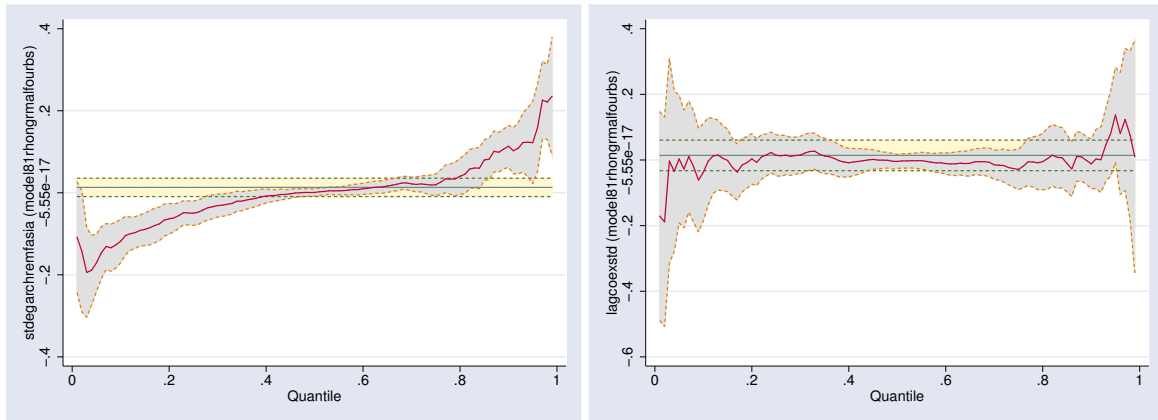
Figure 4.4: Volatility and Lagged Coexceedance. The two graphs picture the coefficients $\beta_3(\tau)$ (estimated volatility of the market index EMF-Asia) and $\beta_4(\tau)$ (one-period lagged value of coexceedance) for $\tau \in \{0.01, \ldots, 0.99\}$ (solid red line). The blue line refers to the LS value.

Table 4.8: Hongkong results. The table presents the coefficient $\beta_1(\tau)$ of the crisis dummy for several conditional quantiles of the coexceedance between Hongkong and ten Asian markets. The upper part refers to the benchmark model only containing a constant and a crisis dummy, the second model also includes a regional index, its volatility and the lagged coexceedance as regressors.

Hongkong	Benchmark Model									
	q2	q4	q6	q8	q10	q90	q92	q94	q96	q98
China	-1.567***	-1.913***	-2.100***	-2.221***	-0.994	0.399	1.097	0.925	1.511**	1.105*
India	-0.077	-0.431	-0.528*	-0.645*	-0.564	-0.486***	-0.525***	-0.625***	-0.638***	-0.862***
Indonesia	-1.998*	-2.378**	-0.244	-0.390	-0.410	-0.297	1.388	1.265	1.772*	1.429*
Japan	-0.794***	-1.092***	-1.012***	-1.111***	-0.938***	-0.503	0.890	0.778	1.056	0.748
Korea	-1.221***	-1.596***	-1.474**	-1.643***	-1.011	-0.387	0.125	-0.017	1.360	0.840
Malaysia	-1.374***	-1.787***	-1.341**	-1.433**	-1.454**	0.626	0.613	0.485	1.830**	1.510*
Philippines	-2.389***	-2.748***	-1.685*	-1.799*	-1.477	0.253	0.449	0.335	2.475**	2.046*
Singapore	-2.355***	-2.734***	-1.728*	-1.907*	-1.091	0.470	1.866	1.742	2.450**	2.087**
Taiwan	-2.252***	-2.570***	-1.812*	-1.927*	-1.019	-0.426	0.450	0.350	0.897	0.578
Thailand	-2.115**	-2.446**	-1.195	-1.303	-0.689	-0.480***	-0.398**	-0.550***	-0.577***	-0.956***
Hongkong	Full Model									
	q2	q4	q6	q8	q10	q90	q92	q94	q96	q98
China	-1.541**	-1.844**	-0.478	-0.516	-0.616	0.028	0.771	0.751	1.508*	1.368*
India	-0.200	-0.185	-0.204	-0.401*	-0.350	0.061	0.504	0.430	0.651*	0.459
Indonesia	-0.246	-0.666	-0.177	-0.205	-0.107	0.127	0.840*	0.619	0.408	-0.130
Japan	-0.713**	-0.985***	-0.630	-0.824**	-0.370	-0.048	0.229	0.145	1.217*	0.870
Korea	-0.906	-1.132*	-0.244	-0.362	-0.477	0.029	0.086	0.036	-0.075	-0.205*
Malaysia	-0.776*	-0.929**	-0.610	-0.677	-0.712	-0.015	0.312	0.226	0.305	-0.004
Philippines	-1.082***	-1.309***	-1.372***	-1.462**	-1.277	0.229	0.377	0.261	1.876**	1.951**
Singapore	-0.966	-1.447**	-0.969	-1.221*	-0.467	0.182	0.132	-0.011	0.938	0.522
Taiwan	-0.946**	-0.966**	-0.891	-0.921	-0.407	0.076	-0.012	-0.052	0.380	0.147
Thailand	-0.404	-0.911**	-0.250	-0.418	-0.306	-0.086	-0.015	-0.106	-0.093	-0.188

* indicates that the coefficient is significantly different from zero at the 90%-level (** at the 95% level, *** at the 99% level); The t-values were calculated by bootstrapping with 2000 repetitions.

Table 4.9: Thailand results. The table presents the coefficient $\beta_1(\tau)$ of the crisis dummy for several quantiles of the coexceedance between Thailand and ten Asian markets. The upper part refers to the benchmark model only containing a constant and a crisis dummy, the second model also includes a regional index, its volatility and the lagged coexceedance as regressors.

	Benchmark Model									
Thailand	q2	q4	q6	q8	q10	q90	q92	q94	q96	q98
China	-0,557	-0,224	-0,121	-0,24	-0,114	-0,053	0,01	-0,115	-0,148	0,065
Hongkong	-0.646	-0.266	-0.302	-0.134	-0.163	-0.199	-0.172	-0.287	-0.050	-0.096
India	-0.238	-0.380	-0.197	-0.218	-0.150	-0.113	-0.023	-0.079	0.177	-0.126
Indonesia	-0.341	-0.247	-0.302	-0.254	-0.271*	-0.104	-0.190	-0.251	-0.298	-0.425
Japan	-0.106	-0.409**	-0.419*	-0.231	-0.230	-0.148	-0.179	-0.131	-0.281	-0.497
Korea	-0.340	-0.306	0.015	0.076	0.010	-0.312***	-0.347***	-0.429**	-0.407	-0.633
Malaysia	-1.027*	-0.319	-0.279	-0.350*	-0.338**	-0.105	-0.144	-0.069	-0.111	0.867
Philippines	-0.669	-0.474	-0.474**	-0.552***	-0.555***	0.106	0.035	0.050	-0.071	-0.273
Singapore	-0.369	0.032	0.047	-0.053	-0.083	-0.219	-0.140	-0.264	-0.276	-0.116
Taiwan	-0.757	-0.301	-0.343	-0.168	-0.247	0.157	0.110	0.024	-0.110	-0.090
	Full Model									
Thailand	q2	q4	q6	q8	q10	q90	q92	q94	q96	q98
China	-0.464**	-0,29	-0,098	-0,087	-0,102	0,085	0,151	0.248*	0,256	0.737*
Hongkong	-0.333	-0.060	-0.101	-0.031	-0.017	-0.048	0.023	-0.005	0.139	0.267
India	-0.177	-0.197	-0.131	-0.124	-0.184*	0.113	0.160	0.291*	0.435**	0.478*
Indonesia	-0.459	-0.013	-0.090	-0.058	-0.057	0.008	-0.013	-0.059	0.002	-0.084
Japan	-0.111	-0.107	-0.151	-0.110	-0.130	-0.033	-0.042	-0.058	0.339	0.312
Korea	-0.011	-0.149	-0.227*	-0.132	-0.015	0.038	0.009	0.065	-0.007	-0.011
Malaysia	0.006	-0.128	-0.122	-0.123	-0.116	-0.035	-0.057	-0.050	-0.021	0.349
Philippines	-0.509	-0.304	-0.323**	-0.260*	-0.190	0.044	0.108	0.201*	0.034	-0.025
Singapore	0.144	0.020	0.002	-0.078	-0.066	-0.017	-0.050	-0.014	-0.050	-0.168
Taiwan	-0.247	-0.232	-0.274**	-0.210*	-0.143	0.090	0.211**	0.157*	0.147	0.009

* indicates that the coefficient is significantly different from zero at the 90%-level (** at the 95% level, *** at the 99% level); The t-values were calculated by bootstrapping with 2000 repetitions.

The upper quantile coefficients have mostly positive values (benchmark and full model) indicating that crisis periods can sometimes also be characterized by extreme positive shocks. An interesting exception are India and Thailand with significantly *negative* values in the benchmark model which completely disappear in the full model. It is evident that the estimated degree of contagion varies depending on the quantile. This is not surprising since contagion is measured conditional on the regime of coexceedances. Note that usually only a certain quantile (e.g. five percent) is reported which would hide the variation observed here. The varying degree of contagion is further analyzed below in subsection 4.2.1 where the evolution of the coexceedances is considered.

The results for the crisis assumed to be associated with occurrences in Thailand are in general less pronounced. The insignificance of the coefficient estimates can be partly explained by the longer crisis period compared to the one assumed in the Hongkong case.

Table 4.10: Coexceedances across regions. The upper part of table presents the crisis dummy coefficient $\beta_1(\tau)$ of the full regression model for several conditional quantiles of the coexceedance between Hongkong and three regional indices. The lower part assumes Thailand to be the crisis breeding element.

	q2	q4	q6	q8	q10	q90	q92	q94	q96	q98
Hongkong										
Latin A.	-1.298**	-1.446**	-0.420	-0.525	-0.602	0.114	0.668	0.574	1.146**	0.738
Europe	-2.347**	-2.563**	-0,927	-1,09	-0,422	0,254	0,405	0,261	2.079*	1,374
USA	0.043	-0.231	-0.237	-0.278	-0.346*	0.315	0.410	0.373	0.638**	0.375
Thailand										
Latin A.	0.016	-0.179	-0.196**	-0.209**	-0.187*	0.102	0.122	0.188	0.169	0.162
Europe	-0.802	-0.031	-0.038	-0.025	-0.092	0.004	0.049	0.188	0.164	0.051
USA	-0.135	-0.000	-0.065	-0.106*	-0.060	0.115	0.080	0.213*	0.195	0.148

4.4.2 Contagion across regions

In this subsection, we want to answer the question whether contagion can also be found across different regions. Therefore, we calculate coexceedances between the Hongkong market as the assumed crisis origin and several regional MSCI indices, namely the United States, Europe and Latin America. We also include Thailand as an alternative crisis-breeding element. In this specification, we use the return and the volatility of the MSCI World index to account for common shocks. Table 4.10 lists the effect of the crisis dummy in the full model for both crisis periods.

The results show that there are larger coexceedances of Hongkong with Latin America and Europe at the lowest quantiles, which indicates evidence of contagion. In contrast, the US market is not strictly affected by shocks originating in the crisis period in Hongkong. Only the 10% quantile slightly indicates contagion, whereas the 2% quantile even has an (insignificant) positive value. Hence, for the assumed crisis period, our result can be seen in line with Bae, Karolyi, and Stulz (2003) who conclude that the US market is insulated from Asian markets. Table 4.10 displays an asymmetry between negative coexceedances and positive coexceedances for the Hongkong crisis, especially for Europe. For the Thailand crisis, there is weaker evidence of any increased coexceedances.

Comparing our results with the findings of Forbes and Rigobon (2002) reveals the fol-

lowing differences: Forbes and Rigobon (2002) find contagion for the Hong Kong crisis in Indonesia, Korea and the Philippines based on a correlation coefficient not corrected for heteroscedasticity. Using the adjusted (corrected for heteroscedasticity) correlation coefficient, no contagion is found in any analyzed Asian market. Although we also correct for high volatility and account for different regimes of coexceedances (which could be interpreted as a correction for heteroscedasticity), we still find contagion in some cases. Our outcomes are thus similar to Corsetti, Pericoli, and Sbracia (2003) who speak of "some contagion, some interdependence".

4.4.2.1 Evolution of coexceedances

In this part, we analyze the evolution of the estimated conditional coexceedances between the Hong Kong market and two regional indices (Europe and USA). The analysis of the coexceedances in a time-varying context (see Chan-Lau, Mathieson, and Yao (2002) for a similar study) provides information on the questions whether extreme joint market movements have increased in recent years, whether negative movements are more pronounced than positive movements and whether the volatility of such movements has increased or stayed rather constant. We present the quantile estimate of the conditional 2-, 50- and 98-percent quantiles of the two coexceedances. Since we use the quantile regression model, our approach is clearly different to the Chan-Lau, Mathieson, and Yao (2002) method that computes moving averages of coexceedances defined as in Bae, Karolyi, and Stulz (2003) with a prespecified threshold.

The plots of the estimated conditional 2-, 50- and 98-percent quantiles along with the realized values are given in figure 4.5 for Hongkong with Europe and the United States. The assumed crisis interval is marked by the grey shaded area. The figures clearly show the differences between the two examined cases. For Europe, there is evidence of contagion expressed by the significantly larger estimated coexceedances during the assumed crisis period. For the United States, the coexceedances are also larger than in tranquil periods, but not larger than in other clusters (periods) of extremes, e.g. at
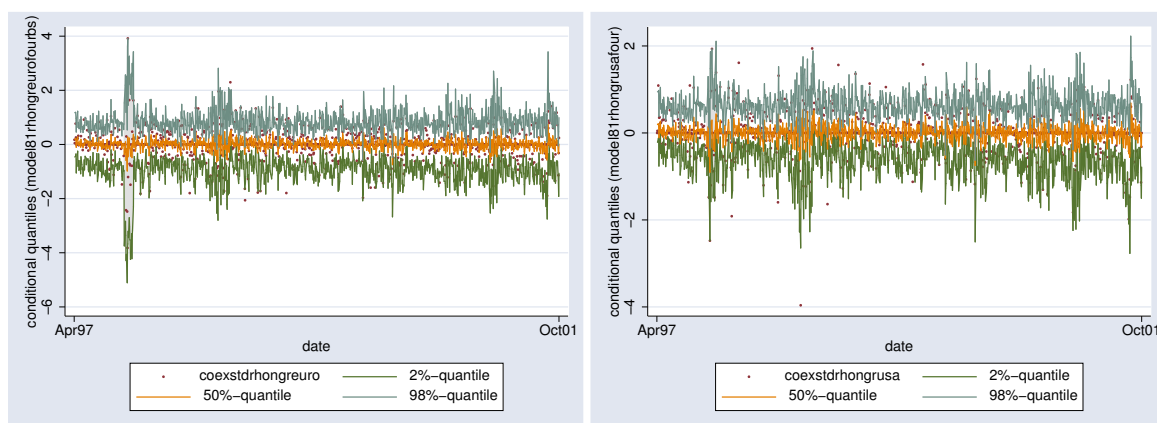
Figure 4.5: Evolution of coexceedances. The left figure shows the computed time-varying coexceedance between Hongkong and Europe as a red dot for every point of time t of the sample. Furthermore, the estimated conditional 2%-, 50%- and 98%-quantiles are superimposed as green, orange and blue lines, respectively. Finally, the underlying crisis period is shaded in grey (in this case: October, 17th until November, 17th 1997). The right graph shows Hongkong and the United States. Please note the different vertical scales.

the end of the sample. Furthermore, there seems to be a slight positive trend of more frequent extreme market movements.

The plots of the conditional quantile estimates for the Thailand crisis period are given in figure 4.6. Both figures show that the crisis period could also be detected endogenously. Even without a crisis dummy in our model, clusters of extreme coexceedances could be uncovered. This could then serve as a justification for the chosen crisis period. This is an important feature of our method and shows an alternative to the crisis detection through outliers proposed by Favero and Giavazzi (2002).

4.4.2.2　Conditional densities

In this part, we present conditional density estimates for the coexceedances of the Hongkong market with two regional stock indices as chosen above. We focus on the Hongkong crisis period since the Thailand crisis interval was shown to be too long (see previous subsection). The conditional densities are calculated by fixing the value of one independent variable to several specific numbers and subsequently computing the
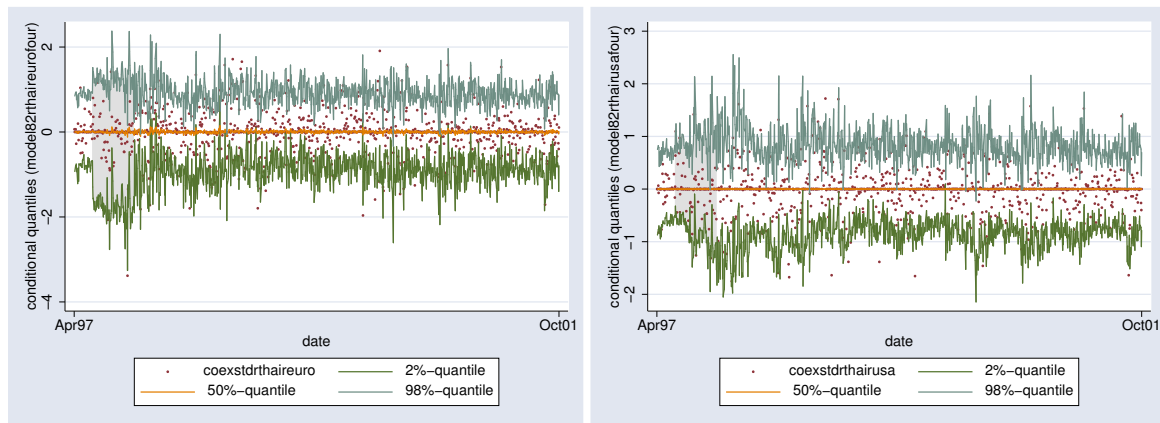
Figure 4.6: Evolution of coexceedances. The figures show the estimated conditional coexceedances between Thailand and Europe and the United States, respectively (compare figure 4.5). The crisis period shaded in grey lasts from July, 2nd until November, 17th 1997 (see text).

density of the coexceedance conditional on these numbers. The according values of the other covariates are calculated by an auxiliary regression (see chapter two for details).

These densities can provide important additional information since they show the distribution of the coexceedance given a certain value of one covariate. This can help obtaining more information about the sources that lead to the occurrences of coexceedances. Figures 4.7 to 4.10 show the density estimates for two pairs of coexceedances (Hong Kong with Europe and USA, respectively) conditional on the four regressors of our QR model. The figures show that the density of the coexceedances in the crisis period ($D^{crisis} = 1$) has a lower mean and a higher variance than the density in the non-crisis period ($D^{crisis} = 0$). There is also some hint of multimodality of the densities during the crisis period which slightly points to the existence of multiple equilibria (see Calvo and Mendoza (2000) and Kodres and Pritsker (2002)). However, the evidence is rather weak.

The figures further show that low (high) values of the world factor lead to lower (higher) expected values of the coexceedances while the variance of the distribution is relatively unaffected. In contrast, the volatility of the world factor affects the variance of the coexceedance but not the mean of the conditional values: large (small) values are asso-
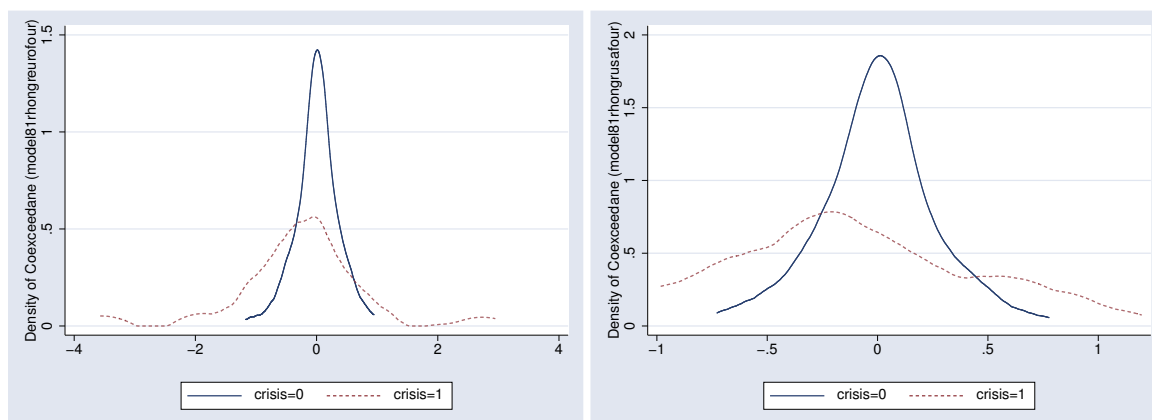
Figure 4.7: Conditional density estimations. The left graph pictures the estimated distribution of the coexceedance between Hongkong and Europe conditional on the value of the crisis dummy. The solid blue line refers to the tranquil period, the short-dashed red line pictures the crisis case. The densities were calculated by applying a kernel density estimation on 99 quantile regression coefficients computed in both cases (see text). The right figure shows the same for the coexceedance between Hongkong and USA.
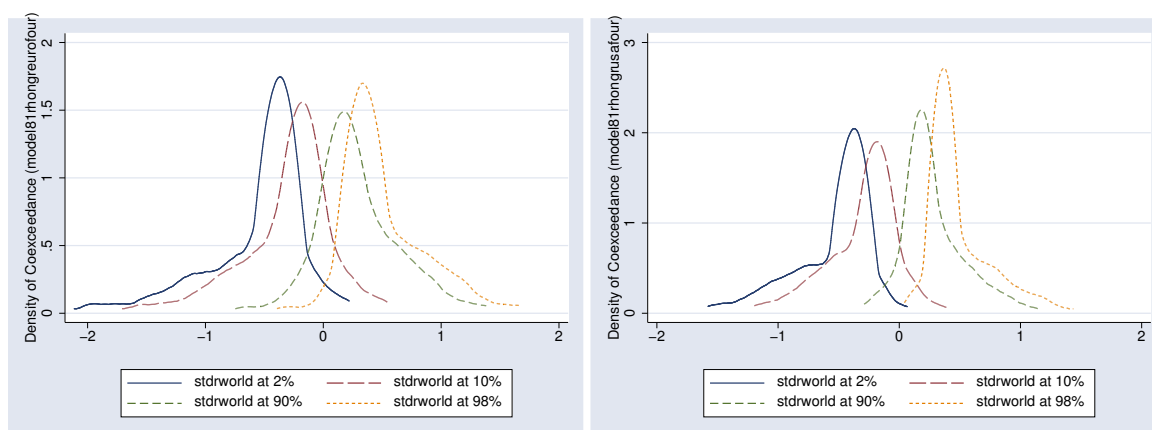

Figure 4.8: Conditional densities. The left figure shows the estimated distribution of the coexceedance between Hongkong and Europe conditional on four different values of the market return (to be precise: the unconditional 2%- (solid blue line), 10%- (long-dashed red line), 90%- (dashed green line) and 98%-quantile (short-dashed orange line) of the MSCI world index). The right graph examines the coexceedance between Hongkong and USA.
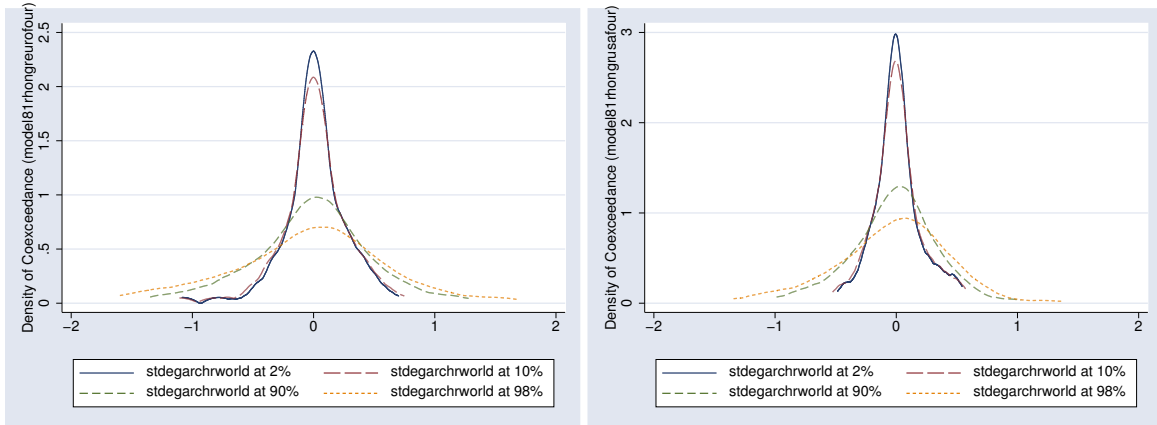
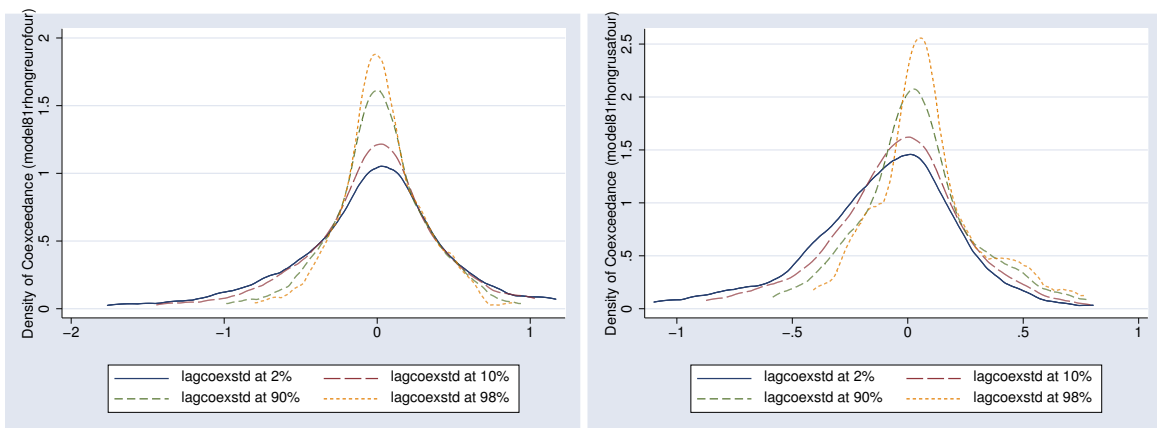Figure 4.9: Density estimation conditional on market volatility (compare figure 4.8)

Figure 4.10: Densities conditional on lagged coexceedance (compare figure 4.8)

ciated with a higher (lower) volatility of the coexceedance. Interestingly, the impact of
the lagged coexceedance is quite similar: negative (positive) lagged coexceedances in-
duce a higher (lower) volatility, whereas the estimated conditional mean remains rather
constant.

4.5   Concluding Remarks

In this chapter, we have proposed a new approach to compute and analyze coexcee-
dances. Time-varying coexceedances and their analysis with the quantile regression
model have various advantages: First, we do not need to count the coexceedances or
prespecify any threshold, second, no distributional assumptions have to be made about
the coexceedances or the underlying returns, third, the approach can estimate linear
and non-linear linkages and is able to correct for different states of shock magnitudes,
and fourth, the inclusion of a dummy variable into the quantile regression model can
directly quantify the degree of contagion. Furthermore, the analysis of the evolution
of extreme coexceedances and the estimation of conditional densities allow use to gain
additional information on the underlying processes, which would remain uncovered if
current methods were employed.

Applying our approach to empirical data, we obtain mixed results within regions: in
some cases we find contagion, in other cases interdependence. This finding also holds
for coexceedances across regions, where we find contagion from Asia to Latin America
and Europe but not to the United States. We have shown that the quantile regression
offers a variety of possibilities to analyze extreme market behavior. Future studies on
conditional quantiles and density estimates could further exploit these new prospects.

Chapter 5

Determinants of Surface Ozone Concentration

Abstract

This chapter proposes the use of the conditional quantile regression approach for the interpretation of the nonlinear relationships between daily maximum 1-hr ozone concentrations and both meteorological and persistence information. When applied to nine years (1992-1999) of data from four monitoring sites in Athens, quantile regression results show that the contributions of the explanatory variables to the conditional distribution of the ozone concentrations vary significantly at different ozone regimes. This evidence of heterogeneity in the ozone values is hidden in an ordinary least square regression that is confined to providing a single central tendency measure. Furthermore, the utilization of an 'amalgated' quantile regression model leads to a significantly improved goodness of fit at all sites. Finally, computation of conditional ozone densities through a simple quantile regression model allows the estimation of complete density distributions that can be used for forecasting next day's ozone concentrations under an uncertainty framework.

113

5.1 Introduction

The study of the impact of meteorology and persistence on surface ozone has attracted
considerable interest in the literature (see e.g. Comrie and Yarnal (1992), Davies, Kelly,
Low, and Pierce (1992), or Derwent, Simmonds, Seuring, and Dimmer (1998)). Statisti-
cal models estimating daily maximum ozone concentrations are numerous and may be
divided into four broad areas: regression-based modelling (Feister and Balzer (1991),
Bloomfield, Royle, Steinberg, and Yang (1996), Hubbard and Cobourn (1998)), extreme
value approaches (Chock and Sluchak (1986), Smith (1989)), neural networks (Can-
non and Lord (2000), Cobourn, Dolcine, French, and Hubbard (2000), Comrie (1997))
and space-time models (Rao, Zalewsky, and Zurbenko (1995), Rao, Zurbenko, Neagu,
Porter, Ku, and Henry (1997)). The main objectives of such models are to obtain ozone
forecasts, investigate and estimate ozone time trends (e.g. Gardner and Dorling (2000)),
increase scientific understanding of the underlying mechanisms in the ozone formation,
and determine potential ozone-related health effects. A model comparison of 15 differ-
ent statistical techniques for ozone forecasting is thoroughly discussed by Schlink, Dor-
ling, Pelikan, Nunnari, Cawley, Junninen, Greig, Foxall, Eben, Chatterton, Vondracek,
Richter, Dostal, Bertucco, Kolehmainen, and Doyle (2003).

Conditional quantile regression was introduced by Koenker and Bassett (1978) and is
gradually emerging as a comprehensive approach to the statistical analysis of linear
and nonlinear models. Quantile regression models have been used in a broad range of
application settings, such as in paediatric medicine (Royston and Wright (1998), Cole
and Green (1992), Cole, Freeman, and Preece (1998), Gasser, Ziegler, Seifert, Prader,
Molinari, and Largo (1994), Heagerty and Pepe (1999)), labor economics (e.g. Buchin-
sky (1994)), financial market analysis (e.g. Engle and Manganelli (2002)) and many
other fields (see Koenker and Hallock (2001)). For a recent review of some typical ap-
plications of quantile regression, the reader is referred to Yu, Lu, and Stander (2003).
By extending the exclusive focus of the least-squares based methods on the estimation
of conditional mean functions with a more general technique for estimating families

of conditional quantile functions, quantile regression is capable of greatly expanding the flexibility of both parametric and nonparametric regression methods. It allows the examination of the entire distribution of the variable of interest rather than a single measure of the central tendency of its distribution.

Additional advantages accruing from using quantile regression models are their flexibility to allow for the covariates to have different impacts at different points of the distribution and the robustness to departures from normality and skewed tails (Mata and Machado (1996)). These latter features are often observed in environmental variables and ozone in particular. Finally, quantile regression can provide information about any linear or nonlinear relationships between the dependent variable and the explanatory variables without an a priori knowledge of the type of (potential) nonlinearities. Despite its advantages and wide range of applications, conditional quantile regression has only been used as median regression in an ozone forecasting study (Schlink, Dorling, Pelikan, Nunnari, Cawley, Junninen, Greig, Foxall, Eben, Chatterton, Vondracek, Richter, Dostal, Bertucco, Kolehmainen, and Doyle (2003)). Yet, with respect to air quality issues, models for mean (or even median) concentration levels may be less relevant from a health standpoint than respective models for upper quantiles representing more extreme concentration levels.

To this end, the purpose of this chapter is to explore the potential of the conditional quantile regression in elucidating the meteorological and persistence influence on the different regimes of daily maximum ozone concentrations recorded in an urban environment susceptible to photochemical pollution such as the Athens metropolitan area.

5.2 The Data

The dataset consists of ozone concentrations recorded at four stations in the Athens basin and was provided by PERPA, a branch of the Hellenic Ministry of Environment, City Planning and Public Works. The Athens basin topography forces two main wind

regimes: the first from the northeast and the second from the southwest. The three residential suburban sites (Liossia, Maroussi and Likovrissi) are located northeast, downwind of the prevailing south-southwesterly winds, and do therefore exhibit the highest ozone concentrations in the basin. The urban site (Smirni) is located upwind, close to the sea front and records lower ozone concentrations. For Smirni, Liossia and Maroussi, an eight-year (1992-1999) data set was used. At Likovrissi, a six-year data set (1994-1999) was available. The analysis focuses on the daily maximum 1-hr ozone concentrations recorded during the 'ozone season' of 01 April through 31 October.

Based on an earlier case study for Athens (Chaloulakou et al., 2003), the meteorological variables considered are: (1) nocturnal wind speed (02:00), (2) morning wind speed (07:00-10:00), (3) afternoon wind speed (13:00-14:00), (4) solar radiation (10:00-14:00), (5) relative humidity (10:00-13:00), (6) upper air temperature at 850 hPa (14:00), (7) change from the previous day of upper air temperature at 850 hPa, (8) range of surface daily temperature ($T_{max} - T_{min}$) and (9) surface maximum temperature. Similar meteorological variables were found to be significant in describing the variation in the daily maximum 1-hr ozone concentrations by several authors (Burrows, Benjamin, Beauchamp, Lord, McCollor, and Thomson (1995), Cannon and Lord (2000), Hubbard and Cobourn (1998), Comrie (1997)). Motivated by the work of Bloomfield, Royle, Steinberg, and Yang (1996) and Davis, Eder, Nychka, and Yang (1998), two perpendicular wind vectors for each type of wind speed (nocturnal, morning and afternoon) are included. The surface meteorological measurements for the period 1992-1999 were obtained from the National Observatory of Athens. The upper air data (i.e., the temperature at 850 hPa) were obtained from the monitoring station at the Athens airport.

Table 5.1 and figure 5.1 presents the histograms of the daily maximum 1-hr ozone concentrations of the datasets at the four sites. The ozone distribution at the urban site Smirni is characterized by much lower values than those at the suburban sites. Furthermore, the ozone concentrations at Smirni more or less resemble a normal distribution, which is clearly not the case for the other stations that exhibit much fatter tails

Table 5.1: Descriptive statistics for the daily one-hour maximum ozone concentrations ($\mu g/m^2$) at the four monitoring stations in the Athens basin

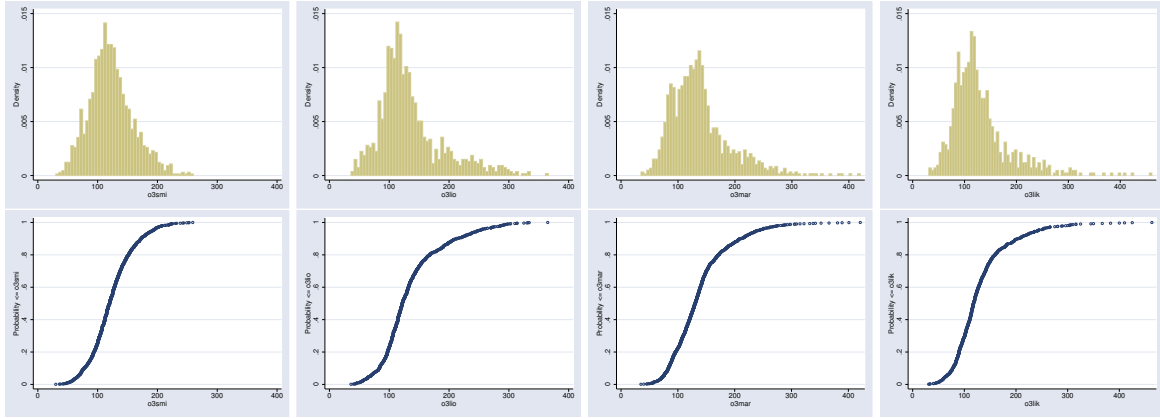| station	obs.	median	mean	std dev	min	max	skewness	kurtosis	autocorr
smirni	1301	120	123.29	36.67	30	259	0.5207	3.4149	0.6125
liossia	1069	122	135.50	53.47	36	365	1.2411	4.5765	0.5124
maroussi	1177	131	138.87	52.51	35	381	1.3534	5.8904	0.5550
likovrissi	839	118	130.65	56.21	31	463	1.8011	7.9693	0.5103

Figure 5.1: Histograms and cumulative distributions of the ozone concentrations at the four monitoring stations (Smirni, Liossia, Maroussi and Likovrissi, respectively)

and are more skewed. In addition, the autocorrelation of ozone concentration at Smirni is higher (0.613) than the autocorrelation at the suburban sites (values slightly greater than 0.5). In our analysis, we will also, thereafter, consider lagged ozone information (i.e., previous day's maximum 1-hr ozone concentration) to account for this persistence in ozone.

5.3 The Quantile Regression Model

A linear quantile regression model (Koenker and Bassett (1978)) assumes that the regressand y (in our case the daily 1-hr maximum ozone concentration) is linearly dependent on K explanatory variables. Furthermore, the τ-th quantile of the error term $\varepsilon_t(\tau)$

conditional on the regressors is assumed to be zero:

$$y_t = \beta_0(\tau) + \sum_{k=1}^{K} \beta_k(\tau) x_{tk} + \varepsilon_t(\tau) \qquad \text{with} \qquad Q_{\varepsilon_t(\tau)}(\tau | x_{t1}, \ldots, x_{tK}) = 0 \qquad (5.1)$$

>From this specification, it follows that the τ-th conditional quantile of y can be written as

$$Q_{y_t}(\tau | x_{t1}, \ldots, x_{tK}) = \beta_0(\tau) + \sum_{k=1}^{K} \beta_k(\tau) x_{tk} \qquad (5.2)$$

In a least squares regression model, the error term is assumed to be independent of the value of the covariates (homoscedasticity). In contrast, quantile regression models allow for the variance of the error term to vary (heteroscedasticity) and make no assumptions about the variance structure (Yu, Lu, and Stander (2003)).

Computationally, quantile regression estimators may be formulated as a linear programming problem and efficiently solved by simplex or barrier methods via an optimization of a piecewise linear objective function in the residuals (Koenker and Hallock (2001)). The constant $\beta_0(\tau)$ and the coefficients $\beta_k(\tau)$ are estimated for 99 different quantiles ($\tau = 0.01, \ldots, 0.99$) using each time the entire data set. A detailed description of the mathematical formulation of conditional quantile regression is provided in chapter two.

The regressors x_k are standardized to have zero mean and unit standard deviation (SD). The ozone data are used in original units ($\mu g\ m^{-3}$). This type of standardization allows one to see at first glance the impact of a one-SD variation in the regressors (ceteris paribus effect) on the daily maximum 1-hr ozone concentrations, and, hence, makes a comparison of the relative importance of the covariates feasible. In total, sixteen regressors are included in the model, comprising nine meteorological variables, six variables related to the two perpendicular components per wind speed, and one measure of persistence (i.e. the lagged ozone concentration).

The meteorological data set exhibits intravariable (e.g. wind speed at different times) and intervariable collinearity (e.g. maximum surface temperature and upper air tem-

perature). However, this collinearity and the joint effect of the regressors can be very different among the ozone regimes (see results below). Hence, we include all meteorological variables in our model and consider the potential collinearity in the interpretation of the results (Chock, Kumar, and Herrmann (1982, 1984)). For inference, we use the bootstrap method to obtain the standard errors since it is valid under more general assumptions than the theoretical analogue.

5.4 Results

5.4.1 Regression coefficients

Figures 5.2 to 5.5 summarize some selected effects (standardized regression coefficients) of the explanatory variables on the daily maximum 1-hr ozone concentrations using the quantile regression and the ordinary least squares (OLS) regression models that include all 16 regressors. Bootstrap estimates of standard error (at the 95% confidence level) were calculated by randomly sampling each dataset with replacement (1000 times). In the quantile regression framework, a regression coefficient is a function of τ, while in the classical regression approach the regression coefficient is a single value for the entire distribution.

The constant in the quantile regression model has the typical and expected shape at all sites. The higher the ozone quantile, the higher the value of the constant in the model. For example at Liossia site, the constant of the model is close to 100 $\mu g \ m^{-3}$ at the 0.25 quantile and over 150 $\mu g \ m^{-3}$ at the 0.75 quantile. At all three suburban sites, the constant is close to 250 $\mu g \ m^{-3}$ for the very high quantiles. The narrow bounds indicate the high degree of confidence about the estimated values for the constant at all sites.

The effect of persistence (i.e. the impact of previous day's maximum ozone) on the daily maximum 1-hr ozone concentration has a common pattern at all four sites: it is positively increasing with the quantile (τ) for most of the range of τ. At Liossia, for example,
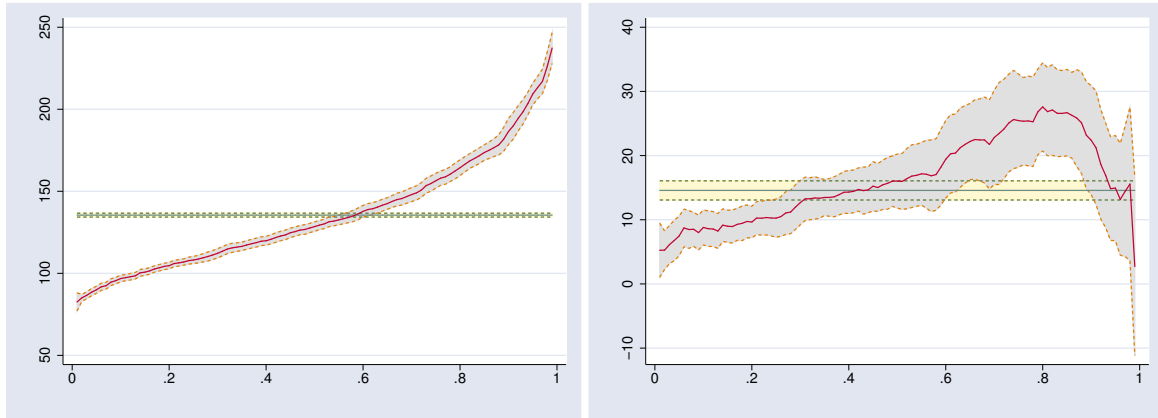
Figure 5.2: Estimated constant and lagged ozone concentration coefficient at Liossia. The standardized quantile regression coefficients (solid red line) are presented with their 95% confidence bounds (shaded in grey). The least squares regression coefficients (solid blue line) are also given with their 95% confidence bounds (shaded in yellow). The vertical axis shows the standardized regression coefficients ($\mu g \ m^{-3}$), the horizontal axis shows the ozone quantile. Please note the different vertical scales.
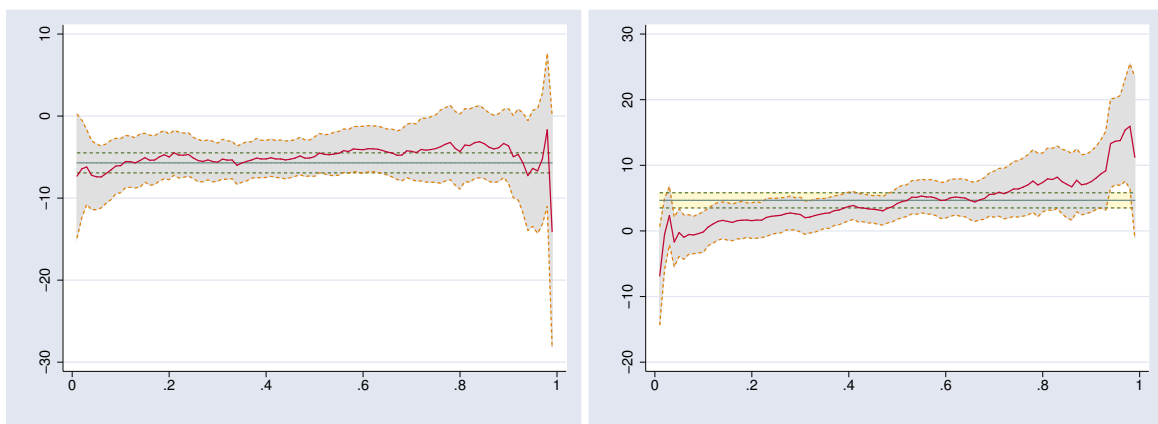

Figure 5.3: Estimated quantile regression effects of relative humidity and range of surface daily temperature at Smirni (compare figure 5.2).
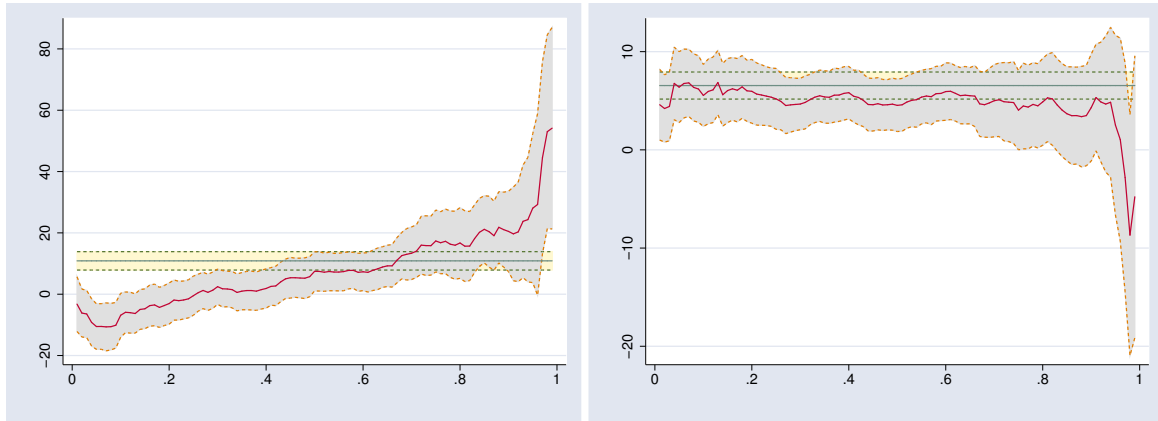
Figure 5.4: Estimated quantile regression effects of upper air temperature and change from previous day's upper air temperature at Maroussi (compare figure 5.2).
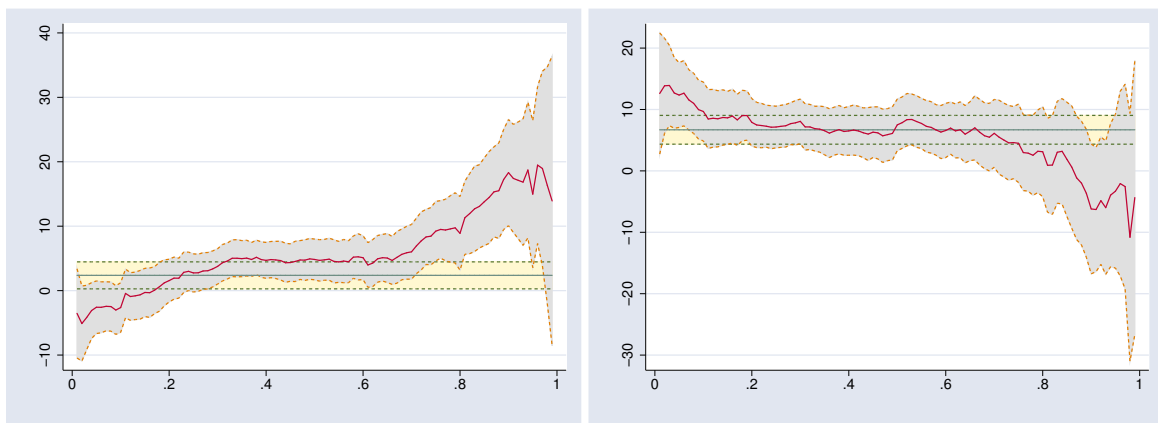

Figure 5.5: Estimated quantile regression effects of afternoon wind speed and solar radiation at Likovrissi (compare figure 5.2).

an increase of one-SD in the previous day's maximum 1-hr ozone concentration would result in a ceteris-paribus increase of 10 $\mu g\ m^{-3}$ in today's maximum ozone at the 0.25 quantile, but in an increase of 25 $\mu g\ m^{-3}$ at the 0.75 quantile. For higher quantiles, however, this effect is reduced, which indicates lower persistence of ozone concentrations at very extreme values of the distribution. Due to the wider confidence bands at high quantiles, the quantification of this effect is a bit less accurate than for low quantiles.

The OLS approach indicates that the statistically insignificant variables (i.e. with t-ratios lower than two in absolute value) at both Likovrissi and Maroussi are the nocturnal wind, relative humidity, range of surface daily temperature and maximum surface temperature. At Liossia and Smirni, the range of surface daily temperature and the morning wind are insignificant, respectively. On the other hand, the quantile regression approach shows that the effects of these meteorological variables on the daily maximum 1-hr ozone concentrations are more complex, which is reflected in the sign, size and significance of the estimated coefficients. We will try to isolate the most significant effects in our discussion. The interpretations are only suggestive of actual physical or chemical causes of ozone variations.

The nocturnal wind (absolute speed and both perpendicular components) is important only for the Smirni site, where previous night's north winds increase the next day's ozone concentration, while south winds decrease it. This may indicate that air entraining ozone travels from the suburban areas towards the urban site at Smirni during the night, increasing thereafter the next day's ozone levels at the site. As expected, this increase is not related to the ozone regime at Smirni, since the pollution is actually transported and not produced locally.

The morning wind speed at Likovrissi becomes significant at the 0.30 quantile onwards, exhibiting a negative effect. The dependence on low morning wind speeds may indicate that middle and higher ozone concentrations at Likovrissi are typified by local production of ozone. However, the effect is random in the upper tail of the ozone distribution.

A similar pattern is seen at Liossia and Maroussi. At Smirni, the effect of the morning wind speed is not statistically different from zero for the entire ozone distribution.

The afternoon wind affects the ozone concentrations at all four sites. At Likovrissi, the effect of the afternoon wind speed is negative at the very low quantiles, but it becomes positive and increases monotonically for the quantiles between 0.25 and 0.90. This may reflect that the very low daily maximum 1-hr ozone concentrations at Likovrissi are more likely to rely on local production processes, while ozone concentrations in the range between 95 and 205 $\mu g \ m^{-3}$ are due to transportation from south-southwest directions. An OLS regression would have averaged the effect of this variable, thus would have veiled the conclusions drawn previously. At Maroussi, afternoon winds coming from south-west direction can increase the ozone concentrations, especially at the higher quantiles. At Liossia, there is a clear positive effect of the afternoon south winds, which is increasing with the quantile. At Smirni, low afternoon winds, particularly from south, increase the ozone concentrations at all points of the distribution.

Solar radiation exhibits a constant positive effect on the ozone quantiles below 0.70 at Likovrissi, which is close to the OLS estimate. The effect of solar radiation is negligible for higher quantiles. Similar behavior is found at the remaining sites. This behavior is consistent with previous studies that show that beyond a certain limit, a further increase of the solar radiation does not affect the daily maximum 1-hr ozone concentrations significantly (Kioutsioukis, Melas, Ziomas, and Skouloudis (2000)).

Relative humidity has no significant effect on the entire ozone distribution at Likovrissi and Maroussi. At Liossia, however, the coefficient of relative humidity is negative and constant at the 0.10-0.40 quantiles, while it is not different from zero at the higher quantiles. At Smirni, relative humidity affects negatively the ozone concentrations up to the 0.70 quantile and becomes insignificant for higher quantiles. The stronger dependence on relative humidity of the ozone concentrations at Smirni, as opposed to the suburban sites, can be explained by the location of Smirni in the vicinity of the seaside.

The effect of the upper air temperature is insignificant at Likovrissi. At Maroussi, there is variation in the magnitude and the sign of the coefficient as we move up the ozone distribution. Specifically, upper air temperature has a negative coefficient at the very low quantiles, it becomes insignificant for the quantiles close to 0.30, and finally flips the sign around the 0.40 quantile. An increasing positive effect is evident at Liossia, where an increase of one-SD in upper air temperature could result in an increase of 40 $\mu g \ m^{-3}$ for the extreme ozone values. This indicates the detection of a photochemical signal for which upper air temperature is a good surrogate. In fact, upper air measurements appear in a number of studies to be most beneficial for estimating the conditions underlying extreme events (Burrows, Benjamin, Beauchamp, Lord, McCollor, and Thomson (1995), Pryor, McKendry, and Steyn (1995)). The negative coefficient of upper air temperature at Smirni seems to be random, given the large confidence interval.

At all sites, the change of the upper air temperature from the previous day seems to have a negligible effect on all ozone quantiles, except for the upper extremes, where the variable has a positive effect at Smirni, a negative effect at Maroussi and Likrovrissi and a negligible effect at Liossia. The effect of the range of surface daily temperature is insignificant at the suburban sites and increasingly positive at Smirni. The maximum surface temperature is insignificant at all sites.

The quantile curves of ozone versus the covariates at each of the four sites (not presented here) are clearly not parallel, which indicates that the distribution of the error term ε depends strongly on the value of the covariate. This confirms that in the present case of ozone modelling the assumption of ordinary regression is severely violated. These results indicate how much information can be veiled if the approach does not account for the varying effects of the determinants of ozone concentrations at the different ozone regimes. In particular, the quantile regression analysis shows that the OLS approach hides a large amount of information about the dependence of the conditional distribution of daily maximum 1-hr ozone concentrations on the lagged ozone concentrations, the morning and afternoon winds and the upper air temperature.
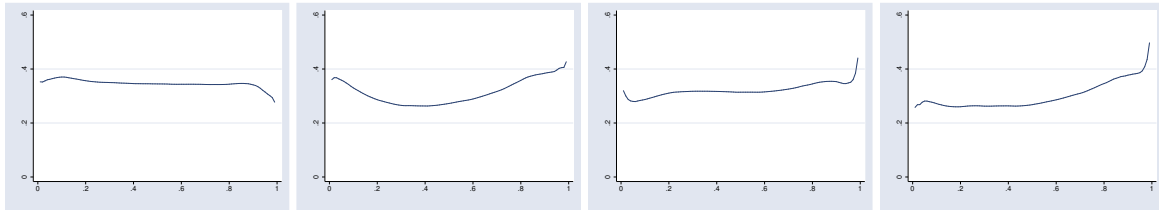
Figure 5.6: Goodness of fit $R^1(\tau)$ of the quantile regression model (vertical axes) as a function of the ozone quantile τ (horizontal axes) at Smirni, Liossia, Maroussi and Likovrissi (from left to right).

5.4.2 Model performance

The relative success of conditional quantile regression models at a specific quantile can be measured in terms of an appropriately weighted sum of absolute residuals and is denoted by $R^1(\tau)$ (Koenker and Machado (1999)). Like the coefficient of determination in the linear regression (R^2), the goodness of fit in the quantile regression $R^1(\tau)$ is in the interval $[0, 1]$. Unlike R^2, that measures a global goodness of fit over the entire conditional distribution, R^1 measures the local goodness of fit as a function of τ. Figure 5.6 illustrates the values of $R^1(\tau)$ at different ozone quantiles for each monitoring site. The quantile regression model of the 16 variables for Smirni exhibits a rather flat R^1 function indicating that the amount of ozone variability explained by the model is equal at all conditional quantiles. At Liossia, the bowl shaped curve of R^1 suggests that the model can explain the ozone variation in the tails of the conditional distribution better than in the center of the distribution. At Maroussi, the goodness of fit is slightly better for the higher than for the lower quantiles. The same pattern, but more pronounced, is apparent at Likovrissi.

As noted earlier, the R^2 of the classical linear regression and the R^1 of the quantile regression are not directly comparable due to their different nature; the former is a global measure, the latter is a local one. Even at the 0.50 quantile that can somehow be seen as a counterpart to the 'mean' ozone behavior (least squares approach), the R^1 and R^2 values are different, although the models' estimates of the daily maximum 1-hr ozone concentrations have similar values in both approaches. This confirms that R^1

and R^2 are not meant to be comparable in absolute values.

It was previously shown that the daily maximum 1-hr ozone concentrations at the suburban sites have skewed distributions with long right tails. As a consequence, the application of the OLS regression leads to an underprediction of the ozone concentrations in the right tail of the distribution. Within the quantile regression framework we suggest adopting the following approach for estimating a global goodness-of-fit to be compared with the R^2 of the classical linear regression: First, we compute the quantile regression coefficients for a small number of quantiles (in this case we chose $\tau \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$ as representative values). Subsequently, we divide the dataset into five uniform intervals according to the ozone values and calculate the predicted ozone concentrations using the quantile regression coefficients of the respective quantile regression function. The rationale in this approach resembles the 'amalgated' linear regression model for ozone by Eder, Davis, and Bloomfield (1994) that combines seven linear regression models for seven meteorological regimes identified by a cluster analysis.

Figure 5.7 presents the scatterplots of the OLS and the quantile regression model estimates versus the observed daily maximum 1-hr ozone concentrations at the four sites. The scatterplots associated with the OLS models reveal the expected tendency to underestimate the highest ozone concentrations (from around 180 $\mu g\ m^{-3}$ on). On the other hand, the results obtained from the amalgated quantile regression models exhibit a marked improvement in the estimated values. For the suburban sites, the amalgated models explain more of the ozone variation (between 72% and 80%) and have lower residual errors (RMSE between 24 and 30 $\mu g\ m^{-3}$) when compared to the OLS models ($R^2 < 0.5$, RMSE around 40 $\mu g\ m^{-3}$). The quantile regression model performance statistics are not significantly affected by using slightly different quantile functions than the five utilized above. Furthermore, the results of the conditional quantile regression approach present an improvement to previous attempts to model the daily maximum 1-hr ozone concentrations in the Athens basin using Classification and Regression
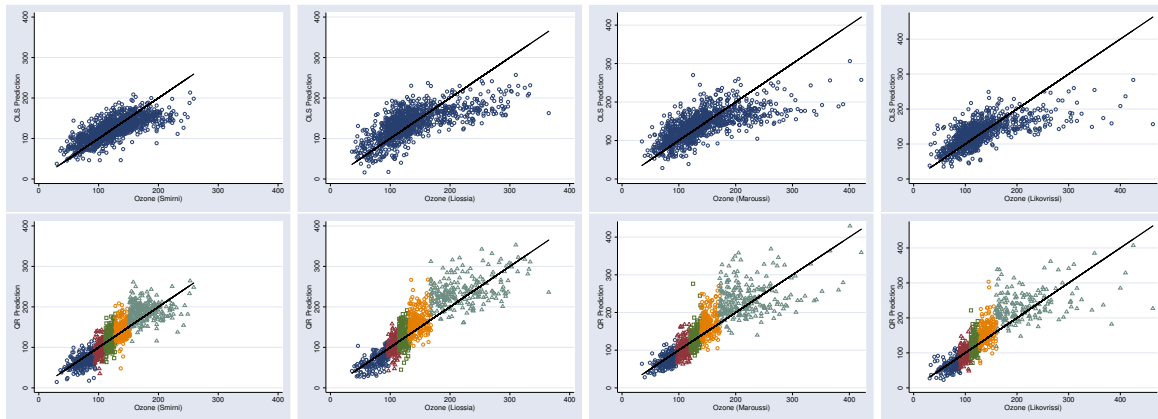
Figure 5.7: Estimated-versus-actual daily maximum 1-hr ozone concentrations at four sites in Athens, using ordinary least squares regression (OLS) models (top row) as well as the 'amalgated' quantile regression (QR) models at five quantiles (0.1, 0.3, 0.5, 0.7, 0.9) (bottom row).

Trees (Kaprara, Karatzas, and Moussiopoulos (2001)) or neural networks (Chaloulakou, Saisana, and Spyrellis (2003), Kioutsioukis, Melas, Ziomas, and Skouloudis (2000)).

5.4.3 Conditional densities

The quantile regression approach can also be applied when the aim of the analysis is to obtain an estimate of the *entire* conditional ozone distribution for *any* value of an explanatory variable considered individually in a simple regression model. The detailed procedure is explained in chapter two (compare Hyndman, Bashtannyk, and Grunwald (1996) and Yu and Jones (1998) who use alternative methods to compute conditional densities). Furthermore, by using one-period-lagged values of a regressor, the outcome allows for one-step forecasts of the ozone concentrations. In contrast to other forecast methods, this approach does not yield a single estimated value of tomorrow's ozone concentration but an estimated density distribution.

Figures 5.8 to 5.10 presents the estimated densities of the daily maximum 1-hr ozone concentrations conditional on (i) the lagged ozone concentration, (ii) the afternoon wind speed, and (iii) the upper air temperature. These explanatory variables were previously shown to be significant in explaining the ozone variation at most sites. The condi-
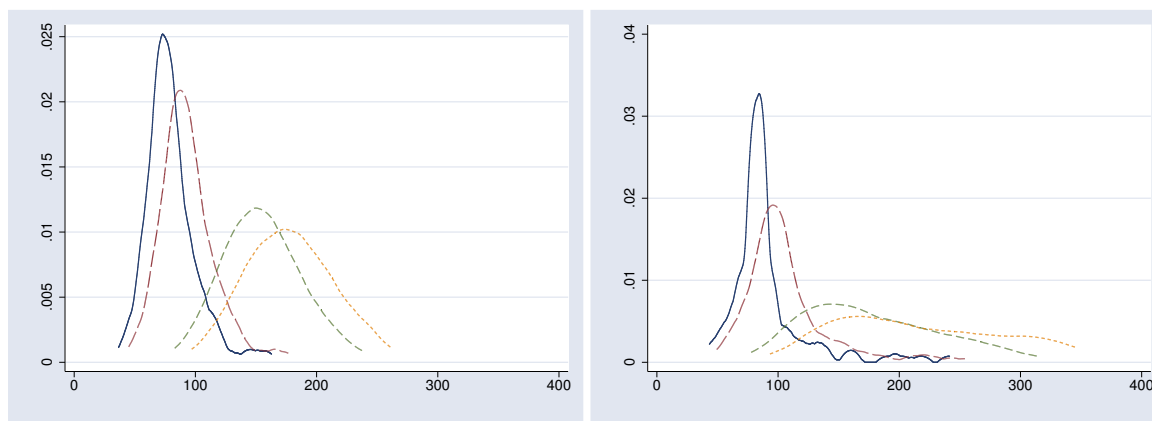
Figure 5.8: Conditional densities. The figure shows the estimated density of daily maximum 1-hr ozone concentration conditional on four different levels of the lagged ozone value at Smirni (left side) and Liossia (right side). The values of the explanatory variables are the 2%- (continuous blue line), 10%-(long-dashed red line), 90%- (dashed green line) and 98%-quantile (short-dashed orange line). The densities were calculated by applying a kernel density estimation on 99 quantile regression coefficients estimated in each case (see chapter two). Vertical axes indicate density, horizontal axes indicate ozone concentration ($\mu g\ m^{-3}$).

tional densities of the ozone concentrations are presented for two low levels (0.02- and 0.10-quantile, solid and long-dashed line) and two high levels (0.90- and 0.98-quantile, dashed and short-dashed line) of the independent variables, respectively. The values have been calculated by applying a kernel density estimation on 99 conditional quantiles in each case (see chapter two).

The results show that conditioning on low lagged ozone concentrations, there is a rather narrow unimodal density distribution of the daily maximum 1-hr ozone concentrations at all sites (Figure 5.8 presents the results only for Smirni and Liossia). However, as the previous day's maximum ozone concentration increases there is a tendency for the upper tail of the distribution to lengthen, which implies that the uncertainty in estimating today's ozone is considerably larger if yesterday's ozone is high. This effect is more pronounced at the suburban sites and shows that the classical regression assumption that a covariate affects only the location of the response distribution, but not its scale or shape, is violated. Similar conclusions were drawn in a case study on temperature

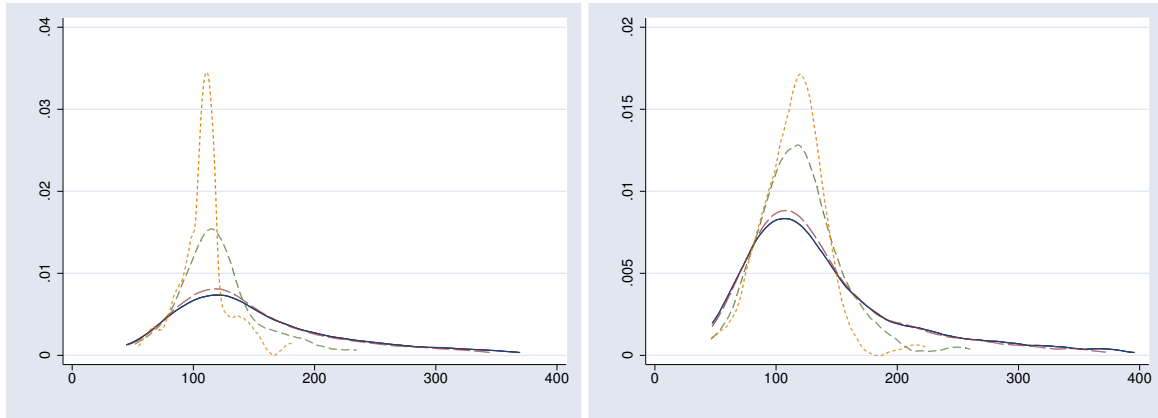Figure 5.9: Conditional densities. The figure shows density estimates of the ozone concentration conditional on afternoon wind speed at Liossia (left side) and Likovrissi (right side). Compare figure 5.8.
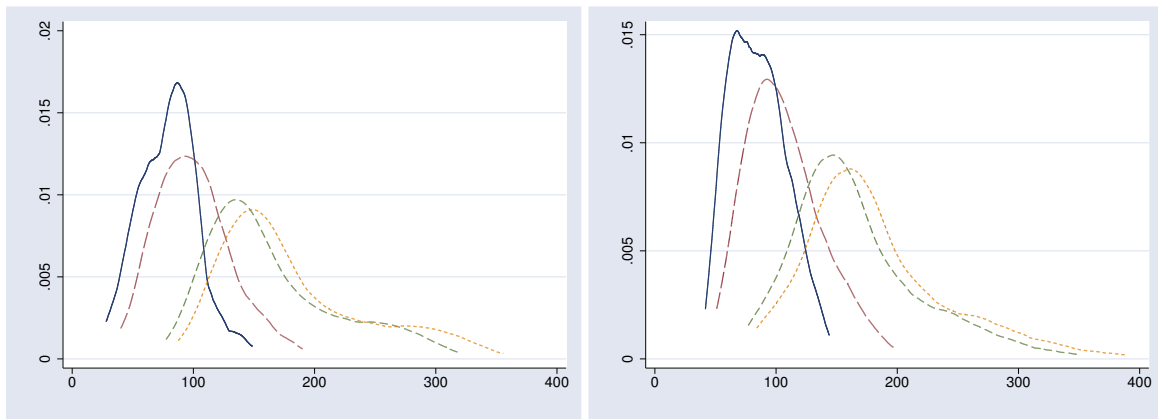

Figure 5.10: Conditional densities. The figure shows density estimates of the ozone concentration conditional on upper air temperature at 850 hPa at Liossia (left side) and Maroussi (right side). Compare figure 5.8.

distributions conducted by Koenker (2001).

The density of ozone distribution conditional on the afternoon wind speed shows a feature which is typical for all wind speed covariates but different from most of the other variables (e.g. upper air temperature): the higher the value of the covariate (the wind speed), the lower is the value of the daily maximum 1-hr ozone concentration, and vice versa. Furthermore, low wind speeds are associated with the whole range of ozone concentrations, while high wind speeds are related to a rather short range of ozone concentrations. These examples show that the analysis of the conditional densities can provide an additional insight into possibly complex relationships between the explanatory variables and the ozone concentrations.

5.5   Concluding Remarks

This chapter, using conditional quantile regression models applied on data from several monitoring sites in Athens, demonstrates that high daily maximum 1-hr ozone concentrations are clearly characterized by different stochastic relationships with meteorology and lagged ozone concentrations than mid or low ozone concentrations. Furthermore, the results confirm the expectation that high ozone concentrations are more persistent than low ozone concentrations. The upper air temperature is more influential for high ozone than for low ozone concentrations. This heterogeneity in the determinants of ozone is well captured by conditional quantile regression models, but is ignored if conventional regression models are employed. Consequently, the quantile regression results for different ozone regimes may provide important insights on the different determinants of ozone concentrations, without undertaking a cluster analysis for meteorological schemes. Furthermore, the application of an 'amalgated' quantile regression model leads to a significant improvement in explaining the ozone concentration. Finally, by calculating conditional densities, we are able to provide an entire density distribution for the forecasted ozone concentration.

Within the framework of describing the effect of meteorology and persistence on daily maximum 1-hr ozone concentrations, there are several arguments that render the quantile regression approach attractive: It is more information rich than ordinary least squares regression. Unlike Classification and Regression Trees, it is not empirical. Moreover, it can treat non-linear mechanisms without using transformed variables and interactions between variables, which in some cases could be cumbersome. Finally, quantile regression can easily be implemented and provides readily interpretable results, unlike neural networks that are often considered as 'black box' approaches. A potential application of our approach could be to forecast ozone values by using lagged explanatory variables or to meteorologically adjust ozone data sets. Future work will further exploit this direction.

Chapter 6

Conclusions

It is difficult to understand why statisticians commonly limit their inquiries to Averages and do not revel in more comprehensive views. Their souls seem as dull to the charm of variety as that of the native of one of our flat English counties, whose retrospect of Switzerland was that, if its mountains could be thrown into its lakes, two nuisances would be got rid of at once.

Sir Francis Galton (1889)

Roughly twenty-five years ago, Roger Koenker and Gilbert W. Bassett, Jr. proposed a new econometric method for the estimation of so-called regression quantiles for a response variable. In the first subsequent years, only few econometric articles dealt with the new procedure. This had several reasons: first, of course, for every new method it takes time to be noticed, discussed, and hopefully accepted by the scientific community (compare for example the time span that elapsed between the invention and widespread application of the Tobit model).

Furthermore, there were also some initially more substantial obstacles that retarded an immediate breakthrough of the quantile regression idea. The first was a slight difficulty of computing, since the solution of a quantile regression problem cannot be formulated as a simple closed expression. Moreover, not much was known about the asymptotic

behavior of the quantile regression process in the beginning, although already Koenker and Bassett (1978) provided some basic results. From this followed that also the derivation of valid inferential assertion seemed demanding in those days.

Fortunately, all of these problems have been extensively addressed and globally solved. The linear programming representation of quantile regression allowed the development of efficient algorithms. In addition, the impressive progression in the availability of fast and cheap computer power has substantially facilitated the practical application of the approach. A multiplicity of papers have analyzed the asymptotic properties of conditional quantiles, so today a well elaborated theory is readily available. This is also true for inferential questions, a variety of procedures have been introduced and successfully applied.

A nice indicator of the wide acceptance quantile regression has reached (rightly) in the meantime might be the average year of publication of applied quantile regression papers. Of course, the attentive reader could (and should) annotate at this point: "why only considering the mean and not the entire (possibly conditional) distribution of publication date?"

We hope to having coherently exposed the great potential inherent in the application of quantile regression and to having accomplished a small contribution to the fast growing empirical literature on conditional quantiles. We want to conclude our study by stating that with quantile regression at hand, one at least does not run the risk to be considered by Sir Francis Galton as a dull native of a flat English county.

Bibliography

ABADIE, A., J. ANGRIST, AND G. IMBENS (2002): "Instrumental Variables Estimates of the Effect of Subsidized Training on the Quantiles of Trainee Earnings," *Econometrica*, 70(1), 91–117.

ABBERGER, K. (1997): "Quantile Smoothing in Financial Time Series," *Statistical Papers*, 38, 125–148.

ABBERGER, K., AND S. HEILER (1997): "Applications of Nonparametric Quantile Regression to Financial Data," University of Konstanz.

ABREVAYA, J. (2001): "The Effects of Demographics and Maternal Behavior on the Distribution of Birth Outcomes," *Empirical Economics*, 26, 247–257.

AIGNER, D. J., T. AMEMIYA, AND D. J. POIRIER (1976): "On the Estimation of Production Frontiers: Maximum Likelihood Estimation of the Parameters of a Discontinuous Density Function," *International Economic Review*, 17(2), 377–396.

ALBRECHT, J., A. BJÖRKLUND, AND S. VROMAN (2003): "Is There a Glass Ceiling in Sweden?," *Journal of Labor Economics*, 21(1), 145–177.

AMEMIYA, T. (1982): "Two Stage Least Absolute Deviations Estimators," *Econometrica*, 50(3), 689–712.

ANDREWS, D. W. K. (1994): "Asymptotics for Semiparametric Econometric Models Via Stochastic Equicontinuity," *Econometrica*, 62(1), 43–72.

ANDREWS, D. W. K., AND M. BUCHINSKY (2000): "A Three-Step Method for Choosing the Number of Bootstrap Repetitions," *Econometrica*, 68(1), 23–51.

——— (2001): "Evaluation of a Three-Step Method for Choosing the Number of Bootstrap Repetitions," *Journal of Econometrics*, 103(1-2), 345–386.

——— (2002): "On the Number of Bootstrap Repetitions for BC_a Confidence Intervals," *Econometric Theory*, 18, 962–984.

ANG, A., AND J. CHEN (2002): "Asymmetric Correlations of Equity Portfolios," *Journal of Financial Economics*, 63, 443–494.

ANGRIST, J. D., V. CHERNOZHUKOV, AND I. FERNÁNDEZ-VAL (2004): "Quantile Regression under Misspecification, with an Application to the U.S. Wage Structure," MIT.

ARIAS, O. (2001): "Heterogeneity in the Determinants of Income Tax Evasion: A Quantile Regression Analysis of U.S. TCMP Data," University of Illinois.

ARIAS, O., K. F. HALLOCK, AND W. SOSA-ESCUDERO (2001): "Individual Heterogeneity in the Returns to Schooling: Instrumental Variables Quantile Regression Using Twins Data," *Empirical Economics*, 26, 7–40.

ATKINSON, A. B., J. GOMULKA, AND N. H. STERN (1990): "Spending on Alcohol: Evidence from the Family Expenditure Survey 1970-1983," *Economic Journal*, 100(402), 808–827.

BAE, K.-H., G. A. KAROLYI, AND R. M. STULZ (2003): "A New Approach to Measuring Financial Contagion," *Review of Financial Studies*, 16(3), 717–763.

BAI, J. (1995): "Least Absolute Deviation Estimation of a Shift," *Econometric Theory*, 11, 403–436.

BAIG, T., AND I. GOLDFAJN (1999): "Financial Market Contagion in the Asian Crisis," *IMF Staff Papers*, 46(2), 167–195.

BARNES, M. L., AND A. W. HUGHES (2002): "A Quantile Regression Analysis of the Cross Section of Stock Market Returns," University of Adelaide.

BARRETO, R. A., AND A. W. HUGHES (2004): "Under Performers and Over Achievers: A Quantile Regression Analysis of Growth," *The Economic Record*, 80(248), 17–35.

BARRODALE, I., AND F. D. K. ROBERTS (1973): "An Improved Algorithm for Discrete L1 Linear Approximation," *SIAM Journal of Numerical Analysis*, 10(5), 839–848.

——— (1974): "Solution of an Overdetermined System of Equations in the L1 Norm," *Communications of the Association for Computing Machinery*, 17, 319–320.

BARTELS, R. H., AND A. R. CONN (1980a): "Algorithm 563: A Program for Linearly Constrained Discrete ℓ_1 Problems," *ACM Transactions on Mathematical Software*, 6(4), 609–614.

——— (1980b): "Linearly Constrained Discrete ℓ_1 Problems," *ACM Transactions on Mathematical Software*, 6(4), 594–608.

BASSETT, G. W., AND H.-L. CHEN (2001): "Portfolio Style: Return-Based Attribution Using Quantile Regression," *Empirical Economics*, 26, 293–305.

BASSETT, G. W., AND R. KOENKER (1978): "Asymptotic Theory of Least Error Absolute Error Regression," *Journal of the American Statistical Association*, 73(363), 618–622.

——— (1982): "An Empirical Quantile Function for Linear Models with iid Errors," *Journal of the American Statistical Association*, 77(378), 407–415.

——— (1986): "Strong Consistency of Regression Quantiles and Related Empirical Processes," *Econometric Theory*, 2, 191–201.

BASSETT, G. W., R. KOENKER, AND G. KORDAS (2004): "Pessimistic Portfolio Allocation and Choquet Expected Utility," University of Illinois.

BASSETT, G. W., M.-Y. S. TAM, AND K. KNIGHT (2002): "Quantile Models and Estimators for Data Analysis," *Metrika*, 55, 17–26.

BEDARD, K. (2003): "School Quality and the Distribution of Male Earnings in Canada," *Economics of Education Review*, 22(4), 395–407.

BEKAERT, G., C. R. HARVEY, AND A. NG (2004): "Market Integration and Contagion," *Journal of Business*, forthcoming.

BERGEZ, J.-E., H. CARDOT, AND F. GARCIA (2003): "Quantile Regression Trees for Prediction in Irrigation Management," in *Proceedings of the 4th EFITA Conference in Debrecen, Hungary*, pp. 747–753.

BICKEL, P. J. (1975): "One-Step Huber Estimates in the Linear Model," *Journal of the American Statistical Association*, 70(350), 428–434.

BICKEL, P. J., AND D. A. FREEDMAN (1981): "Some Asymptotic Theory for the Bootstrap," *Annals of Statistics*, 9(6), 1196–1217.

BILIAS, Y., S. CHEN, AND Z. YING (2000): "Simple Resampling Methods for Censored Regression Quantiles," *Journal of Econometrics*, 99(2), 373–386.

BILIAS, Y., AND M. HALIASSOS (2004): "The Distribution of Gains from Access to Stocks," University of Cyprus.

BILLIO, M., AND L. PELIZZON (2003): "Contagion and Interdependence in Stock Markets: Have They Been Misdiagnosed?," *Journal of Economics and Business*, 55, 405–426.

BLOOMFIELD, P., J. A. ROYLE, L. J. STEINBERG, AND Q. YANG (1996): "Accounting for Meteorological Effects in Measuring Urban Ozone Levels and Trends," *Atmospheric Environment*, 30(17), 3067–3077.

BLOOMFIELD, P., AND W. S. STEIGER (1983): *Least Absolute Deviations: Theory, Applications, and Algorithms*. Birkhauser.

BLUNDELL, R., P. PASHARDES, AND G. WEBER (1993): "What Do We Learn About Consumer Demand Patterns from Micro Data?," *American Economic Review*, 83(3), 570–597.

BLUNDELL, R., AND J. L. POWELL (2003a): "Endogeneity in Nonparametric and Semiparametric Regression Models," in *Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress*, ed. by M. Dewatripont, L. P. Hansen, and S. J. Turnovsky, vol. II, pp. 312–358. Cambridge University Press.

——— (2003b): "Endogeneity in Semiparametric Binary Response Models," University College London.

BOFINGER, E. (1975): "Estimation of a Density Function using Order Statistics," *Australian Journal of Statistics*, 17(1), 1–7.

BOYER, B. H., M. S. GIBSON, AND M. LORETAN (1999): "Pitfalls in Tests for Changes in Correlations," International Finance Discussion Paper 597, Federal Reserve Board.

BUCHINSKY, M. (1991): "The Theory and Practice of Quantile Regression," Ph.D. Dissertation, Harvard University.

——— (1994): "Changes in the U.S. Wage Structure 1963-1987: Application of Quantile Regression," *Econometrica*, 62(2), 405–458.

——— (1995a): "Estimating the Asymptotic Covariance Matrix for Quantile Regression Models: A Monte Carlo Study," *Journal of Econometrics*, 68, 303–338.

——— (1995b): "Quantile Regression, Box-Cox Transformation Model, and the U.S. Wage Structure, 1963-1987," *Journal of Econometrics*, 65, 109–154.

——— (1996): "Bootstrapping Quantile Regression Model and other Methods," NSF Symposium on the Bootstrap, University of California.

——— (1998a): "The Dynamics of Changes in the Female Wage Distribution in the USA: A Quantile Regression Approach," *Journal of Applied Econometrics*, 13, 1–30.

——— (1998b): "Recent Advances in Quantile Regression - A Practical Guide for Empirical Reserach," *Journal of Human Resources*, 33(1), 88–126.

——— (2001): "Quantile Regression with Sample Selection: Estimating Women's Return to Education in the U.S.," *Empirical Economics*, 26, 87–113.

BUCHINSKY, M., AND J. HAHN (1998): "An Alternative Estimator for the Censored Quantile Regression Model," *Econometrica*, 66(3), 653–671.

BUCHINSKY, M., AND P. LESLIE (1997): "Educational Attainment and the Changing U.S. Wage Structure: Some Dynamic Implications," Yale University.

BURROWS, W. R., M. BENJAMIN, S. BEAUCHAMP, E. R. LORD, D. MCCOLLOR, AND B. THOMSON (1995): "CART Decision-Tree Statistical Analysis and Prediction of Summer Season Maximum Surface Ozone for the Vancouver, Montreal, and Atlantic Regions of Canada," *Journal of Applied Meteorology*, 34(8), 1848–1862.

CADE, B. S. (2003): "Quantile Regression Models of Animal Habitat Relationships," Ph.D. Dissertation, Colorado State University.

CADE, B. S., AND J. D. RICHARDS (2001): "User Manual for Blossom Statistical Software," U.S. Geological Survey.

CADE, B. S., J. W. TERRELL, AND R. L. SCHROEDER (1999): "Estimating Effects of Limiting Factors with Regression Quantiles," *Ecology*, 80(1), 311–323.

CAI, Z. (2002): "Regression Quantiles for Time Series," *Econometric Theory*, 18, 169–192.

CALVO, G. A., AND E. G. MENDOZA (2000): "Rational Contagion and the Globalization of Securities Markets," *Journal of International Economics*, 51(1), 79–113.

CALVO, S., AND C. M. REINHART (1996): "Capital Flows to Latin America: Is There Evidence of Contagion Effects," in *Private Capital Flows to Emerging Market*, ed. by G. Calvo, M. Goldstein, and E. Hochreiter. Institute of International Economics, Washington DC.

CANNON, A. J., AND E. R. LORD (2000): "Forecasting Summertime Surface-Level Ozone Concentrations in the Lower Fraser Valley of British Columbia: An Ensemble Neural Network Approach," *Journal of the Air & Waste Management Association*, 50, 322–339.

CANTELLI, F. (1933): "Sulla Probabilita coma Limita della Frequenza," *Rendiconti Accademia Lincei*, 26(1), 39ff.

CASHIN, P., M. S. KUMAR, AND C. J. MCDERMOTT (1995): "International Integration of Equity Markets and Contagion Effects," IMF Working Paper WP/95/110.

CHAKRABORTY, B. (2003): "On Multivariate Quantile Regression," *Journal of Statistical Planning and Inference*, 110, 109–132.

CHAKRABORTY, B., AND P. CHAUDHURI (1998): "On an Adaptive Transformation-Retransformation Estimate of Multivariate Location," *Journal of the Royal Statistical Society, Series B*, 60(1), 145–157.

CHALOULAKOU, A., M. SAISANA, AND N. SPYRELLIS (2003): "Comparative Assessment of Neural Networks and Regression Models for Forecasting Summertime Ozone in Athens," *Science of the Total Environment*, 313(1-3), 1–13.

CHAMBERLAIN, G. (1994): "Quantile Regression, Censoring, and the Structure of Wages," in *Advances in Econometrics, Sixth World Congress*, ed. by C. A. Sims, pp. 171–209. Cambridge University Press.

CHAN-LAU, J. A., D. J. MATHIESON, AND J. Y. YAO (2002): "Extreme Contagion in Equity Markets," IMF Working Paper WP/02/98.

CHANG, I. J., AND A. S. WEIGEND (1999): "Nonlinear Prediction of Conditional Percentiles for Value-at-Risk," New York University.

CHAUDHURI, P. (1991): "Nonparametric Estimates of Regression Quantiles and their Local Bahadur Representation," *Annals of Statistics*, 19(2), 760–777.

——— (1996): "On a Geometric Notion of Quantiles for Multivariate Data," *Journal of the American Statistical Association*, 91(434), 862–872.

CHAUDHURI, P., K. DOKSUM, AND A. SAMAROV (1997): "On Average Derivative Quantile Regression," *Annals of Statistics*, 25(2), 715–744.

CHAUDHURI, P., AND W.-Y. LOH (2002): "Nonparametric Estimation of Conditional Quantiles Using Quantile Regression Trees," *Bernoulli*, 8(5), 561–576.

CHAY, K. Y. (1995): "Evaluating the Impact of the 1964 Civil Rights Act on the Economic Status of Black Men using Censored Longitudinal Earnings Data," Princeton University.

CHAY, K. Y., AND B. E. HONORÉ (1998): "Estimation of Semiparametric Censored Regression Models: An Application to Changes in Black-White Earnings Inequality During the 1960s," *Journal of Human Resources*, 33(1), 4–38.

CHAY, K. Y., AND J. L. POWELL (2001): "Semiparametric Censored Regression Models," University of California.

CHEN, J.-E. (2001): "Investigations on Quantile Regression: Theories and Applications for Time Series Models," National Chung Cheng University.

CHEN, M.-Y., AND J.-E. CHEN (2002): "Application of Quantile Regression to Estimation of Value at Risk," National Chung-Cheng University.

CHEN, S., AND S. KHAN (1998): "Quantile Estimation of Non-Stationary Panel Data Censored Regression Models," Hong Kong University of Science and Technology.

——— (1999): "Semiparametric Estimation of a Semilinear Censored Regression Model," Hong Kong University of Science and Technology.

——— (2000): "Estimation of a Nonparametric Censored Regression Model," Hong Kong University of Science and Technology.

CHEN, S. H., AND S. KHAN (2003): "Nonparametric Estimation of Average Volatility Differentials in Selection Models with an Application to Returns to Schooling," State University of New York.

CHEN, X. R., Y. WU, AND L. C. ZHAO (1995): "A Necessary Condition for the Consistency of L_1 Estimates in Linear Models," *Indian Journal of Statistics, Series A*, 57(3), 384–392.

CHERNOZHUKOV, V. V. (2000): "Conditional Extremes and Near-Extremes: Concepts, Asymptotic Theory, and Economic Applications," Ph.D. Dissertation, Stanford University.

——— (2002a): "Extremal Quantile Regression," Massachusetts Institute of Technology.

——— (2002b): "Inference on Quantile Regression Process, An Alternative," Massachusetts Institute of Technology.

CHERNOZHUKOV, V. V., AND C. HANSEN (2001): "An IV Model of Quantile Treatment Effects," Massachusetts Institute of Technology.

——— (2002): "Inference for Distributional Effects Using Instrumental Quantile Regression," Massachusetts Institute of Technology.

CHERNOZHUKOV, V. V., AND H. HONG (2002): "Three-Step Censored Quantile Regression and Extramarital Affairs," *Journal of the American Statistical Association*, 97(459), 872–882.

——— (2003): "An MCMC Approach to Classical Estimation," *Journal of Econometrics*, 115, 293–346.

CHERNOZHUKOV, V. V., AND L. UMANTSEV (2001): "Conditional Value-at-Risk: Aspects of Modeling and Estimation," *Empirical Economics*, 26, 271–292.

CHOCK, D. P., S. KUMAR, AND R. W. HERRMANN (1982): "An Analysis of Trends in Oxidant Air Quality in the South Coast Air Basin of California, U.S.A.," *Atmospheric Environment*, 16(11), 2615–2624.

——— (1984): "An Update on Oxidant Trends in the South Coast Air Basin of California," *Atmospheric Environment*, 18(10), 2131–2134.

CHOCK, D. P., AND P. S. SLUCHAK (1986): "Estimating Extreme Values of Air Quality Data using Different Fitted Distributions," *Atmospheric Environment*, 20(5), 989–993.

CHRISTENSEN, L. R., D. W. JORGENSON, AND L. J. LAU (1975): "Transcendental Logarithmic Utility Functions," *American Economic Review*, 65(3), 367–383.

CIZEK, P. (2003): "Quantile Regression," in *XploRe Application Guide*, ed. by W. Härdle, Z. Hlávka, and S. Klinke, pp. 19–48. Springer.

COBOURN, W. G., L. DOLCINE, M. FRENCH, AND M. C. HUBBARD (2000): "A Comparison of Nonlinear Regression and Neural Network Models for Ground-Level Ozone Forecasting," *Journal of the Air & Waste Management Association*, 50, 1999–2009.

COLE, T. J. (1988): "Fitting Smoothed Centile Curves to Reference Data," *Journal of the Royal Statistical Society, Series A*, 151(3), 385–418.

COLE, T. J., J. V. FREEMAN, AND M. A. PREECE (1998): "British 1990 Growth Reference Centiles for Weight, Height, Body Mass Index and Head Circumference Fitted by Maximum Penalized Likelihood," *Statistics in Medicine*, 17(4), 407–429.

COLE, T. J., AND P. J. GREEN (1992): "Smoothing Reference Centile Curves: The LMS Method and Penalized Likelihood," *Statistics in Medicine*, 11, 1305–1319.

COMRIE, A. C. (1997): "Comparing Neural Networks and Regression Models for Ozone Forecasting," *Journal of the Air & Waste Management Association*, 47, 653–663.

COMRIE, A. C., AND B. YARNAL (1992): "Relationships Between Synoptic-Scale Atmospheric Circulation and Ozone Concentrations in Metropolitan Pittsburg, Pennsylvania," *Atmospheric Environment*, 26(3), 301–312.

CORSETTI, G., M. PERICOLI, AND M. SBRACIA (2003): "Some Contagion, Some Interdependence: More Pitfalls in the Tests of Financial Contagion," European University Institute.

CRAIG, S. G., AND P. T. NG (2000): "Using Quantile Smoothing Splines to Identify Employment Subcenters in a Multicentric Urban Area," University of Houston.

CUNNINGHAM, B. M. (2003): "The Distributional Heterogenity of Growth Effects: Some Evidence," *The Manchester School*, 71(4), 417–447.

DAVIES, T. D., P. M. KELLY, P. S. LOW, AND C. E. PIERCE (1992): "Surface Ozone Concentrations in Europe: Links with Regional-Scale Atmospheric Circulation," *Journal of Geophysical Research*, 97(D9), 9819–9832.

DAVIS, J. M., B. K. EDER, D. NYCHKA, AND Q. YANG (1998): "Modeling the Effects of Meteorology on Ozone in Houston using Cluster Analysis and Generalized Additive Models," *Atmospheric Environment*, 32(14-15), 2505–2520.

DAVIS, R. A., AND W. T. M. DUNSMUIR (1997): "Least Absolute Deviation Estimation for Regression with ARMA Errors," *Journal of Theoretical Probability*, 10(2), 481–497.

DE ANGELIS, D., P. HALL, AND G. A. YOUNG (1993): "Analytical and Bootstrap Approximations to Estimator Distributions in L1 Regression," *Journal of the American Statistical Association*, 88(424), 1310–1316.

DE GOOIJER, J. G., A. GANNOUN, AND D. ZEROM (2002): "A Multivariate Quantile Predictor," Universiteit van Amsterdam.

DEATON, A. (1997): *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*. Johns Hopkins University Press.

DEATON, A., AND J. MUELLBAUER (1980): "An Almost Ideal Demand System," *American Economic Review*, 70(3), 312–326.

DERWENT, R. G., P. G. SIMMONDS, S. SEURING, AND C. DIMMER (1998): "Observation and Interpretation of the Seasonal Cycles in the Surface Concentrations of Ozone and Carbon Monoxide at Mace Head, Ireland from 1990 to 1994," *Atmospheric Environment*, 32(2), 145–157.

DING, W., AND S. F. LEHRER (2002): "New Evidence on Education Production Functions and the Class Size Debate," University of Michigan.

EDER, B. K., J. M. DAVIS, AND P. BLOOMFIELD (1994): "An Automated Classification Scheme Designed to Better Elucidate the Dependence of Ozone on Meteorology," *Journal of Applied Meteorology*, 33(10), 1182–1199.

EDWARDS, S. (1998): "Interest Rate Volatility, Contagion and Convergence: An Empirical Investigation of the Cases of Argentina, Chile and Mexico," *Journal of Applied Economics*, 1(1), 55–86.

EFRON, B. (1979): "Bootstrap Methods: Another Look at the Jackknife," *Annals of Statistics*, 7(1), 1–26.

――― (1982): *The Jackknife, the Bootstrap, and Other Resampling Plans*. Society for Industrial & Applied Mathematics.

――― (1991): "Regression Percentiles Using Asymmetric Squared Error Loss," *Statistica Sinica*, 1, 93–125.

EFRON, B., AND R. J. TIBSHIRANI (1993): *An Introduction to the Bootstrap*. Chapman and Hall.

EIDE, E. R., D. D. GOLDHABER, AND M. H. SHOWALTER (2003): "Does Catholic High School Attendance Lead to Attendance at a More Selective College?," University of Washingtion.

EIDE, E. R., AND M. H. SHOWALTER (1998a): "The Effect of School Quality of Student Performance: A Quantile Regression Approach," *Economics Letters*, 58, 345–350.

――― (1998b): "Factors Affecting the Transmission of Earnings Across Generations: A Quantile Regression Approach," *Journal of Human Resources*, 34(2), 253–267.

EL BANTLI, F., AND M. HALLIN (1999): "L1 Estimation in Linear Models with Heterogenous White Noise," *Statistics & Probability Letters*, 45, 305–315.

――― (2002): "Estimation of the Innovation Quantile Density Function of an AR(q) Process Based on Autoregression Quantiles," *Bernoulli*, 8(2), 255–274.

ELLIS, S. P. (1998): "Instability of Least Squares, Least Absolute Deviation and Least Median of Squares Linear Regression (with discussion)," *Statistical Science*, 13(4), 337–344.

EMBRECHTS, P., A. J. MCNEIL, AND D. STRAUMANN (2002): "Correlation and Dependence in Risk Management: Properties and Pitfalls," in *Risk Management: Value at Risk and Beyond*, ed. by M. H. A. Dempster, pp. 176–223. Cambridge University Press.

ENGEL, E. (1857): "Die Productions- und Consumptionsverhältnisse des Königreichs Sachsen," *Zeitschrift des Statistischen Büros des Königlich Sächsischen Ministeriums des Innern*, 8/9, reprinted as appendix to "Die Lebenskosten Belgischer Arbeiter-Familien Früher und Jetzt" (1895), *International Statistical Institute Bulletin*, 9(1), 1-124.

ENGLE, R. F., AND S. MANGANELLI (2002): "CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles," University of California.

FALARIS, E. M. (2003): "A Quantile Regression Analysis of Wages in Panama," University of Delaware.

FAVERO, C. A., AND F. GIAVAZZI (2002): "Is the International Propagation of Financial Shocks Non-Linear? Evidence from the ERM," *Journal of International Economics*, 57(1), 231–246.

FEISTER, U., AND K. BALZER (1991): "Surface Ozone and Meteorological Predictors on a Sub-Regional Scale," *Atmospheric Environment*, 25(A9), 1781–1790.

FERREIRA, M. S. (2003): "Quantile Unit Root Test with an Application to the Purchasing Power Parity," University of Illinois.

FERSTERER, J., AND R. WINTER-EBMER (2002): "Are Austrian Returns to Education Falling Over Time?," University of Linz.

FERTIG, A. R. (2001): "Trends in Intergenerational Earnings Mobility," Princeton University.

FITZENBERGER, B. (1997a): "A Guide to Censored Quantile Regression," in *Handbook of Statistics*, ed. by G. Maddala, and C. Rao, vol. 15, pp. 405–437. Elsevier Science.

——— (1997b): "The Moving Blocks Bootstrap and Robust Inference for Linear Least Squares and Quantile Regressions," *Journal of Econometrics*, 82(2), 235–287.

——— (2003): "Gender Wage Differences Across Quantiles Accounting for Sample Selection," University of Mannheim.

FITZENBERGER, B., R. HUJER, T. E. MACURDY, AND R. SCHNABEL (2001): "Testing for Uniform Wage Trends in West-Germany: A Cohort Analysis Using Quantile Regressions for Censored Data," *Empirical Economics*, 26, 41–86.

FITZENBERGER, B., AND C. KURZ (2002): "New Insights on Earnings Trends Across Skill Groups and Industries in West Germany," University of Mannheim.

FITZENBERGER, B., AND G. WUNDERLICH (2002): "Gender Wage Differences in West Germany: A Cohort Analyis," *German Economic Review*, 3(4), 379–414.

FORBES, K. J., AND R. RIGOBON (2002): "No Contagion, Only Interdependence: Measuring Stock Market Comovements," *Journal of Finance*, 57(5), 2223–2261.

FREEDMAN, D. A. (1981): "Bootstrapping Regression Models," *Annals of Statistics*, 9(6), 1218–1228.

FRIEDMAN, M. (1975): *There's No Such Thing as a Free Lunch*. Open Court.

FURNO, M. (2003): "ARCH Tests and Quantile Regressions," Università di Cassino.

GALBRAITH, J. W., S. ZERNOV, AND V. ZINDE-WALSH (2001): "Conditional Quantiles of Volatility in Equity Index and Foreign Exchange Data," McGill University Montréal.

GALFALVY, H. C., X. HE, AND D. G. SIMPSON (2000): "Conditional Quantile Curves with an Application to Infrastructure Studies," University of Illinois.

GALTON, F. (1889): *Natural Inheritance*. Macmillan.

GANNOUN, A., J. SARACCO, AND K. YU (2003): "Nonparametric Prediction by Conditional Median and Quantiles," *Journal of Statistical Planning and Inference*, 117(2), 207–223.

GARCÍA, J., P. J. HERNÁNDEZ, AND A. LÓPEZ-NICOLÁZ (2001): "How Wide is the Gap? An Investigation of Gender Wage Differences Using Quantile Regression," *Empirical Economics*, 26, 149–167.

GARDNER, M. W., AND S. R. DORLING (2000): "Meteorologically Adjusted Trends in UK Daily Maximum Surface Ozone Concentrations," *Atmospheric Environment*, 34, 171–176.

GASSER, T., P. ZIEGLER, B. SEIFERT, A. PRADER, L. MOLINARI, AND R. LARGO (1994): "Measures of Body Mass and of Obesity from Infancy to Adulthood and their Appropriate Transformation," *Annals of Human Biology*, 21(2), 111–125.

GEYER, C. J. (1996): "On the Asymptotics of Convex Stochastic Optimization," University of Minnesota.

GLIVENKO, V. (1933): "Sulla Determinazione Empirica delle Leggi di Probabilita," *Giornale dell'Istituta Italiano degli Attuari*, 4, 92ff.

GOMANEE, K., S. GIRMA, AND O. MORRISSEY (2003): "Aid, Public Spending and Human Welfare: Evidence from Quantile Regression," Liverpool University.

GONZÁLEZ, X., AND D. MILES (2001): "Wage Inequality in a Developing Country: Decrease in Minimum Wage or Increase in Education Returns," *Empirical Economics*, 26, 135–148.

GOULD, W. W. (1992): "Quantile Regression with Bootstrapped Standard Errors (sg11.1)," *Stata Technical Bulletin*, 9, 19–21.

———— (1997): "Interquantile and Simultaneous Quantile Regression (sg70)," *Stata Technical Bulletin*, 38, 14–22.

GREEN, H. M., AND A. S. KOZEK (2001): "Modelling Weather Data by Approximate Regression Quantiles," Macquarie University.

GREENE, W. H. (2002): *Econometric Analysis*. Prentice Hall, 5th edn.

GUIMARAES, J. F. (2002): "Quantile Regression and the Gender Wage Gap over Time: Are All Women Swimming Upstream?," University of Illinois.

GUTENBRUNNER, C., AND J. JURECKOVA (1992): "Regression Rank Scores and Regression Quantiles," *Annals of Statistics*, 20(1), 305–330.

GUTENBRUNNER, C., J. JURECKOVA, R. KOENKER, AND S. PORTNOY (1993): "Test of Linear Hypotheses Based on Regression Rank Scores," *Journal of Nonparametric Statistics*, 2, 307–333.

HAHN, J. (1995): "Bootstrapping Quantile Regression Estimators," *Econometric Theory*, 11, 105–121.

———— (1997): "Baysian Bootstrap of the Quantile Regression Estimator: A Large Sample Study," *International Economic Review*, 38(4), 795–808.

HAIRE, S. L., C. E. BOCK, AND B. S. CADE (2000): "Paradigm Shift in Theory and Methods: Regression Quantile Analysis Enables New Insights for Ecology," in *Proceedings of the 4th International Conference on Integrating Geographic Information Systems and Environmental Modeling: Problems, Prospects, and Needs for Research*, ed. by B. Parks, K. M. Clarke, and M. P. Crane.

HAIRE, S. L., C. E. BOCK, B. S. CADE, AND B. C. BENNETT (2000): "The Role of Landscape and Habitat Charactericstics in Limiting Abundance of Grassland Nesting Songbirds in an Urban Open Space," *Landscape and Urban Planning*, 48, 65–82.

HALL, P., AND S. J. SHEATHER (1988): "On the Distribution of a Studentized Quantile," *Journal of the Royal Statistical Society, Series B*, 50(3), 381–391.

HALLIN, M., AND J. JURECKOVA (1999): "Optimal Tests for Autoregressive Models based on Autoregression Rank Scores," *Annals of Statistics*, 27(4), 1385–1414.

HALLOCK, K. F., R. MADALOZZO, AND C. G. RECK (2003): "Uncovering Heterogeneity in Managerial Pay - Firm Performance Relationships Using Quantile Regression," University of Illinois.

HAMAO, Y., R. W. MASULIS, AND V. NG (1990): "Correlations in Price Changes and Volatility across International Stock Markets," *Review of Financial Studies*, 3(2), 281–307.

HARTOG, J., P. T. PEREIRA, AND J. A. C. VIERA (2001): "Changing Returns to Education in Portugal During the 1980s and Early 1990s: OLS and Quantile Regression Estimators," *Applied Economics*, 33, 1021–1037.

HASAN, M. N., AND R. KOENKER (1997): "Robust Rank Tests of the Unit Root Hypothesis," *Econometrica*, 65(1), 133–161.

HASTIE, T., AND C. LOADER (1993): "Local Regression: Automatic Kernel Carpentry," *Statistical Science*, 8(2), 120–129.

HE, X. (1997): "Quantile Curves Without Crossing," *The American Statistician*, 51(2), 186–192.

HE, X., B. FU, AND W. K. FUNG (2003): "Median Regression of Longitudinal Data," *Statistics in Medicine*, 22, 3655–3669.

HE, X., AND F. HU (2002): "Markov Chain Marginal Bootstrap," *Journal of the American Statistical Association*, 97(459), 783–795.

HE, X., AND M.-O. KIM (2002): "On Marginal Estimation in a Semiparametric Model for Longitudinal Data with Time-Independent Covariates," *Metrika*, 55, 67–74.

HE, X., AND S. PORTNOY (1998): "Asymptotics of the Deepest Line," in *Applied Statistical Science III*, ed. by S. E. Ahmed, M. Ahsanullah, and B. K. Sinha, pp. 71–81. Nova Science.

HE, X., AND Q.-M. SHAO (1996): "A General Bahadur Representation of M-Estimators and its Application to Linear Regression with Nonstochastic Designs," *Annals of Statistics*, 24(6), 2608–2630.

HE, X., AND L.-X. ZHU (2003): "A Lack-of-Fit Test for Quantile Regression," *Journal of the American Statistical Association*, 98(464), 1013–1022.

HEAGERTY, P. J., AND M. S. PEPE (1999): "Semiparametric Estimation of Regression Quantiles with Application to Standardizing Weight for Height and Age in US Children," *Journal of the Royal Statistical Society, Series C (Applied Statistics)*, 48(4), 533–551.

HECKMAN, J. L. (1979): "Sample Selection Bias as a Specification Error," *Econometrica*, 47(1), 153–162.

HENDRICKS, W., AND R. KOENKER (1992): "Hierarchical Spline Models for Conditional Quantiles and the Demand for Electricity," *Journal of the American Statistical Association*, 87(417), 58–68.

HERCE, M. A. (1996): "Asymptotic Theory of LAD Estimation in a Unit Root Process with Finite Variance Errors," *Econometric Theory*, 12, 129–153.

HÁJEK, J., AND Z. SIDÁK (1967): *Theory of Rank Tests*. Academia.

HJORT, N. L., AND D. POLLARD (1993): "Asymptotics for Minmizers of Convex Processes," University of Oslo.

HOROWITZ, J. L. (1992): "A Smoothed Maximum Score Estimator for the Binary Response Model," *Econometrica*, 60(3), 505–531.

HOGG, R. V. (1975): "Estimates of Percentile Regression Lines Using Salary Data," *Journal of the American Statistical Association*, 70(349), 56–59.

HONORÉ, B. E., AND L. HU (2004): "On the Performance of Some Robust Instrumental Variables Estimators," *Journal of Business and Economics Statistics*, 22(1), 30–39.

HONORÉ, B. E., S. KHAN, AND J. L. POWELL (2002): "Quantile Regression under Random Censoring," *Journal of Econometrics*, 109, 67–105.

HOROWITZ, J. L. (1998): "Bootstrap Methods for Median Regression Models," *Econometrica*, 66(6), 1327–1351.

HOROWITZ, J. L., AND S. LEE (2002): "Semiparametric Methods in Applied Econometrics: Do the Models Fit the Data?," *Statistical Modelling*, 2, 3–22.

HU, L. (2003): "Dependence Patterns across Financial Markets: A Mixed Copula Approach," Ohio State University.

HUBBARD, M. C., AND W. G. COBOURN (1998): "Development of a Regression Model to Forecast Ground-Level Ozone Concentration in Louisville, KY," *Atmospheric Environment*, 32(14-15), 2637–2647.

HUBER, P. J. (1967): "The Behavior of Maximum Likelihood Estimates Under Nonstandard Conditions," in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, ed. by L. M. LeCam, and J. Neyman, vol. 1, pp. 221–233. University of California Press.

HUNTER, D., AND K. LANGE (1998): "An Optimization Transfer Algorithm for Quantile Regression," University of Michigan.

HUSKOVA, M. (1994): "Some Sequential Procedures Based on Regression Rank Scores," *Journal of Nonparametric Statistics*, 3, 285–298.

HYNDMAN, R. J., D. M. BASHTANNYK, AND G. K. GRUNWALD (1996): "Estimating and Visualizing Conditional Densities," *Journal of Computational and Graphical Statistics*, 5(4), 315–336.

JONES, A. M., AND A. L. NICOLÁS (2002): "The Importance of Individual Heterogeneity in the Decomposition of Measures of Socioeconomic Inequality in Health: An Approach Based on Quantile Regression," University of York.

JURECKOVA, J. (1977): "Asymptotic Relations of M-Estimates and R-Estimates in Linear Regression Model," *Annals of Statistics*, 5(3), 464–472.

JURECKOVA, J., AND L. B. KLEBANOV (1999): "Trimmed, Bayesian and Admissible Estimators," *Statistics and Probability Letters*, 42, 47–51.

JURECKOVA, J., AND B. PROCHAZKA (1994): "Regression Quantiles and Trimmed Least Squares Estimation in Nonlinear Regression Models," *Journal of Nonparametric Statistics*, 3, 201–222.

KAHN, L. M. (1998): "Collective Bargaining and the Interindustry Wage Structure: International Evidence," *Economica*, 65(260), 507–534.

KAN, K., AND W.-D. TSAI (2004): "Obesity and Risk Knowledge," IEAS Working Paper No. 04-A002, Institute of Economics, Academia Sinica.

KAPRARA, A., K. KARATZAS, AND N. MOUSSIOPOULOS (2001): "Maximum Ozone Level Prediction in Athens with the Aid of the CART System, a Modelling Study," in *Procedings of the 7th International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes in Beligerate, Italy*, pp. 193–196.

KEE, H. J. (2004): "An Investigation on the Gender Wage Gap in the Australian Labour Market," University of Adelaide.

KEEN, M. (1986): "Zero Expenditures and the Estimation of Engel Curves," *Journal of Applied Econometrics*, 1(3), 277–286.

KHAN, S. (2001): "Two-Stage Rank Estimation of Quantile Index Models," *Journal of Econometrics*, 100, 319–355.

KHAN, S., AND J. L. POWELL (1999): "Two-Step Estimation of Semiparametric Censored Regression Models," University of Rochester.

KIM, T.-H., AND C. MULLER (2003): "Two Stage Quantile Regression when the First Stage is Based on Quantile Regression," University of Nottingham.

KIM, T.-H., AND H. WHITE (2003): "Estimation, Inference, and Specifcation Testing for Possibly Misspecified Quantile Regression," *Advances in Econometrics*, 17, 107–132.

KING, M. A., AND S. WADHWANI (1990): "Transmission of Volatility between Stock Markets," *Review of Financial Studies*, 3(1), 5–33.

KIOUTSIOUKIS, I., D. MELAS, I. C. ZIOMAS, AND A. SKOULOUDIS (2000): "Predicting Peak Photochemical Pollutant Concentrations with a Combination of Neural Network Models," *Neural Netword World*, 10(6), 909–916.

KNIGHT, C. A., AND D. D. ACKERLY (2002): "Variation in Nuclear DNA Content Across Environmental Gradients: A Quantile Regression Analysis," *Ecology Letters*, 5, 66–76.

KNIGHT, K. (1998): "Limiting Distributions for L1 Regression Estimators under General Conditions," *Annals of Statistics*, 26(2), 755–770.

———— (1999): "Asymptotics for L_1-Estimators of Regression Parameters under Heteroscedasticity," *Canadian Journal of Statistics*, 27(3), 497–507.

———— (2000): "What are the Limiting Distributions of Quantile Estimators?," University of Toronto.

———— (2002): "On the Second Order Behaviour of the Bootstrap of L1 Regression Estimators," University of Toronto.

KOCHERGINSKY, M., X. HE, AND Y. MU (2003): "Practical Confidence Intervals for Regression Quantiles," University of Chicago.

KODRES, L. E., AND M. PRITSKER (2002): "A Rational Expectations Model of Financial Contagion," *Journal of Finance*, 57(2), 769–799.

KOENKER, R. (1994): "Confidence Intervals for Regression Quantiles," in *Proceedings of the Fifth Prague Symposium on Asymptotic Statistics*, ed. by P. Mandl, and M. Huskova, pp. 349–359. Physica.

———— (1997): "Rank Tests for Linear Models," in *Handbook of Statistics*, ed. by G. Maddala, and C. Rao, vol. 15, pp. 175–199. Elsevier Science.

———— (2000): "Galton, Edgeworth, Frisch, and Prospects for Quantile Regression in Econometrics," *Journal of Econometrics*, 95, 347–374.

———— (2001): "Linear Hypothesis: Regression (Quantile)," in *Statistics Section of the International Encyclopedia of the Social & Behavioral Sciences*, ed. by S. Fienberg, and J. Kadane, vol. 13, pp. 8893–8899. Elsevier Science.

———— (2003): "Quantile Regression for Longitudinal Data," University of Illinois.

———— (2004): "The QUANTREG Package v3.35 (Quantile Regression and Related Methods)," R Foundation for Statistical Computing.

KOENKER, R., AND G. BASSETT (1978): "Regression Quantiles," *Econometrica*, 46(1), 33–50.

———— (1982a): "Robust Tests for Heteroscedasticity Based on Regression Quantiles," *Econometrica*, 50(1), 43–62.

———— (1982b): "Tests of Linear Hypotheses and L1 Estimation," *Econometrica*, 50(6), 1577–1584.

KOENKER, R., AND G. W. BASSETT (1984): "Four (Pathological) Examples in Asymptotic Statistics," *The American Statistician*, 38(4), 209–212.

KOENKER, R., AND Y. BILIAS (2001): "Quantile Regression for Duration Data: A Reappraisal of the Pennsylvania Reemployment Bonus Experiments," *Empirical Economics*, 26, 199–220.

KOENKER, R., AND V. D'OREY (1987): "Algorithm AS 229: Computing Regression Quantiles," *Journal of the Royal Statistical Society, Series C (Applied Statistics)*, 36(3), 383–393.

——— (1994): "A Remark on Algorithm AS 229: Computing Dual Regression Quantiles and Regression Rank Scores," *Journal of the Royal Statistical Society, Series C (Applied Statistics)*, 43(2), 410–414.

KOENKER, R., AND O. GELING (2001): "Reappraising Medfly Longevity: A Quantile Regression Survival Analysis," *Journal of the American Statistical Association*, 96(454), 458–468.

KOENKER, R., AND K. F. HALLOCK (2000): "Quantile Regression - An Introduction," University of Illinois.

——— (2001): "Quantile Regression," *Journal of Economic Perspectives*, 15(4), 143–156.

KOENKER, R., AND J. A. F. MACHADO (1999): "Goodness of Fit and Related Inference Processes for Quantile Regression," *Journal of the American Statistical Assocation*, 94(448), 1296–1310.

KOENKER, R., P. NG, AND S. PORTNOY (1994): "Quantile Smoothing Splines," *Biometrika*, 81(4), 673–680.

KOENKER, R., AND B. J. PARK (1996): "An Interior Point Algorithm for Nonlinear Quantile Regression," *Journal of Econometrics*, 71(2), 265–283.

KOENKER, R., AND S. PORTNOY (1987): "L-Estimation for Linear Models," *Journal of the American Statistical Association*, 82(399), 851–857.

——— (1999): "Quantile Regression," Unpublished Manuscript, University of Illinois.

KOENKER, R., AND Z. XIAO (2002a): "Inference on the Quantile Regression Process," *Econometrica*, 70(4), 1583–1612.

——— (2002b): "Quantile Autoregression, Unit Roots and Asymmetric Interest Rate Dynamics," University of Illinois.

KOENKER, R., AND Q. ZHAO (1994): "L-Estimation for Linear Heteroscedastic Models," *Journal of Nonparametric Statistics*, 3, 223–235.

——— (1996): "Conditional Quantile Estimation and Inference for ARCH Models," *Econometric Theory*, 12, 793–813.

KOLTCHINSKII, V. I. (1997): "M-Estimation, Convexity and Quantiles," *Annals of Statistics*, 25(2), 435–477.

KOMUNJER, I. (2002): "Trois Essais d'Econometrie Financiere," Ph.D. Dissertation, HEC Paris.

——— (2003): "Efficient Conditional Quantile Estimation and Specification Testing," California Institute of Technology.

KORDAS, G. (2002a): "Credit Scoring Using Binary Quantile Regression," University of Illinois.

——— (2002b): "Smoothed Binary Quantile Regression," University of Pennsylvania.

KORDAS, G., AND S. F. LEHRER (2003): "Matching using Semiparametric Propensity Scores," University of Pennsylvania.

KOUL, H. L., AND A. K. E. SALEH (1995): "Autoregression Quantiles and Related Rank-Scores Processes," *Annals of Statistics*, 23(2), 670–689.

LEE, S. (2003): "Efficient Semiparametric Estimation of a Partially Linear Quantile Regression Model," *Econometric Theory*, 19, 1–31.

——— (2004): "Endogeneity in Quantile Regression Models: A Control Function Approach," University College London.

LEE, S. B., AND K. J. KIM (1993): "Does the October 1987 Crash Strengthen the Co-Movements among National Stock Markets?," *Review of Financial Economics*, 3(1), 89–103.

LEE, T.-H., AND Y. YANG (2004): "Bagging Binary Predictors for Time Series," University of California.

LEGGETT, D., AND S. CRAIGHEAD (2000): "Risk Drivers Revealed: Quantile Regression and Insolvency," in *Proceedings of the 10th AFIR Colloquium in Tromso, Norway*, pp. 211–238.

LEVIN, J. (2001): "For Whom the Reduction Count: A Quantile Regression Analysis of Class Size and Peer Effects on Scholastic Achievement," *Empirical Economics*, 26, 221–246.

LING, S., AND M. MCALEER (2001): "Regression Quantiles for Unstable Autoregressive Models," Osaka University.

LIPSITZ, S. R., G. M. FITZMAURICE, G. MOLENBERGHS, AND L. P. ZHAO (1997): "Quantile Regression Methods for Longitudinal Data with Drop-Outs: Application to CD4 Cell Counts of Patients Infected with the Human Immunodeficiency Virus," *Journal of the Royal Statistical Society, Series C (Applied Statistics)*, 48(4), 463–476.

LONGIN, F., AND B. SOLNIK (2001): "Extreme Correlation of International Equity Markets," *Journal of Finance*, 56(2), 649–676.

LORETAN, M., AND W. B. ENGLISH (2000): "Evaluating Correlation Breakdowns During Periods of Market Volatility," International Finance Discussion Paper 658, Federal Reserve Board.

LUCCHETTI, R. (2002): "Inconsistency of Naive GMM Estimation for QR Models with Endogenous Regressors," *Economics Letters*, 75, 179–185.

MA, L., AND R. KOENKER (2004): "Quantile Regression Methods for Recursive Structural Equation Models," University College London.

MACHADO, J. A. F., AND J. MATA (2000): "Counterfactual Decomposition of Changes in Wage Distributions using Quantile Regression," Universidade Nova de Lisboa.

——— (2001): "Earning Functions in Portugal 1982-1994: Evidence from Quantile Regression," *Empirical Economics*, 26, 115–134.

MACHADO, J. A. F., AND P. PARENTE (2003): "Multvariate Inferences using the Bootstrap," Universidade Nova de Lisboa.

MACHADO, J. A. F., AND J. M. C. SANTOS SILVA (2001): "Identification with Averaged Data and Implications for Hedonic Regression Studies," WP 10-01, Banco de Portugal.

——— (2002): "Quantiles for Counts," Universidade Nova de Lisboa.

MADALOZZO, R. (2002): "The Firm Size Wage Premium: A Quantile Analysis," Ibmec Business School.

MALONEY, W. F., W. CUNNINGHAM, AND M. BOSCH (2003): "Who Suffers Income Falls During Crises? An Application of Quantile Regression Analysis to Mexico 1992-1995," The World Bank.

MALONEY, W. F., AND E. P. RIBEIRO (1999): "Efficiency Wage and Union Effects in Labor Demand and Wage Structure in Mexico: An Application of Quantile Analysis," The World Bank.

MANNING, W. G., L. BLUMBERG, AND L. H. MOULTON (1995): "The Demand for Alcohol: The Differential Response to Price," *Journal of Health Economics*, 14, 123–148.

MANSKI, C. (1985): "Semiparametric Analysis of Discrete Response: Asymptotic Properties of the Maximum Score Estimator," *Journal of Econometrics*, 27, 313–333.

MARTINS, P. S. (2004): "Industry Wage Premia: Evidence From the Wage Distribution," *Economics Letters*, 83, 157–163.

MATA, J., AND J. A. F. MACHADO (1996): "Firm Start-Up Size: A Conditional Quantile Approach," *European Economic Review*, 40(6), 1305–1323.

MCFADDEN, D. (1974): "The Measurement of Urban Travel Demand," *Journal of Public Economics*, 3, 303–328.

MELLO, M., AND A. NOVO (2002): "The New Empirics of Economic Growth: Quantile Regression Estimation of Growth Equations," University of Illinois.

MELLO, M., AND R. PERRELLI (2003): "Growth Equations: A Quantile Regression Exploration," *Quarterly Review of Economics and Finance*, 43, 643–667.

MELLY, B. (2001): "The Theory and Practice of Quantile Regression," University of St. Gallen.

——— (2003): "Public-Private Sector Wage Differentials in Germany: Evidence from Quantile Regression," University of St. Gallen.

——— (2004): "Decomposition of Differences in Distribution using Quantile Regression," University of St. Gallen.

MIZERA, I., AND J. A. WELLNER (1998): "Necessary and Sufficient Conditions for Weak Consistency of the Median of Independent but not Identically Distributed Random Variables," *Annals of Statistics*, 26(2), 672–691.

MONTENEGRO, C. (2001): "Wage Distribution in Chile: Does Gender Matter? A Quantile Regression Approach," The World Bank, Development Research Group.

MORILLO, D. (1999a): "Income Mobility with Nonparametric Quantiles: A Comparison of the U.S. and Germany," University of Illinois.

——— (1999b): "Recursive Estimation of Regression Quantiles," University of Illinois.

MOSTELLER, F., AND J. W. TUKEY (1977): *Data Analysis and Regression: A Second Course in Statistics*. Addison-Wesley.

MUKHERJEE, K. (1999): "Asymptotics of Quantiles and Rank Scores in Nonlinear Time Series," *Journal of Time Series Analysis*, 20(2), 173–192.

NAHM, J.-W. (2001): "Nonparametric Quantile Regression Analysis of R&D-Sales Relationship for Korean Firms," *Empirical Economics*, 26, 259–270.

NELSON, R. B. (1999): *An Introduction to Copulas*. Springer, Lecture Notes in Statistics 139.

NEWEY, W. K. (2001): "Flexible Simulated Moment Estimation of Nonlinear Errors-in-Variables Models," *Review of Economics and Statisics*, 83(4), 616–627.

NEWEY, W. K., AND J. L. POWELL (1987): "Asymmetric Least Squares Estimation and Testing," *Econometrica*, 55(4), 819–847.

——— (1990): "Efficient Estimation of Linear and Type I Censored Regression Models Under Conditional Quantile Restrictions," *Econometric Theory*, 6(3), 295–317.

NGUYEN, B. T., J. W. ALBRECHT, S. B. VROMAN, AND M. D. WESTBROOK (2003): "A Quantile Regression Decomposition of Urban-Rural Inequality in Vietnam," Georgetown University.

NIELSEN, H. S., AND M. ROSHOLM (2001): "The Public-Private Sector Wage Gap in Zambia in the 1990s: A Quantile Regression Approach," *Empirical Economics*, 26, 169–182.

NIEMIRO, W. (1992): "Asymptotics for M-Estimators Defined by Convex Minimization," *Annals of Statistics*, 20(3), 1514–1533.

OBERHOFER, W. (1982): "The Consistency of Nonlinear Regression Minimizing the L1-Norm," *Annals of Statistics*, 10(1), 316–319.

OTSU, T. (2003): "Empirical Likelihood for Quantile Regression," University of Wisconsin-Madison.

PAGAN, A., AND A. ULLAH (1999): *Nonparametric Econometrics*. Cambridge University Press.

PAPASTEFANOU, G. (2001): "The ZUMA Data File Version of the GfK ConsumerScan Household Panel," in *Social and Economic Research with Consumer Panel Data*, ed. by G. Papastefanou, P. Schmidt, A. Börsch-Supan, H. Lüdtke, and U. Oltersdorf. ZUMA-Nachrichten Spezial Band 7.

PARK, B.-J. (2002): "On the Quantile Regression Based Tests for Asymmetry in Stock Return Volatility," *Asian Economic Journal*, 16(2), 175–191.

PARZEN, E. (1979): "Nonparametric Statistical Data Modeling," *Journal of the American Statistical Association*, 74(365), 105–121.

PARZEN, M. I., L. J. WEI, AND Z. YING (1994): "A Resampling Method Based on Pivotal Estimating Functions," *Biometrika*, 81(2), 341–350.

PERACCHI, F. (2001): *Econometrics*. John Wiley.

PEREIRA, P. T., AND P. S. MARTINS (2000): "Does Education Reduce Wage Inequality? Quantile Regression Evidence from Fifteen European Countries," Universidade Nova de Lisboa.

PERICOLI, M., AND M. SBRACIA (2003): "A Primer on Financial Contagion," *Journal of Economic Surveys*, 17(4), 571–608.

PHILLIPS, P. C. B. (1991): "A Shortcut to LAD Estimator Asymptotics," *Econometric Theory*, 7, 450–463.

POLLARD, D. (1991): "Asymptotics for Least Absolute Deviation Regression Estimators," *Econometric Theory*, 7, 186–199.

POON, S.-H., M. ROCKINGER, AND J. TAWN (2004): "Extreme Value Dependence in Financial Markets: Diagnostics, Models, and Financial Implications," *Review of Financial Studies*, 17(2), 581–610.

PORTNOY, S. (1991a): "Asymptotic Behavior of Regression Quantiles in Nonstationary, Dependent Cases," *Journal of Multivariate Analysis*, 38(1), 100–113.

——— (1991b): "Asymptotic Behavior of the Number of Regression Quantile Breakpoints," *SIAM Journal on Scientific Computing*, 12, 867–883.

——— (2003): "Censored Regression Quantiles," *Journal of the American Statistical Association*, 98(464), 1001–1012.

PORTNOY, S., AND J. JURECKOVA (1999): "On Extreme Regression Quantiles," University of Illinois.

PORTNOY, S., AND R. KOENKER (1989): "Adaptive L-Estimation for Linear Models," *Annals of Statistics*, 17(1), 362–381.

——— (1997): "The Gaussian Hare and the Laplacian Tortoise: Computability of Squared-Error versus Absolute Error Estimators (with discussion)," *Statistical Science*, 12(4), 279–300.

POTERBA, J. M., AND K. S. RUEBEN (1994): "The Distribution of Public Sector Wage Premia: New Evidence using Quantile Regression Methods," NBER Working Paper 4734.

POWELL, J. L. (1983): "The Asymptotic Normality of Two-Stage Least Absolute Deviations Estimators," *Econometrica*, 51(5), 1569–1576.

——— (1984): "Least Absolute Deviation Estimation for the Censored Regression Model," *Journal of Econometrics*, 25, 303–325.

——— (1986): "Censored Regression Quantiles," *Journal of Econometrics*, 32, 143–155.

——— (1991): "Estimation of Monotonic Regression Models Under Quantile Restrictions," in *Nonparametric and Semiparametric Methods in Econometrics and Statistics: Proceedings of the Fifth International Symposium in Economic Theory and Econometrics*, ed. by W. A. Barnett, J. Powell, and G. E. Tauchen, pp. 357–384. Cambridge University Press.

PRYOR, S. C., I. G. MCKENDRY, AND D. G. STEYN (1995): "Synoptic-Scale Meteorological Variability and Surface Ozone Concentrations in Vancouver, British Columbia," *Journal of Applied Meteorology*, 34(8), 1824–1833.

RANGVID, B. S. (2003): "Do Schools Matter? The Influence of School Inputs on Student Performance and Outcomes," Ph.D. Dissertation, Aarhus School of Business.

RAO, S. T., E. ZALEWSKY, AND I. G. ZURBENKO (1995): "Determining Temporal and Spatial Variations in Ozone Air Quality," *Journal of the Air & Waste Management Association*, 45, 57–61.

RAO, S. T., I. G. ZURBENKO, R. NEAGU, P. S. PORTER, J. Y. KU, AND R. F. HENRY (1997): "Space and Time Scales in Ambient Ozone Data," *Bulletin of the American Meteorological Society*, 78(10), 2153–2166.

RECK, C. (2003): "Heterogeneity and Black-White Labor Market Differences: Quantile Regression with Censored Data 1979-2001," University of Illinois Urbana-Champaign.

RIBEIRO, E. P. (2001): "Asymmetric Labor Supply," *Empirical Economics*, 26, 183–197.

RIGOBON, R. (2003): "On the Measurement of the International Propagation of Shocks: Is the Transmission Stable?," *Journal of International Economics*, 61, 261–283.

ROGERS, W. H. (1992): "Quantile Regression Standard Errors (sg11)," *Stata Technical Bulletin*, 9, 16–19.

——— (1993): "Calculation of Quantile Regression Standard Errors (sg11.2)," *Stata Technical Bulletin*, 13, 18–19.

RONNING, G. (1988): "Möglichkeiten und Grenzen der Ökonometrischen Nachfrageanalyse," *Wirtschaftswissenschaftliches Studium*, 17, 65–72.

——— (1992): "Share Equations in Econometrics: A Story of Repression, Frustration and Dead Ends," *Statistical Papers*, 33, 307–334.

RONNING, G., AND K. F. ZIMMERMANN (1991): "Mikroökonomie, Mikroökonometrie und Wirtschaftspolitik," *Ifo-Studien: Zeitschrift für empirische Wirtschaftsforschung*, 37(3-4), 189–200.

ROUSSEEUW, P. J., AND M. HUBERT (1999): "Regression Depth (with discussion)," *Journal of the American Statistical Association*, 94(445), 388–433.

ROYSTON, P., AND D. G. ALTMAN (1994): "Regression Using Fractional Polynomials of Continuous Covariates: Parsimonious Parametrc Modelling," *Journal of the Royal Statistical Society, Series C (Applied Statistics)*, 43(3), 429–467.

ROYSTON, P., AND E. M. WRIGHT (1998): "A Method for Estimating Age-Specific Reference Intervals ('Normal Ranges') based on Fractional Polynomials and Exponential Transformation," *Journal of the Royal Statistical Society, Series A*, 161(1), 79–101.

——— (2000): "Goodness-of-Fit Statistics for Age-Specific Reference Intervals," *Statistics in Medicine*, 19, 2943–2962.

RUPPERT, D., AND R. J. CARROLL (1980): "Trimmed Least Squares Estimation in the Linear Model," *Journal of the American Statistical Association*, 75(372), 828–838.

SAKOV, A., AND P. J. BICKEL (2000): "An Edgeworth Expansion for the m out of n Bootstrapped Median," *Statistics and Probability Letters*, 49(3), 217–223.

SANKARASUBRAMANIAN, A., AND U. LALL (2003): "Flood Quantiles in a Changing Climate: Seasonal Forecasts and Causal Relations," *Water Resources Research*, 39(5), 1134–1146.

SCHLINK, U., S. DORLING, E. PELIKAN, G. NUNNARI, G. CAWLEY, H. JUNNINEN, A. GREIG, R. FOXALL, K. EBEN, T. CHATTERTON, J. VONDRACEK, M. RICHTER, M. DOSTAL, L. BERTUCCO, M. KOLEHMAINEN, AND M. DOYLE (2003): "A Rigorous Inter-Comparison of Ground-Level Ozone Predictions," *Atmospheric Environment*, 37(23), 3237–3253.

SCHULTZ, T. P., AND G. MWABU (1998): "Labor Unions and the Distribution of Wages and Employment in South Africa," *Industrial and Labor Relations Review*, 51(4), 680–703.

SCHUNK, D. (2003): "The Pennsylvania Reemployment Bonus Experiments: How a Survival Model Hepls in the Analysis of the Data," University of Mannheim.

SHEATHER, S. J., AND J. S. MARITZ (1983): "An Estimate of the Asymptotic Standard Error of the Sample Median," *Australian Journal of Statistics*, 25, 109–122.

SIDDIQUI, M. M. (1960): "Distribution of Quantiles in Samples from a Bivariate Population," *Journal of Research of the National Bureau of Standards, Section B*, 64, 145–150.

SILVERMAN, B. W. (1986): *Density Estimation for Statistics and Data Analysis*. Chapman and Hall.

SKANS, O. N. (2002): "Labour Market Effects of Working Time Reductions and Demographic Changes," Ph.D. Dissertation, Uppsala University.

SMIRNOV, N. V. (1952): "Limit Distributions for the Terms of a Variational Series," *American Mathematical Society Translations, Series I*, 11, 82–143.

SMITH, R. L. (1989): "Extreme Value Analysis of Environmental Time Series: An Application to Trend Detection in Ground-Level Ozone," *Statistical Science*, 4(4), 367–377.

STATA CORPORATION (2003): *Stata Reference Manual Release 8*. Stata Press.

STONE, R. (1954): "Linear Expenditure Systems and Demand Analysis: An Application to the Pattern of British Demand," *Economic Journal*, 64(255), 511–527.

SUN, Y. (2003): "Modeling Health Production Function and Decomposing Health Inequalities Based on Semiparametric Quantile Regression Approach," University of Guelph.

TAM, M.-Y. S., G. W. BASSETT, AND U. SUKHATME (2002): "New Selection Indices for University Admission: A Quantile Approach," University of Illinois.

TANNURI-PIANTO, M. (2001): "Relative Earnings of Immigrants and Natives under Changes in the U.S. Wage Structure, 1970-1990: A Quantile Regression Approach," Universidade de Brasília.

TANNURI-PIANTO, M., AND D. M. PIANTO (2002): "Formal-Informal Earnings Differentials in Brazil. A Semi-Parametric Approach," University of Brasília.

TAYLOR, J. W. (1999): "A Quantile Regression Approach to Estimating the Distribution of Multiperiod Returns," *Journal of Derivatives*, 7(1), 64–78.

——— (2000): "A Quantile Regression Neural Network Approach to Estimating the Conditional Density of Multiperiod Returns," *Journal of Forecasting*, 19, 299–311.

TAYLOR, J. W., AND D. W. BUNN (1999): "A Quantile Regression Approach to Generating Prediction Intervals," *Management Science*, 45(2), 225–237.

TSIONAS, E. G. (2003): "Bayesian Quantile Inference," *Journal of Statistical Computation and Simulation*, 9, 659–674.

TUKEY, J. W. (1965): "Which Part of the Sample Contains the Information?," *Proceedings of the National Academy of Sciences of the United States of America*, 53, 127–134.

VAN DER VAART, A. W. (1998): *Asymptotic Statistics*. Cambridge University Press.

VAPNIK, V. N. (1998): *Statistical Learning Theory*. Wiley-Interscience.

VARIAN, H. R. (1992): *Microeconomic Analysis*. W. W. Norton & Company, 3rd edn.

VARIYAM, J. N., J. BLAYLOCK, AND D. SMALLWORLD (2002): "Characterizing the Distribution of Macronutrient Intake among U.S. Adults: A Quantile Regression Approach," *American Journal of Agricultural Economics*, 84(2), 454–466.

VENABLES, W. N., AND D. M. SMITH (2003): "An Introduction to R (Version 1.8.1)," R Foundation for Statistical Computing.

WANG, F. (2002): "The Application of Quantile Regression in Analysis of Gender Earnings Gap in China," East Carolina University.

WEISS, A. A. (1991): "Estimating Nonlinear Dynamic Models Using Least Absolute Error Estimation," *Econometric Theory*, 7, 46–68.

WELSH, A. H. (1988): "Asymptotically Efficient Estimation of the Sparsity Function at a Point," *Statistics and Probability Letters*, 6, 427–432.

WELSH, A. H., R. J. CARROLL, AND D. RUPPERT (1994): "Fitting Heteroscedastic Regression Models," *Journal of the American Statistical Association*, 82(425), 100–116.

WHANG, Y.-J. (2003): "Consistent Specification Testing for Quantile Regression Models," Korea University, Seoul.

WÖSSMANN, L. (2004): "How Equal Are Educational Opportunities? Family Background and Student Achievment in Europe and the United States," University of Munich.

WRIGHT, E. M., AND P. ROYSTON (1997): "A Comparison of Statistical Methods for Age-Related Reference Intervals," *Journal of the Royal Statistical Society, Series A*, 160(1), 47–69.

YANG, S. (1999): "Censored Median Regression Using Weighted Empirical Survival and Hazard Functions," *Journal of the American Statistical Association*, 94(445), 137–145.

YEE, T. W. (2004): "VGAM Family Functions for Quantile Regression," University of Auckland.

YING, Z., S. H. JUNG, AND L. J. WEI (1995): "Survival Analysis with Median Regression Models," *Journal of the American Statistical Association*, 90(429), 178–184.

YU, K. (2002): "Quantile Regression using RJMCMC Algorithm," *Computational Statistics and Data Analysis*, 40(2), 303–315.

YU, K., AND M. C. JONES (1998): "Local Linear Quantile Regression," *Journal of the American Statistical Association*, 93(441), 228–237.

YU, K., AND Z. LU (2002): "Local Linear Additive Quantile Regression," University of Plymouth.

YU, K., Z. LU, AND J. STANDER (2001): "Quantile Regression: Applications and Current Research Areas," University of Plymouth.

———— (2003): "Quantile Regression: Applications and Current Research Areas," *Journal of the Royal Statistical Society, Series D (The Statistician)*, 52(3), 331–350.

YU, K., AND R. A. MOYEED (2001): "Baysian Quantile Regression," *Statistics & Probability Letters*, 54(4), 437–447.

ZHAO, L., C. RAO, AND X. CHEN (1993): "A Note on the Consistency of M-Estimators in Linear Models," in *Stochastic Processes*, ed. by S. Cambains, pp. 359–367.

ZHAO, Q. (2001): "Asymptotically Efficient Median Regression in the Presence of Heteroscedasticiy of Unknown Form," *Econometric Theory*, 17, 765–784.

ZHOU, K. Q., AND S. L. PORTNOY (1996): "Direct Use of Regression Quantiles to Construct Confidence Sets in Linear Models," *Annals of Statistics*, 24(1), 287–306.

———— (1998): "Statistical Inference on Heteroscedastic Linear Models Based on Regression Quantiles," *Journal of Nonparametric Statistics*, 9, 239–260.