Quantile Regression Methods

BERND FITZENBERGER and RALF ANDREAS WILKE

Abstract

Quantile regression is emerging as a popular statistical approach, which complements the estimation of conditional mean models. While the latter only focuses on one aspect of the conditional distribution of the dependent variable, the mean, quantile regression provides more detailed insights by modeling conditional quantiles. Quantile regression can therefore detect whether the partial effect of a regressor on the conditional quantiles is the same for all quantiles or differs across quantiles. Quantile regression can provide evidence for a statistical relationship between two variables even if the mean regression model does not.

We provide a short informal introduction into the principle of quantile regression which includes an illustrative application from empirical labor market research. This is followed by briefly sketching the underlying statistical model for linear quantile regression based on a cross-section sample. We summarize various important extensions of the model including the nonlinear quantile regression model, censored quantile regression, and quantile regression for time-series data. We also discuss a number of more recent extensions of the quantile regression model to censored data, duration data, and endogeneity, and we describe how quantile regression can be used for decomposition analysis. Finally, we identify several key issues, which should be addressed by future research, and we provide an overview of quantile regression implementations in major statistics software. Our treatment of the topic is based on the perspective of applied researchers using quantile regression in their empirical work.

INTRODUCTION

We consider the linear regression model

$$y_i = x_i\beta + u_i,$$

with observations $i = 1, \dots, n$ and $x_i = (1, x_{2i}, \dots, x_{Ki})$ is $1xK$, includes a constant, and β is $Kx1.y$ and the regressors (covariates) x are observed, the error term u is not observed, and β is to be estimated. The error term is assumed to be zero in expectation given any value of the covariates, and it is independent of the covariates. The common approach to estimate the parameters of such a model is ordinary least squares (OLS), which estimates the conditional mean

Emerging Trends in the Social and Behavioral Sciences. Edited by Robert Scott and Stephen Kosslyn.
 ISBN 978-1-118-90077-2.

function $E(y \mid x)$. This is the average value of y given the observed covariates. A single parameter β_k is therefore informative about the partial relationship between the covariate x_k (with some abuse of notation) and the average value of y holding all other covariates constant. It is therefore an estimate of the average effect in the population from which the observations i are randomly sampled.

Quantile regression (QR) follows a somewhat different approach. Instead of estimating the average population effect, it estimates the effect at conditional quantiles of y given x (y and x are random variables with observations y_i and x_i). This is the conditional quantile function

$$q_{y|x}(\tau) \; := x\beta_\tau,$$

for quantile $\tau \in (0, 1)$. Similar to OLS, the linear QR model assumes that the conditional quantile function is linear in the parameters β_τ, but the parameters can vary in τ. A single parameter $\beta_{\tau k}$ is the change in the conditional quantile of y in response to a 1 unit increase in x_k holding all other covariates constant. If we consider $\tau = 0.5$, it is the change in the conditional median of y given x due to a 1 unit increase in x_k. QR is thus more informative than mean regression models as it considers the entire distribution of the dependent variable. Of course, there is nothing wrong with conditional mean models. However, they only focus on one feature of the conditional distribution as a function of covariates. As a matter of fact if the effect varies across quantiles and even changes its sign, the mean model may suggest no effect of a covariate on the mean, but the QR would reveal a more complete picture with non-constant effects across quantiles. Having not observed an effect on the mean may lead applied researchers to the conclusion that the variable does not play a role in the model, but this may not be true. A prominent example for application of QR in social sciences is wage regression where individual wages are explained by a number of covariates. When QR is applied to these models, it allows the effect of covariates to vary across quantiles. For example, an additional year of education may well have a different effect on lower (τ small) and higher (τ large) quantiles of the conditional wage distribution.

As an example, we estimate a wage equation with the log(wage) as the dependent variable and a number of independent variables with a sample of 369,389 full-time working employees in Germany in June 2004. The sample is an extract from German administrative labor market data (IAB Employment Sample 2004).

Figure 1 shows how the estimated QR coefficients vary across quantiles and how they relate to the OLS estimates (dashed line). For example, QR estimates the conditional quantile function of females being 40% lower than that of males at the first decile (quantile $\tau = 0.1$), while being only 15% lower at the ninth decile (quantile $\tau = 0.9$). The OLS estimate suggests that the

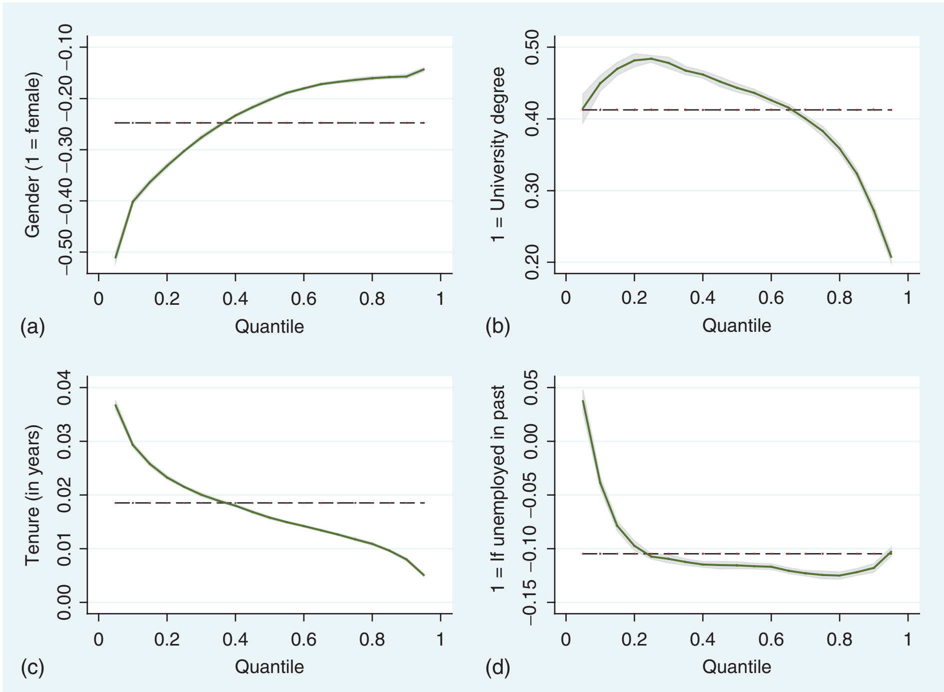

Figure 1 Estimated QR coefficients (solid line) with 95% confidence intervals (gray area). It also contains the estimated coefficient of the mean OLS regression (dashed line).

average wage for a female is around 25% lower than that for a male with the same other characteristics. The results therefore suggest that gender differences in wages may be smaller in higher paid jobs, in contrast to the so-called glass ceiling hypothesis. However, these results should not be overinterpreted because of the restrictive set of covariates used.

Roger Koenker, who is the key contributor to the foundational research of QR, has written several seminal articles, his widely used 2005 text book "Quantile Regression", and various surveys on QR methods. His textbook provides all key references up to the year 2004. Together with coauthors, Roger Koenker has contributed many computational resources to the open-source statistical package R. His work provides formal presentations of the material, detailed examples, and an introduction to computer code.

FOUNDATIONAL RESEARCH

Roger Koenker and Gilbert Bassett introduced the linear QR model in 1978 in their seminal article in *Econometrica* as a generalization of the estimation of an empirical sample quantile.

In a sample of n observations $\{y_1, \ldots, y_n\}$, the τ − quantile $[\tau \in (0,1)]$ of y is that value $q_y(\tau)$ for which at most a share of $(1-\tau)\cdot 100\,\%$ of the observations lie above that value and at most a share of $\tau \cdot 100\,\%$ of the observations lie below that value; Thus $q_y(\tau)$ cuts the observations into the lowest $\tau \cdot 100\,\%$ of the observations and the highest $(1-\tau)\cdot 100\,\%$ of the observations. For instance, the median corresponds to $\tau = 0.5$, the first decile to $\tau = 0.1$, and the ninth decile to $\tau = 0.9$. Quantiles are an alternative form to represent the distribution of a statistical variable, such that $F_y(q_y(\tau)) := \tau$, where $F_y(.)$ is the distribution function with $F_y(y) := \mathrm{P}(Y \leq y)$. Quantiles [formally the quantile process as a function of $\tau \in (0,1)$] represent the possible nonunique inversion of the distribution function.

The main insight to introduce QR is that the determination of an empirical quantile $q_y(\tau)$ can be viewed as the outcome of the following minimization exercise:

$$q_y(\tau) \ := \arg\min_q \left[\tau \sum_{i\ : y_i\ >\ q} |y_i - q| + (1-\tau) \sum_{i\ : y_i < q} |y_i - q| \right].$$

Implicit in this formulation is the focus on absolute differences from the location parameter q and the asymmetric weighting with τ and $(1-\tau)$, depending on whether an observation lies above or below q. This is the famous check function depicted in Figure 2.

For linear QR, we make q a function of covariates x_i and model the conditional quantile of the response variable y, given x_i as a linear function of $x_i\beta_\tau$. Thus, the determination of the linear QR amounts to the following minimization exercise:

$$\beta_\tau \ := \arg\min_\beta \left[\tau \sum_{i\ : y_i\ >\ x_i\beta} |y_i - x_i\beta| + (1-\tau) \sum_{i\ : y_i <\ x_i\beta} |y_i - x_i\beta| \right].$$

Linear QR coefficients describe the change in the conditional quantile of the response variable when a covariate changes by 1 unit. Analogous to a sample quantile, the implied regression relationship is such that at most a share of $(1-\tau)\cdot 100\,\%$ of the observations lie above the regression line and at most a share of $\tau \cdot 100\,\%$ of the observations lie below. The calculation of the regression coefficients corresponds to a linear program, which implies many properties of linear QRs. For instance, if the matrix of covariates has full rank K, then there will be at least K observations i with linearly independent vectors of covariates x_i such that the deviation from the regression line is exactly zero ($y_i = x_i\beta_\tau$). This is the so-called interpolation property.

The fact that the minimization problem cannot be solved by simple calculus methods is no restriction to today's computer power. To calculate linear QRs, effective algorithms based on refinements of the simplex method to solve

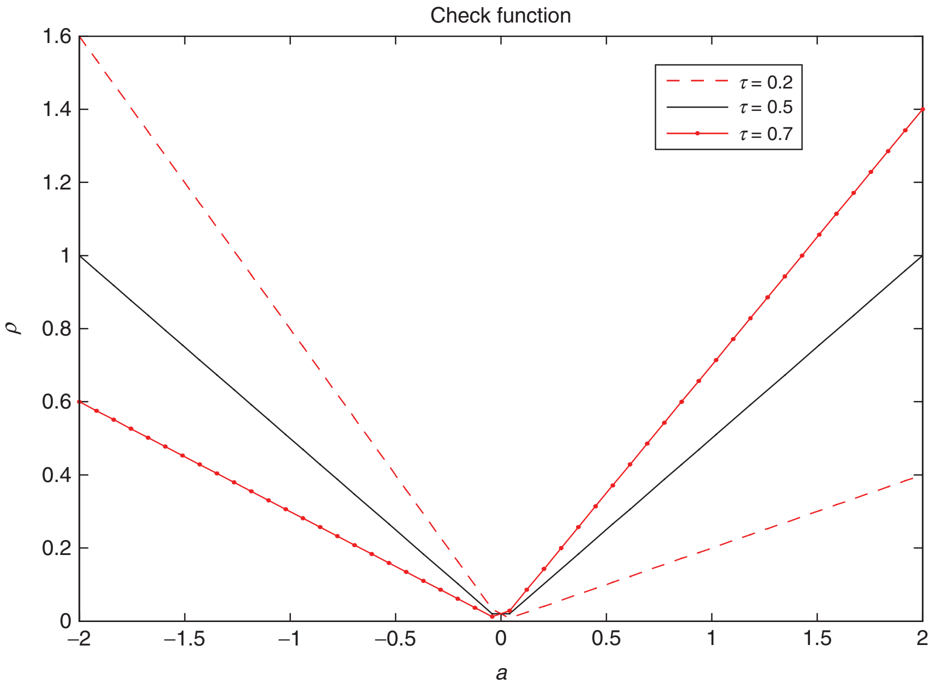

Figure 2 Check function.

linear programs are available. Interior point methods together with preprocessing of the data provide effective alternatives for very large data sets.

The asymptotic variance–covariance matrix is in fact very similar to OLS regression, provided the response variable follows a continuous distribution around the true conditional quantiles of interest. Instead of the variance of the error term (as in the OLS case), the asymptotic variance–covariance matrix involves the density of the response variable at the conditional population quantile. The statistical theory of QR also provides the joint distribution of the coefficient estimates at different quantiles, where the covariance matrix across quantiles has a similar structure as the variance–covariance matrix at the individual quantiles. Assuming a constant conditional distribution of the response variable around the conditional quantile allows one to estimate constant conditional densities based on the estimated residuals around the conditional QR (excluding the exact zeroes resulting from the interpolation property). In the heteroscedasticity case (which is the case when QR is interesting, see next paragraph), it would be necessary to use observation-specific density estimates, which would be computationally difficult and cumbersome. In his 2005 textbook, Roger Koenker discusses a simple and elegant alternative based on estimating QR slightly above and slightly below the quantile τ of interest, and then uses the implied conditional quantiles to obtain local density estimates. In practice, many researchers resort to bootstrapping methods to obtain asymptotically heteroscedasticity robust inference. A pairwise bootstrap of the estimation of QR at different quantiles (one estimates the QR at all quantiles of interest for the same resample) automatically provides estimates of the covariance across the different quantiles.

Estimating linear QRs for various values of τ provides a parsimonious picture of how the conditional distribution of the response variable changes with the covariates (see earlier example). In fact, due to the linear programming structure of the estimation problem, it is straightforward to calculate the entire process of QRs as a function of τ in a given sample, because the QR only change at finitely many values of τ. However, due to the combinatorical complexity of the problem, researchers rarely calculate the entire process in practice. Instead, they report the QR coefficients at selected equispaced quantiles, for example, for each decile ($\tau = 0.1, 0.2, \ldots, 0.9$) or each percentile ($\tau = 0.01, 0.02, 0.03, \ldots, 0.99$). When slope coefficients change across quantiles (which can be investigated by means of standard Wald tests), this is an indication of heteroscedasticity, that is, the conditional dispersion of the response variable changes with the covariates. In many applications (e.g., the effect of unions on wages), one would expect that the effect of a covariate changes along the conditional distribution (e.g., the effect of unions may be stronger in the lower part of the distribution than in the upper part of the distribution), that is, heteroscedasticity (changing dispersion) is a meaningful finding (e.g., unions reduce the dispersion of wages by increasing wages in the lower part of the distribution more strongly than in the upper part of the distribution).

An important advantage of QR compared with OLS regression relates to the equivariance property of quantiles under strictly monotone transformations, for which the τth quantile of the values of the function corresponds to the function value evaluated at the τth quantile of the original value. For instance, if we know that the conditional median of the logarithm of wages is a linear function of the covariates, we know also that the median of wages in levels corresponds to the exponential function applied to this linear function, thus modeling log wages entails modeling wages in levels. It is well known that this is not the case for OLS regression, because the expected value of wages is not equal to the exponential function applied to the expected value of log wages. The equivariance property of quantiles allows for more general strictly monotonous functions, such as the Box–Cox transformation defined for positive responses. However, computational issues may arise, because the inverse of the Box–Cox transformation may not be strictly positive for all data points, as pointed out by the authors and Xuan Zhang in 2010.

An apparent limitation of linear QR is the fact that nonparallel regression lines at different quantiles are bound to cross somewhere, thus for some values of the covariates, the predicted values at a higher quantile (e.g., the median $\tau = 0.5$) lies below the predicted value at a lower quantile (e.g., for $\tau = 0.49$). At the average values of the covariates, the ordering of predicted quantiles is preserved. One can use a sizeable incidence of quantile crossing among the observed values of the covariates as an indication for the need to respecify the model in a more flexible way, for example, by introducing nonlinear terms or nonparametric components as covariates. Holger Dette and Stanislav Volgushev (Victor Chernozhukov, Ivan Fernandez-Val, and Alfred Galichon (2010) have also written on this issue) discuss rearrangement methods (these involve smoothing of the estimates building on isotone regression techniques) to impose monotonicity of the predicted quantiles (not necessarily for coefficient estimates, for which the problem cannot be resolved). However, these methods may not resolve a problem of misspecification. It has been our experience that quantile crossing should be used as a guidance for misspecification of the model and that quantile crossing is often not a serious problem, if one allows for a sufficiently flexible specification (see section titled “Nonlinear Models”).

Table 1 in the appendix provides an overview of the implementation of QR in various statistical software packages. The “quantreg” package in R, developed by Roger Koenker, is most important for the dissemination of state-of-the-art QR techniques.

Table 1
Summary of Functionality of Major Statistical Packages for QR Analysis

	Statistical package			
Method	**R**	**STATA**	**TSP**	**Matlab**
Linear QR	• (quantreg)	• (qreg, sqreg)	• (lad with option)	• (Hunter, quantreg)
Nonlinear QR	• (quantreg)	—	—	• (Hunter)
Nonparametric QR	• (quantreg)	—	—	—
Censored QR	• (quantreg)	• (clad, cqiv)	• (lad with option)	—
Bootstrap for QR	• (quantreg)	• (sqreg with option)	• (lad with option)	• (quantreg)
IV QR	—	• (cqiv, Hansen)	—	• (Hansen)
Decomposition	—	• (various ado files)	—	—
Panel QR	• (rqpd)	—	—	—
Competing Risks QR	cmprskQR	—	—	—

•, implemented (requires additional package/command).

CUTTING-EDGE RESEARCH

QR is nowadays applied to a variety of more advanced models for cross-section, time-series, and panel data. Extensions of linear QR have been developed since the mid-1980, but most of this research was conducted after the year 2000 and it is still gradually developing. A broader process of knowledge transfer from method-based research into broader applied research in economics, and other social sciences did not start before the year 2005 but since then it is increasing in pace and still picking up. Being economists, it seems to us that empirical research in biostatistics has started to pick up QR methods, in particular in survival analysis taking account of censoring. Here, we present an overview of important model extensions and fields of applications of more advanced QR models. Roger Koenker contributed to some of these extensions, and the references given to his work cover most of these extensions.

Nonlinear Models

The linear functional form of the conditional quantile function might be too restrictive in an application. Roger Koenker outlines the estimation of nonlinear quantile functions in his textbook. The QR model can be extended to allow for nonlinear relationships, that is, $q_{y|x}(\tau)=g_{\tau}(x)$ with g_{τ} being some nonlinear function that satisfies some regularity conditions. In the abovementioned minimization exercise, $g_{\tau}(x)$ replaces the predicted values $x_i\beta_{\tau}$. Nonlinear, strictly monotonous transformations (such as the logarithm or a Box–Cox transformation) of the response variable to achieve a linear QR (as discussed above) are a special case. The parametric, nonlinear QR model can be applied if g_{τ} is known subject to some unknown parameters. This is analogous to nonlinear least squares regression.

A nonparametric QR on a set of continuous covariates can be applied if g_{τ} is unknown. A local smoothing approach such as kernel smoothing can be used to locally estimate the unknown g_{τ}, the estimates being subject to the curse of dimensionality. (The convergence rate of the estimator becomes slower when the number of covariates increases.) Estimating a weighted linear QR just on an intercept (the residuals are weighted by kernel weights) for continuous covariates produces the QR alternative of a Nadaraya–Watson (local constant) kernel regression. Local linear (local polynomial) QR can be estimated using a local linear (local polynomial) approximation in the weighted QR. Note that using a local constant kernel regression with the same bandwidth for all quantiles resolves by construction the quantile crossing problem. The literature also involves semiparametric specifications with additive nonparametric components estimated using backfitting techniques, involving an iterative procedure. During an iteration step, each component of the

QR specification is estimated recursively based on the previous estimates of all other components.

Censored Quantile Regression

The QR method can also be applied to a censored regression. Here, the conditional quantile function for y corresponds either to the censoring value or it is linear in the covariates. A prominent example from social sciences involves labor supply, which is either zero or positive and many individuals supply zero hours of work. Another example involves health expenditures, which are either zero or positive. Models for such response variables can be estimated by censored quantile regressions (CQR). The interpretation of estimation results regarding the observed censored values requires the computation of partial effects accounting for censoring.

We first consider the simple case of right censoring (e.g., top coding of wages or right censoring of durations of ongoing spells), where for censored observations we only know that the statistical variable of interest exceeds a certain known threshold. If this threshold is constant for all observations and the τth quantile of the censored observations lies below the threshold, we know that the τth quantile of the censored observations corresponds to the τth quantile of the uncensored observations. Here, the censored observations correspond to the actual values of the variable of interest, if they are not censored, and to the censoring value, if they are censored, that is, $y_i = \min(y_i^*, c)$ where y_i is the observed censored value, y_i^* is the true uncensored value, and c is the censoring threshold.

Roger Koenker discusses three approaches to CQR in his 2008 article. The first approach developed by James Powell (1986) involves the case of fixed censoring where the censoring values may vary across observations, but it is known for all observations. The estimator replicates the censoring mechanism in the regression specification for the censored observation as $g_\tau(x_i) = \min(x_i\beta_\tau, c_i)$, where c_i is the observation-specific, known censoring value. This is a special case of nonlinear QR. The Powell estimator provides a semiparametric alternative to the standard Tobit estimator for the censored regression model, which relies on the assumption of a normally distributed error term and which is inconsistent under heteroscedasticity. However, the calculation of the Powell estimator is difficult when there is a lot of censoring. Various modifications of the estimator have been suggested in the literature to overcome these difficulties. Two appealing approaches have been suggested by Steve Portnoy (2003) and by Limin Peng and Yijian Huang (2008). These involve regression versions of nonparametric estimators of distribution functions under independent censoring (Kaplan–Meier and Nelson–Aalen estimator), where the censoring values are only known for those observations that are censored (the case of random censoring).

DURATION MODELS

The response variable in duration analysis (or survival analysis) is time until an event or failure occurs (single-risk model). QR is an attractive approach to analyze the distribution of a duration as it can allow for different effects of covariates on lower and higher quantiles of the conditional distribution. For example, when the response variable is the duration of unemployment and we want to study the effect of a training programme on the job-taking time of unemployed people, we would expect that such a training programme at the first instance increases shorter unemployment periods, because the unemployed are locked into the programme. For longer unemployment periods, we would hope to find a shortening effect of the training programme once the training is completed. Thus, it is conceivable that even the sign of the estimated coefficient of training varies across quantiles. Standard parametric and semiparametric duration models such as proportional hazard models do not possess this degree of flexibility as they typically model the effect of a covariate by one single parameter. While linear QR can be directly applied to duration data, these data are often characterized by being censored. In the presence of censoring a CQR can be estimated. Roger Koenker did some pioneering work on QR for duration models, which is reviewed in his 2005 textbook, and his 2008 article on CQR has a focus on applications in duration analysis. Our 2006 survey discusses the usefulness and the limitations of QR for duration analysis in the presence of independent censoring.

TIME-SERIES MODELS

QR is also becoming increasingly popular for the empirical analysis with time-series data. Roger Koenker and Zhijie Xiao consider a class of quantile autoregression models where the covariates involve the lags of the response variable in discrete time. Such QR time-series models allow for a systematic influence of the lagged dependent variable on the location, scale, and shape of the distribution of the response variable. For the analysis of univariate time series, the models include the autoregressive model (both stationary processes and processes with unit roots) and the autoregressive conditional heteroscedasticity (ARCH) model. Such models allow for asymmetric dynamics and local persistence in time series and thus may bridge the gap between stationary and integrated time-series processes. Such models have been applied to macroeconomic time-series data and financial data. A standard generalized autoregressive conditional heteroscedasticity (GARCH) model implies a symmetric persistence in the second moment of a time series, irrespective of the direction of change. A quantile autoregression model allows for asymmetric dynamics, implying different responses in the conditional scale depending on whether there was a strong downside or upside movement of the response variable in the recent past. Downside movements or upside movements in the recent past may involve different degrees of persistence (for instance, unit root behavior may only exist in some part of the conditional distribution of the response variable). Such effects may prove important in the analysis of financial data. Extensions of quantile autoregression models to the analysis of multivariate time series are possible, including the estimation of quantile vector autoregressions and quantile co-integrating regressions.

KEY ISSUES FOR FURTHER RESEARCH

We discuss some pertinent issues for further research that are related to our own research. Due to space constraints, the choice has to be somewhat eclectic.

Censored Quantile Regression

We raise four issues: First, Roger Koenker describes the major computational difficulties involved and discusses some practical solutions (see above). Somewhat practical approaches exist for the CQR model under fixed and random censoring, although there does not seem to exist a consensus in the literature on what works best (correspondingly, popular software package use different algorithms to calculate CQR, often with little justification of the particular choice). Second, it should be noted that identification of CQR is a tenuous issue, because the CQR line involves an extrapolation based on functional form assumptions into the censored part of the data. A more substantive analysis of this issue would be useful. Third, the random censoring case assumes that the censoring values are independent of the response variable (at least conditional on the covariates). CQR models for random censoring are therefore not applicable to all empirical problems. While there has been some work for CQR if observations are dependent, there is still a gap to accommodate various forms of dependent censoring in the CQR model. Existing studies impose stringent assumptions on the regression models and are plagued by high computational costs due to multiple step algorithms. Fourth, two-limit CQR allowing for censoring both from above and below (in an analogy to the two-limit tobit model) is a straightforward extension under fixed censoring, and similar algorithms can be used. Under random censoring, at most one of the two censoring points is observed (and neither one is observed for uncensored observations) and the rationale for the algorithms suggested by Steve Portnoy (2003) and by Limin Peng and Yijian Huang (2008) breaks down (see above).

Duration Models

So far, QR models for duration data have been mainly used for single-risk models with independent right censoring. Here, the QR model estimates the conditional quantile functions of the duration distribution conditional on a time-invariant set of covariates. We think that there are a number of extensions still to be analyzed. These include multiple spell QR duration models (this issue is related to panel data applications, see in the following section) and QR models that explicitly allow for unobserved heterogeneity and time-varying covariates. Duration analysis based on modeling hazard rates can take these issues into account, typically under the restrictive proportional hazard assumption. Regarding the last point, QR can condition on a predetermined time path of the covariates, but the analysis may quickly involve a large number of coefficients to be estimated. Although applications of QR for competing risks models with possibly dependent competing risks do exist, the applications cannot estimate the duration distributions of the separate competing risks (transitions to a certain destination state) but rather estimate the cumulative incidence function of the separate risks. The cumulative incidence—the duration of observed transitions into a certain destination state—is a common descriptive tool (especially in Biostatistics), but in the absence of knowledge about the dependence structure between risks, it is difficult to infer the duration distributions of the competing risks, which are often objects of prime interest in the social sciences.

Decomposition Analysis and Unconditional QR

For a long time, a major restriction of QR in applied research was that conditional quantiles do not aggregate directly to the unconditional distribution of the response variable. While the overall mean is the weighted average of cell means (or the fitted value of an OLS regression at the sample mean of the covariates), it is not possible to calculate the aggregate as τth quantile cannot be calculated based on the conditional τth quantiles. Analogously, the estimated QR coefficient at a certain quantile does not reflect the effect of a uniform shift of one covariate by 1 unit on the aggregate quantile. Relating conditional and unconditional quantiles is important for so-called Blinder–Oaxaca type decomposition analysis. An example is the analysis of gender differences in the wage distribution, which may be due to gender differences in the distribution of covariates (characteristics) and gender differences in the QR coefficients. To distinguish the two effects, researchers estimate counterfactual wage distributions (e.g., the wage distribution resulting if females exhibited male characteristics and still female coefficients applied). This counterfactual distribution can be calculated by determining the conditional distribution function of the response variable based on the full process of QR coefficients, explicitly aggregating the conditional distribution functions (by the law of total probability) and then inverting the aggregate distribution function. This approach is described by Victor Chernozhukov, Ivan Fernandez-Val, and Blaise Melly, who also suggest its extension to sequential decompositions of the contribution of individual covariates to the differences in aggregate distributions. A simple alternative approach for decomposition analysis would make use of inverse probability weighting to calculate counterfactual wage distributions while balancing characteristics between males and females. This approach is very important in the literature on (unconditional) quantile treatment effects under the conditional independence assumption. A third alternative denoted as unconditional QR is introduced by Serjo Firpo, Nicole Fortin, and Thomas Lemieux. The basic idea is to estimate a discrete choice model of whether the individual response does not exceed the aggregate τth quantile. Based on the fitted probabilities and the density of the aggregate distribution at the aggregate τth quantile, it is possible to study the effect of uniform shifts in covariates on the aggregate τth quantile. The approach can be used for decomposition analysis. Estimating a linear probability model allows to implement a sequential decomposition. The literature is missing a comprehensive comparison of the different approaches to estimate counterfactual distributions. One approach requires the specification of conditional QRs, while the other approach requires the specification of the unconditional QR. Unconditional QR may prove useful in duration analysis. This idea has not yet been explored.

Endogeneity

A huge literature has emerged on the estimation of QR under endogeneity of some of the covariates. Just as OLS, QR estimation is inconsistent in such a situation, and various approaches using instrumental variables (in analogy to the two-stage least squares estimator of models of the conditional mean) have been explored, and most of them are described by Roger Koenker in Chapter 8.8 of his 2005 textbook (including the following ones). For the case with a discrete endogenous covariate and a discrete instrument, Alberto Abadie, Josh Angrist, and Guido Imbens (2002) have developed a weighted QR estimator, where the weights identify the compliers. Victor Chernozhukov and Christian Hansen (2008) use the fact that when the true QR coefficient of the endogenous covariate is known, the instrument should not affect the response variable. Andrew Chesher develops a general treatment of recursive structural equation models in a QR setting with continuous response variables and continuous endogenous covariates. He is concerned with the estimation of the causal effect of a change in the τ_2th quantile in the reduced form equation of an endogenous covariate on the τ_1th quantile of the response variable. These effects can be estimated as functions of nonparametric derivative estimates, and they can be aggregated, provided necessary support conditions hold, to more conventional causal effects. Instead, many researchers, including Roger Koenker, suggest control function approaches, where the residual of the reduced form regression of the endogenous covariate is used as an additional covariate in the structural equation of interest to control for the endogeneity. Our reading of the active literature on these topics suggests that the appropriate estimation approach for QR under endogeneity of some of the covariates may be very specific to the application of interest, because most approaches can easily become infeasible in realistic applications with more complex models. More guidance is needed here.

Panel Data Models

Whereas fixed-effects OLS regressions for longitudinal data provides consistent coefficient estimates for time-varying covariates, fixed-effects QR (i.e., QR with a dummy variable for each panel observation) suffers from the incidental parameter problem and do not provide consistent coefficient estimates for time-varying covariates. Furthermore, in an actual application it should be made sufficiently clear what it means to assume a separate individual-specific effect for each quantile of the conditional distribution. Effectively, a fixed-effects QR models the conditional distribution of the response variable around the individual-specific effect. Roger Koenker (as discussed in Section 8.7.2 of his 2005 textbook) suggests a common individual-specific effect for all quantiles, when the ~number of panel observations is small, and to implement a shrinkage estimator using an *l1*-penalty for the individual-specific effects. This way the estimation problem still involves a linear program and the estimation may set individual-specific effects to zero, depending on the size of the penalty. All individual-specific effects are set to zero for a sufficiently large penalty, while a sufficiently small penalty replicates fixed-effects QR. It remains an open question how to choose the size of the penalty. Furthermore, econometricians are concerned about the fact that the resulting coefficient estimates are not consistent for a fixed number of time periods. As an alternative to fixed-effects estimation, Jason Abrevaya and Christian Dahl suggest a correlated random-effects model for QR on panel data but similar conceptual issues arise for this model. More methodological research is needed regarding the proper modeling choices of individual-specific effects in QR. Given the state of literature, we advise applied researchers to clarify the role of individual-specific effects in QR in the context of their substantive question of interest and to avoid mechanical analogies to fixed-effects OLS estimation.

APPENDIX: SOFTWARE IMPLEMENTATIONS

See Table 1

Sources:

- R package quantreg: http://cran.r-project.org/web/packages/quantreg/quantreg.pdf
- R package rqpd: http://rqpd.r-forge.r-project.org
- R package cmprskQR: http://cran.r-project.org/web/packages/cmprskQR/index.html
- STATA package CQIV: http://ideas.repec.org/c/boc/bocode/s457478.html
- Stata codes for decomposition analysis: (i) Chernozhukov, Fernandez-Val, Melly: http://www.econ.brown.edu/fac/Blaise_Melly/code_counter.html; (ii) Firpo, Fortin, Lemieux: http://faculty.arts.ubc.ca/nfortin/datahead.html
- TSP: http://www.tspintl.com
- Matlab package: http://sites.stat.psu.edu/~dhunter/code/qrmatlab/
- Matlab function: https://www.mathworks.co.uk/matlabcentral/fileexchange/32115-quantreg-m-quantile-regression
- Matlab functions by C. Hansen: http://faculty.chicagobooth.edu/christian.hansen/research/#Code

REFERENCES

Abadie, A., Angrist, J., & Imbens, G. (2002). Instrumental variables estimates of the effect of subsidized training on the quantiles of trainee earnings. *Econometrica*, *70*(1), 91–117.

Chernozhukov, V., Fernández-Val, I., & Galichon, A. (2010). Quantile and probability curves without crossing. *Econometrica*, *78*(3), 1093–1125.

Chernozhukov, V., & Hansen, C. (2008). Instrumental variable quantile regression: A robust inference approach. *Journal of Econometrics*, *142*(1), 379–398.

Peng, L., & Huang, Y. (2008). Survival analysis with quantile regression models. *Journal of the American Statistical Association*, *103*, 637–649.

Portnoy, S. (2003). Censored regression quantiles. *Journal of the American Statistical Association*, *98*, 1001–1012.
Powell, J. L. (1986). Censored regression quantiles. *Journal of Econometrics*, *32*(1), 143–155.

FURTHER READING

Abrevaya, J., & Dahl, C. M. (2008). The effects of birth inputs on birthweight. *Journal of Business & Economic Statistics* (American Statistical Association), *26*, 379–397.
Chernozhukov, V., Fernandez-Val, I., & Melly, B. (2013). Inference on counterfactual distributions. *Econometrica*, *81*, 2205–2268.
Dette, H., & Volgushev, S. (2008). Non-crossing non-parametric estimates of quantile curves. *Journal of the Royal Statistical Society: Series B*, *70*, 609–627.
Fitzenberger, B., & Wilke, R. A. (2006). Using quantile regression for duration analysis. *Allgemeines Statistisches Archiv*, *90*, 105–120.
Fitzenberger, B., Wilke, R. A., & Zhang, X. (2010). Implementing Box–Cox quantile regression. *Econometric Reviews*, *29*(2), 158–181.
Firpo, S., Fortin, N. M., & Lemieux, T. (2009). Unconditional quantile regressions, *Econometrica*, *77*(3), 953–973.
Koenker, R. (2005). *Quantile regression*. Cambridge, England: Cambridge University Press.
Koenker, R. (2008). Censored quantile regression redux. *Journal of Statistical Software*, *27*, 1–24.
Koenker, R., & Bassett, G. (1978). Regression quantiles. *Econometrica*, *46*, 33–50.
Koenker, R., & Xiao, Z. (2006). Quantile autoregression. *Journal of the American Statistical Association*, *101*, 980–990.

BERND FITZENBERGER SHORT BIOGRAPHY

Bernd Fitzenberger has been a Full Professor of statistics and econometrics at the University of Freiburg since 2007, where he teaches econometrics and labor economics as well as an associate editor of “Empirical Economics”. He previously held positions at the universities of Dresden, Mannheim, and Frankfurt. He obtained his PhD in Economics, a Master of Science in Statistics from Stanford University, and his diploma in Economics from the University of Konstanz.

His research focuses on labor economics and microeconometric methods. Currently, he works on the following topics: wage structure, unions/wage bargaining, evaluation of active labor market policies, vocational training, school-to-work transitions, employment of mothers after first birth, and quantile regression. He is a research associate of the ZEW (Mannheim), a research fellow at IZA (Bonn) and at ROA (Maastricht), a research affiliate at the Institute for Fiscal Studies (London) as well as an editor of “Empirical

Economics" and a member of the editorial board of the "Journal of the European Economic Association" and the "Journal for Labour Market Research."

He has published around 45 per reviewed papers in international journals in Economics and Statistics and in edited volumes. He has been co-editor of six per reviewed special issues of international journals in Economics and Statistics.

RALF ANDREAS WILKE SHORT BIOGRAPHY

Ralf Andreas Wilke is a Full Professor of Applied Econometrics at Copenhagen Business School since 2014. From 2001 until 2014 he was a Reader at the Department of Economics and Related Studies at the University of York. Previously, he was an Associate Professor at the University of Nottingham (2011) and he held Lectureships at the Universities of Nottingham (2007–2011) and Leicester (2006–2007). He was a Research Economist at the Centre for European Economic Research in Mannheim (2002–2006). He obtained his PhD in Economics from the University of Dortmund in 2002. He has a Research Master Degree from the Toulouse School of Economics and he graduated from the University of Bonn.

He is a Research Professor at the Institute for Employment Research, Nuremberg, since 2013, and a Research Fellow of the Centre of European Economic Research, Mannheim, since 2009. He is an associate editor of "Empirical Economics".

He has published around 25 per reviewed papers in international journals in Economics and Statistics.

RELATED ESSAYS

Social Epigenetics: Incorporating Epigenetic Effects as Social Cause and Consequence *(Sociology)*, Douglas L. Anderton and Kathleen F. Arcaro

To Flop Is Human: Inventing Better Scientific Approaches to Anticipating Failure *(Methods)*, Robert Boruch and Alan Ruby

Repeated Cross-Sections in Survey Data *(Methods)*, Henry E. Brady and Richard Johnston

Ambulatory Assessment: Methods for Studying Everyday Life *(Methods)*, Tamlin S. Conner and Matthias R. Mehl

The Evidence-Based Practice Movement *(Sociology)*, Edward W. Gondolf

The Use of Geophysical Survey in Archaeology *(Methods)*, Timothy J. Horsley

Network Research Experiments *(Methods)*, Allen L. Linton and Betsy Sinclair

Longitudinal Data Analysis *(Methods)*, Todd D. Little *et al*.

Data Mining *(Methods)*, Gregg R. Murray and Anthony Scime