ORIGINAL PAPER

Quantile ratio regression

Alessio Farcomeni[1] · Marco Geraci[2,3]

Abstract

We introduce quantile ratio regression. Our proposed model assumes that the ratio of two arbitrary quantiles of a continuous response distribution is a function of a linear predictor. Thanks to basic quantile properties, estimation can be carried out on the scale of either the response or the link function. The advantage of using the latter becomes tangible when implementing fast optimizers for linear regression in the presence of large datasets. We show the theoretical properties of the estimator and derive an efficient method to obtain standard errors. The good performance and merit of our methods are illustrated by means of a simulation study and a real data analysis; where we investigate income inequality in the European Union (EU) using data from a sample of about two million households. We find a significant association between inequality, as measured by quantile ratios, and certain macroeconomic indicators; and we identify countries with outlying income inequality relative to the rest of the EU. An R implementation of the proposed methods is freely available.

Keywords Conditional quantiles · Distribution-free methods · Income inequality · L-statistics · Palma ratio · Poverty

1 Introduction

The ratio of quantiles of a distribution is an important measure of distributional features that finds application in several fields, most notably for the study of economic inequalities. Two popular measures are the ratio of the 80th and 20th income percentiles and the ratio of the 90th and 40th income percentiles. See also Cowell (2011), Chancel and Piketty (2021) and Chancel et al. (2022) for other quantile ratios of interest. The quartile ratio is found in biomedical

✉ Alessio Farcomeni
alessio.farcomeni@uniroma2.it

1 Department of Economics and Finance, University of Rome Tor Vergata, Rome, Italy

2 MEMOTEF Department, Sapienza University of Rome, Rome, Italy

3 Department of Epidemiology and Biostatistics, University of South Carolina, Columbia, SC, USA

applications (Feinberg et al. 2015), and the decile ratios in cost-effectiveness analysis (Kumar et al. 2018). Other applications of quantile ratios range from environmental (Rustomji et al. 2009) to ceramic materials studies (Das and Maiti 1999). To preempt any misunderstanding, we remark that this article does not deal with quantile *share* ratios (Beach and Davidson 1983; Langel and Tillé 2011). One such a ratio is the *Palma ratio*, that is, the ratio of the total income in the hands of the top 10% earners to that of the lowest 40%. As compared to quantile share ratios, quantile ratios are clearly more robust to outliers and, in our opinion, more readily interpretable.

Despite the widespread interest, quantile ratios have been used mostly as descriptive statistics. Few works that deal explicitly with inferential methods for quantile ratios are confined to the unconditional (as opposed to conditional) case. Prendergast and Staudte (2017) developed distribution-free asymptotic confidence intervals for unconditional ratios. In a related work, Prendergast and Staudte (2018) built upon (unconditional) quantile ratios to derive an inequality index that is akin to the Gini coefficient. A general and elegant non-parametric method to study wage inequality by means of an Oaxaca-Blinder-type decomposition is proposed in Firpo et al. (2018). On the other hand, a distribution-free method to make inference on conditional quantile ratios seems to be lacking, despite it being potentially useful in many applica-

tions (particularly for the study of economic inequalities); it is the objective of the present work to make such a proposal.

In general, the use of nonparametric bootstrap is quite common in the literature about inequality indices (ratios of income shares, Gini and Theil coefficients, and so on) (Biewen 2002; Buhmann et al. 1988). As far as conditional quantile ratios go, a poor man's solution for parametric estimation would be to fit a quantile regression model on the log-transformed outcome for the two quantile levels of interest and then take their difference (see Sect. 2.1 for more details). The calculation of standard errors would easily follow from well-established results for regression quantiles (Koenker and Bassett 1978a). However, this strategy would bring little flexibility to model's specification and, perhaps more importantly, it would always result in a loss of efficiency. The model would in general be misspecified and possibly therefore would not provide a good fit (Feng et al. 2023). Quantile crossing might also arise.

Another possible approach to the modeling of inequality would be a fully parametric one, say via (Generalized) Linear Models (GLMs). In one specification, given a distribution from the exponential family (e.g., gamma or normal), the expectation of the outcome is linked to the linear predictor. Conditional quantiles are subsequently estimated from the fitted GLM and their ratios are regressed against the same linear predictor upon a suitable transformation. In this modeling strategy, which we call GLM-based, bias and loss of efficiency may result from an incorrect distributional assumption. In a different specification, when some form of clustering is present (e.g., spatial), a naïve model would regress sample quantiles on aggregated covariates, potentially resulting in ecological bias. Related to the modeling of inequality via mean regression, see Olkin and Yitzhaki (1992).

In the present paper, we propose a distribution-free model for conditional quantile ratios and develop inferential methods in a parametric setting. Interpretation of the parameters is simple, while formal hypothesis testing is carried out with classical Wald-type tests. Our work is connected to inference for L-statistics, defined as linear combinations of order statistics (e.g., see Koenker 2005). Indeed, robust measures of location, quantile differences for assessing variability (e.g., interquartile range) (Goldman and Kaplan 2018) as well as other quantile-based measures of shape (Groeneveld and Meeden 1984; Groeneveld 1998; Jones et al. 2011) can ideally be linked to quantile ratios upon a logarithmic transformation. We also mention the studies of Dominici et al. (2005) and Cheng and Wu (2010) as unrelated to ours, since they concern the ratio of the same quantile in two different populations. Naturally, quantile regression is a very active area of research in general. We refer the interested reader to some recent contributions (Sherwood and Li 2022; Far-

comeni et al. 2022; Merlo et al. 2022; Redivo et al. 2023) and reviews (Koenker 2017; Waldmann 2018).

The present work is motivated by an original application to economic (income) inequality in Europe. The analysis is based on about two million households sampled in several cross-sectional waves of the European Union Survey on Income and Living Conditions (EU-SILC) database. We looked at inequality in 19 countries and how inequality has changed over a time period spanning the years before and after the Global Financial Crisis. We identified outlying countries that experienced more economic inequality than the rest of the European Union. We also investigated the role of macroeconomic indicators in explaining differences between countries. This is the first study to allow inferential statements about the (adjusted) association between these indicators and household-level income inequality.

In the following, we give a formal definition of the Quantile Ratio Regression (QRR) model (Sect. 2.1) and then propose an estimation procedure to fit QRR (Sect. 2.2), along with a scalable algorithm (Sect. 2.3); we derive standard errors for the parameter estimates (Sect. 2.4); we establish the theoretical properties of our methods (Sect. 3); we investigate the performance and behaviour of the proposed procedure via a simulation study (Sect. 4) and an analysis of the EU-SILC data (Sect. 5); and we conclude with final remarks (Sect. 6).

An R implementation of QRR estimation and related methods is freely available in the Qtools package (Geraci 2016). A brief (but exhaustive) tutorial is given in "Appendix".

2 Quantile ratio regression

2.1 The model

Let Y denote an absolutely continuous and strictly positive random variable with conditional cumulative distribution function (CDF) $F_Y(y|\boldsymbol{x}) = \Pr(Y \leq y \,|\boldsymbol{x})$, where $\boldsymbol{x}$ is a p-dimensional vector of covariates. Let

$$H_Y(\tau|\boldsymbol{x}) \equiv F_Y^{-1}(\tau|\boldsymbol{x}) \tag{1}$$

be the corresponding conditional quantile function. In classical quantile regression, the aim is to find a satisfactory model for H as a function of the covariates $\boldsymbol{x}$ for some $0 < \tau < 1$. Here our interest lies in the ratio of two conditional quantiles at some fixed levels $0 < \tau_2 < \tau_1 < 1$, with the assumption that $H_Y(\tau_1|\boldsymbol{x})/H_Y(\tau_2|\boldsymbol{x}) > 1$. We also assume that, for an appropriate known and monotone link function $g : (1, \infty) \rightarrow \mathbb{R}$, such a ratio is a linear function of the covariates and a parameter $\boldsymbol{\beta}$, namely

$$g\left\{\frac{H_Y(\tau_1|\boldsymbol{x})}{H_Y(\tau_2|\boldsymbol{x})}\right\} = \boldsymbol{x}^\top \boldsymbol{\beta}. \tag{2}$$

Though different choices for g are possible, our preference in this article is for $g(u) = \log(u - 1)$, which corresponds to an inverse logit transform since $\pm\text{logit}(1/u) = \pm \log(1/u/(1 - 1/u)) = \pm \log(1/(u-1))$. Not only does this transformation ensure that there is no quantile crossing between τ_2 and τ_1, but it also allows one to interpret each element of the vector $\boldsymbol{\beta}$ as an adjusted per unit log fold increase in the quantile ratio. We stress that we do not make any parametric assumption on the distribution of Y and that the model assumptions are *local*, namely, they involve only the specific quantile ratio but leave unspecified the rest of the conditional CDF.

We note that model (2) entails the following:

$$H_Y(\tau_1|\boldsymbol{x}) = H_Y(\tau_2|\boldsymbol{x}) \cdot g^{-1}(\boldsymbol{x}^\top\boldsymbol{\beta}),$$
$$H_Y(\tau_2|\boldsymbol{x}) = H_Y(\tau_1|\boldsymbol{x}) \cdot \{g^{-1}(\boldsymbol{x}^\top\boldsymbol{\beta})\}^{-1}. \tag{3}$$

More specifically, if $g(u) = \log(u - 1)$, then

$$H_Y(\tau_1|\boldsymbol{x}) = H_Y(\tau_2|\boldsymbol{x}) \cdot \{1 + \exp(\boldsymbol{x}^\top\boldsymbol{\beta})\}$$

and

$$H_Y(\tau_2|\boldsymbol{x}) = H_Y(\tau_1|\boldsymbol{x}) \cdot \{1 + \exp(\boldsymbol{x}^\top\boldsymbol{\beta})\}^{-1}$$

are two nonlinear conditional quantile equations (Koenker and Park 1996) that share the same parameter $\boldsymbol{\beta}$.

To appreciate how our proposal differs from a crude approach where the numerator and denominator are obtained from a common 'global' model for the conditional quantile function (1), it is perhaps most natural to consider an exponential model of the type $H_Y(\tau|\boldsymbol{x}) = \exp\{\boldsymbol{x}^\top\boldsymbol{\gamma}(\tau)\}$, in which case the transformation equivariance property of quantiles implies that

$$\log\left\{\frac{H_Y(\tau_1|\boldsymbol{x})}{H_Y(\tau_2|\boldsymbol{x})}\right\} = H_{\log(Y)}(\tau_1|\boldsymbol{x}) - H_{\log(Y)}(\tau_2|\boldsymbol{x}) = \boldsymbol{x}^\top\tilde{\boldsymbol{\beta}},$$

where $\tilde{\boldsymbol{\beta}} = \boldsymbol{\gamma}(\tau_1) - \boldsymbol{\gamma}(\tau_2)$ is the parameter of interest. Hence, it would be extremely easy to estimate $\tilde{\boldsymbol{\beta}}$ by fitting the same model on $\log Y$ for τ_1 and τ_2 using standard estimation routines for linear conditional quantiles. However, this model entails twice as many parameters as model (2). Standard errors of the final parameters would invariably be larger than those obtained with our proposal. A smaller issue with this crude approach is that it does not guarantee monotonicity of the quantiles. Should quantile crossing arise, it would need to be addressed using *ad hoc* solutions (e.g., He 1997; Chernozhukov et al. 2010) that may have limitations in terms of scope, and induce further bias.

2.2 Estimation

Let $\boldsymbol{y} = (y_1, \ldots, y_n)^\top$ denote a vector of n independent observations, associated with an $n \times p$ matrix of covariates $\boldsymbol{x}$ whose generic ith row vector is $\boldsymbol{x}_i$, $i = 1, \ldots, n$. Also, define $\xi_j \equiv H_Y(\tau_j|\boldsymbol{x})$, $j = 1, 2$, the true conditional quantile of Y and $\hat{\xi}_j$ its corresponding prediction. The parameter $\boldsymbol{\beta}$, common to the two equations in (3), can be estimated by either $H_Z(\tau_1|\boldsymbol{x}) = g^{-1}\{\boldsymbol{x}^\top\boldsymbol{\gamma}(\tau_1)\}$, where $Z = Y/\hat{\xi}_2$, or $H_Z(\tau_2|\boldsymbol{x}) = \{g^{-1}(\boldsymbol{x}^\top\boldsymbol{\gamma}(\tau_2))\}^{-1}$, where $Z = Y/\hat{\xi}_1$. Say we use the former. At first, the predictions $\hat{\xi}_2$ can be obtained by means of an appropriate method for estimating the conditional quantile function (1) such as linear, non-parametric (Yu and Jones 1998; Racine and Li 2017), or transformation-based (Geraci and Jones 2015) regression. In our experience (see Sect. 4) a linear model provides a satisfactory initial value $\hat{\xi}_2$ in the sense that the resulting estimate $\hat{\boldsymbol{\gamma}}(\tau_1)$ from the first model is generally a stationary point for the second model (see further below). If a refinement is needed, i.e., if $\hat{\boldsymbol{\gamma}}(\tau_1)$ is a not a solution for $H_Z(\tau_2|\boldsymbol{x})$, one could fit the latter with $Z = Y/\hat{\xi}_1$ where $\hat{\xi}_1 = g^{-1}\{\boldsymbol{x}^\top\hat{\boldsymbol{\gamma}}(\tau_1)\} \cdot \hat{\xi}_2$ is obtained from the first model. The new predictions $\hat{\xi}_2 = \{g^{-1}(\boldsymbol{x}^\top\hat{\boldsymbol{\gamma}}(\tau_2))\}^{-1} \cdot \hat{\xi}_1$ are subsequently fed into the first model and a new estimate $\hat{\boldsymbol{\gamma}}(\tau_1)$ is obtained. A single refinement step is usually sufficient, when needed.

We now illustrate why the solution $\hat{\boldsymbol{\gamma}}(\tau_1)$ is, in general, a stationary point for the fitting of $H_Z(\tau_2|\boldsymbol{x}) = \{g^{-1}(\boldsymbol{x}^\top\boldsymbol{\gamma}(\tau_2))\}^{-1}$. We consider a nonlinear g since not only is it more relevant to our discussion but also because the demonstration is trivial when g is linear. We first need to briefly expound the nonlinear estimation algorithm of Koenker and Park (1996), as adapted in Koenker (2005, p. 211) and implemented in the R function `nlrq` of the `quantreg` package (Koenker 2023). Consider the problem

$$\hat{\boldsymbol{\gamma}}_n(\tau_2) = \arg\min_{\boldsymbol{\gamma} \in \Gamma} \sum_{i=1}^n \rho_{\tau_2}(r_i(\boldsymbol{\gamma})), \tag{4}$$

where $\Gamma \subset \mathbb{R}^p$ is compact, $\rho_\tau(r) = r(\tau - I(r < 0))$, $r_i(\boldsymbol{\gamma}) = z_i - \{g^{-1}(\boldsymbol{x}_i^\top\boldsymbol{\gamma})\}^{-1}$ and $z_i = y_i/\hat{\xi}_{i,1}$. The solution (4) is obtained by solving, with respect to $\boldsymbol{\delta}_\gamma$, a sequence of locally linearized problems of the type $r(\boldsymbol{\gamma}) = \boldsymbol{j}(\boldsymbol{\gamma})\boldsymbol{\delta}_\gamma + \epsilon$, where $\boldsymbol{j}(\boldsymbol{\gamma}) = \partial r(\boldsymbol{\gamma})/\partial\boldsymbol{\gamma}$. The expression for the step $\boldsymbol{\delta}_\gamma$ is given in Koenker (2005). Then, the algorithm updates $\boldsymbol{\gamma}$ with $\boldsymbol{\gamma} + \lambda\boldsymbol{\delta}_\gamma$, where the step length $\lambda \in [0, 1]$ minimizes $\sum_{i=1}^n \rho_{\tau_2}(r_i(\boldsymbol{\gamma} + \lambda\boldsymbol{\delta}_\gamma))$. The algorithm proceeds until convergence (see Koenker and Park 1996, for the conditions and criteria of convergence).

Now, given the estimate $\hat{\boldsymbol{\gamma}}(\tau_1)$ and the predictions $\hat{\xi}_{i,1} = g^{-1}\{\boldsymbol{x}_i^\top\hat{\boldsymbol{\gamma}}(\tau_1)\} \cdot \hat{\xi}_{i,2}$ obtained from $H_Z(\tau_1|\boldsymbol{x}) = g^{-1}\{\boldsymbol{x}^\top\boldsymbol{\gamma}(\tau_1)\}$,

at the very first iteration of the algorithm to solve the minimization problem in (4), we have

$$
\sum_{i=1}^{n} \rho_{\tau_2}\left(r_i\left(\hat{\boldsymbol{\gamma}}(\tau_1) + \lambda \boldsymbol{\delta}_\gamma\right)\right)
$$

$$
= \sum_{i=1}^{n} \rho_{\tau_2}\left(y_i/\hat{\xi}_{i,1} - 1/g^{-1}\{\boldsymbol{x}_i^\top \hat{\boldsymbol{\gamma}}(\tau_1) + \lambda \boldsymbol{x}_i^\top \boldsymbol{\delta}_\gamma\}\right)
$$

$$
= \sum_{i=1}^{n} \rho_{\tau_2}\left(\frac{y_i/\hat{\xi}_{i,2}}{g^{-1}\{\boldsymbol{x}_i^\top \hat{\boldsymbol{\gamma}}(\tau_1)\}} - \frac{1}{g^{-1}\{\boldsymbol{x}_i^\top \hat{\boldsymbol{\gamma}}(\tau_1) + \lambda \boldsymbol{x}_i^\top \boldsymbol{\delta}_\gamma\}}\right). \tag{5}
$$

Note that $H_{Y/\hat{\xi}_2}(\tau_2|\boldsymbol{x}) = \xi_2/\hat{\xi}_2$, hence the empirical marginal quantile τ_2 of $Y/\hat{\xi}_2$ is, in general, equal to 1. This implies that (5) is minimized only when $\lambda = 0$ since

$$
\sum_{i=1}^{n} \rho_{\tau_2}\left(\frac{y_i/\hat{\xi}_{i,2}}{g^{-1}\{\boldsymbol{x}_i^\top \hat{\boldsymbol{\gamma}}(\tau_1)\}} - \frac{1}{g^{-1}\{\boldsymbol{x}_i^\top \hat{\boldsymbol{\gamma}}(\tau_1)\}}\right)
$$

$$
\propto \sum_{i=1}^{n} \rho_{\tau_2}\left(y_i/\hat{\xi}_{i,2} - 1\right),
$$

in which case the algorithm stops immediately and returns $\hat{\boldsymbol{\gamma}}(\tau_2) = \hat{\boldsymbol{\gamma}}(\tau_1)$. Of course, if the initial estimate $\hat{\xi}_2$ does not yield $\hat{H}_{Y/\hat{\xi}_2}(\tau_2) = 1$, the estimation will need to be refined as discussed above.

As far as asymptotics are concerned, for fixed ξ_j, $j = 1, 2$, the estimator $\hat{\boldsymbol{\gamma}}_n(\tau_j)$, $j = 2, 1$, respectively, is consistent and asymptotically normal (Koenker 2005). With slight abuse of notation, let $h = g^{-1}$ whenever h refers to an expression involving $\tau = \tau_1$ and $h = 1/g^{-1}$ when it involves $\tau = \tau_2$. Thanks to the Bahadur representation

$$
\sqrt{n}\left(\hat{\boldsymbol{\gamma}}_n(\tau_j) - \boldsymbol{\gamma}_0(\tau_j)\right) = \boldsymbol{D}_1^{-1} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \dot{h}_i \psi_{\tau_j}\left(r_i(\tau_j)\right)
$$

$$
+ o_p(1), \tag{6}
$$

where $\boldsymbol{D}_1 = n^{-1} \sum f_i\left(\xi_{i,j}\right) \dot{h}_i \dot{h}_i^\top$, f_i is the density of Z_i conditional on $\boldsymbol{x}_i$ and $\xi_{i,j} = H_{Z_i}(\tau_j|\boldsymbol{x}_i)$, $\dot{h}_i = \partial h\left(\boldsymbol{x}_i^\top \boldsymbol{\gamma}\right)/\partial \boldsymbol{\gamma}|_{\boldsymbol{\gamma}=\boldsymbol{\gamma}_0}$, $r_i(\tau_j) = y_i/\hat{\xi}_{i,j} - h\left(\boldsymbol{x}_i^\top \boldsymbol{\gamma}_0(\tau)\right)$ and $\psi_\tau(r) = \tau - I(r < 0)$.

2.3 A scalable estimation strategy

Given our choice for g, the estimation procedure introduced in the previous section requires solving a nonlinear problem, whose computation may represent a bottleneck. In order to scale inference to millions of observations we can linearize the problem by applying the equivariance property of quantiles. This is convenient due to the existence of fast computation algorithms for linear problems such as

Frisch–Newton-type algorithms (Koenker 2005) and algorithms based on smoothing (e.g., Bottai et al. 2015; Fernandes et al. 2021; He et al. 2023). In our real data example, which involves a sample of almost two million, we used the algorithm proposed by He et al. (2023).

For the linearly transformed estimation problem, we propose an estimation algorithm based on fitting the *linear* quantile model $H_{g(Z)}(\tau_1|\boldsymbol{x}) = \boldsymbol{x}^\top \boldsymbol{\gamma}(\tau_1)$, where $Z = Y/\hat{\xi}_2$ given the current value of $\hat{\xi}_2$, to obtain an estimate $\hat{\boldsymbol{\gamma}}(\tau_1)$ and predictions $\hat{\xi}_1 = g^{-1}\{\boldsymbol{x}^\top \hat{\boldsymbol{\gamma}}(\tau_1)\} \cdot \hat{\xi}_2$. As with the estimation procedure in Sect. 2.2, one should check (and possibly, refine) the starting solution by fitting the *linear* quantile model $H_{g(Z)}(1 - \tau_2|\boldsymbol{x}) = \boldsymbol{x}^\top \boldsymbol{\gamma}(1 - \tau_2)$, where $Z = \hat{\xi}_1/Y$ given the current value of $\hat{\xi}_1$, with estimates $\hat{\boldsymbol{\gamma}}(1 - \tau_2)$ and predictions $\hat{\xi}_2 = \{g^{-1}(\boldsymbol{x}^\top \hat{\boldsymbol{\gamma}}(1 - \tau_2))\}^{-1} \cdot \hat{\xi}_1$. For the latter estimation we have used the property $1/H_Y(\tau) = H_{1/Y}(1 - \tau)$ for a positive, continuous Y. Note that existence of the estimators is not guaranteed as it could happen that $Z \leq 1$. Luckily, population quantiles are invariant to censoring from below up to the quantile of interest. Therefore we can censor observations whenever $Y/\hat{\xi}_2 \leq 1$ and $\hat{\xi}_1/Y \leq 1$. An implicit assumption here is that the (population) quantiles at τ_1 and τ_2 are sufficiently separated, which is the rule in empirical applications.

The advantage of this approximate estimation resides in the fact that, rather than (4), we now deal with

$$
\hat{\boldsymbol{\gamma}}_n(\tau) = \arg\min_{\boldsymbol{\gamma} \in \mathbb{R}^p} \sum_{i=1}^{n} \rho_\tau\{g(z_i) - \boldsymbol{x}^\top \boldsymbol{\gamma}\}, \tag{7}
$$

where $z_i = y_i/\hat{\xi}_{i,2}$ for $\tau = \tau_1$ or $z_i = \hat{\xi}_{i,1}/y_i$ for $\tau = 1 - \tau_2$. Solving (7) can be accomplished by means of fast interior point algorithms. When this is computationally prohibitive, an alternative route is smoothing the loss function ρ_τ to achieve differentiability and thus apply gradient-based methods. For example, He et al. (2023) use convolution smoothing, which, among other things, has the advantage of yielding a strongly convex function (as opposed to other smoothing techniques such as that proposed by, e.g., Horowitz 1998).

If common regularity conditions for the linear quantile model hold (e.g., see Koenker 2005, p. 120), then

$$
\sqrt{n}\{\hat{\boldsymbol{\gamma}}_n(\tau_1) - \boldsymbol{\gamma}_0(\tau_1)\} \xrightarrow{d} \mathcal{N}(0, \boldsymbol{\Omega}_1), \tag{8}
$$

where $\boldsymbol{\Omega}_1 = \tau_1(1 - \tau_1)\boldsymbol{D}_1^{-1}\boldsymbol{D}_0\boldsymbol{D}_1^{-1}$, $\boldsymbol{D}_0 = \lim_{n\to\infty} n^{-1} \sum \boldsymbol{x}_i\boldsymbol{x}_i^\top$, $\boldsymbol{D}_1 = \lim_{n\to\infty} n^{-1} \sum f_i\left(\xi_{i,1}\right) \boldsymbol{x}_i \boldsymbol{x}_i^\top$, f_i is the density of Z_i conditional on $\boldsymbol{x}_i$ and $\xi_{i,1} = H_{Z_i}(\tau_1|\boldsymbol{x}_i)$, with $Z_i = Y_i/\hat{\xi}_{i,2}$. Similarly, for $\hat{\boldsymbol{\gamma}}(1 - \tau_2)$ we have

$$
\sqrt{n}\{\hat{\boldsymbol{\gamma}}_n(1 - \tau_2) - \boldsymbol{\gamma}_0(1 - \tau_2)\} \xrightarrow{d} \mathcal{N}(0, \boldsymbol{\Omega}_2), \tag{9}
$$

where $\Omega_2 = \tau_2(1 - \tau_2)D_1^{-1}D_0D_1^{-1}$, $D_0 = \lim_{n\to\infty} n^{-1} \sum x_i x_i^\top$,
$D_1 = \lim_{n\to\infty} n^{-1} \sum f_i(\xi_{i,2}) x_i x_i^\top$, f_i is the density of Z_i conditional on x_i and $\xi_{i,2} = H_{Z_i}(1 - \tau_2|x_i)$, with $Z_i = \hat{\xi}_{i,1}/Y_i$.

2.4 Standard errors

Standard error estimation can be carried out by means of an ordinary xy-pair bootstrap, a method we investigated in the simulation study of Sect. 4 and which proved to work well even at smaller sample sizes. Alternatives include smoother approaches, like the Bayesian bootstrap (e.g., Hahn 1997). However, bootstrap may not be feasible when the sample size is in the order of the hundreds of thousands or more. We derive a method for direct estimation of the standard errors that scales well and has a satisfactory performance, especially at larger sample sizes, as illustrated in Sect. 4.

Starting from the linearized version of our algorithm (Sect. 2.3), we smooth the objective functions through an appropriate kernel so as to make them differentiable and, ultimately, lay the groundwork for scalability. Here we show expressions based on a Gaussian kernel, though in principle other kernels can be used. Conditionally on $\hat{\xi}_2$, we have that the gradient of the smoothed objective function is given by

$$
\frac{\partial}{\partial \gamma} \rho_{\tau_1}^{(s)}(Z - x^\top \gamma)
$$
$$
= x^\top \Phi\left(\frac{x^\top \gamma(\tau_1) - \log(Y/\hat{\xi}_2 - 1)}{b} \right), \tag{10}
$$

where Φ denotes the CDF of the standard normal and b is a smoothing parameter (or bandwidth). Analogously, the gradient of the smoothed objective function conditionally on $\hat{\xi}_1$ is given by

$$
\frac{\partial}{\partial \gamma} \rho_{1-\tau_2}^{(s)}(Z - x^\top \gamma)
$$
$$
= x^\top \Phi\left(\frac{x^\top \gamma(1 - \tau_2) - \log(\hat{\xi}_1/Y - 1)}{b} \right). \tag{11}
$$

The fact that $\hat{\gamma}$ is a stationary point for both equations allows us to approximate the smoothed gradient as the sum of (10) and (11). The information matrix is then obtained through numerical differentiation of this approximated p-dimensional gradient, which is rather quick. Standard errors are obtained as the square root of the diagonal elements of its inverse. As for the bandwidth, in our implementation we set $b = ((\log(n) + p)/n)^{0.4}$, which minimizes the Gaussian approximation error at each step according to He et al. (2023).

3 Theoretical properties

In this section we show consistency and asymptotic normality of the estimator. With the same abuse of notation as before, we let $h = g^{-1}$ whenever h refers to an expression involving $\tau = \tau_1$ and $h = 1/g^{-1}$ when it involves $\tau = \tau_2$. Also, let X denote the design matrix for all observations. For reasons of space we restrict our theoretical results to the nonlinear algorithm in Sect. 2.2, where the linear approximation with convolution smoothing in Sect. 2.3 can be accommodated with some slight changes.

Theorem 1 *Assume the usual regularity conditions for cross-sectional quantile regression models, and that $X^\top X/n$ is positive definite and converges to a positive definite matrix C. Assume initial predictions are consistent. Then, $\hat{\gamma}$ converges in probability to β.*

Proof By definition,

$$
\beta = \lim_n \arg\min_\gamma \frac{\sum_{i=1}^n \rho_{\tau_2}\left(\frac{y_i}{\xi_{i,1}} - h(x_i, \gamma) \right)}{n}
$$
$$
= \lim_n \arg\min_\gamma \frac{\sum_{i=1}^n \rho_{\tau_1}\left(\frac{y_i}{\xi_{i,2}} - h(x_i, \gamma) \right)}{n}. \tag{12}
$$

Similarly, for all n,

$$
\hat{\gamma} = \arg\min_\gamma \frac{\sum_{i=1}^n \rho_{\tau_2}\left(\frac{y_i}{\hat{\xi}_{i,1}} - h(x_i, \gamma) \right)}{n}
$$
$$
= \arg\min_\gamma \frac{\sum_{i=1}^n \rho_{\tau_1}\left(\frac{y_i}{\hat{\xi}_{i,2}} - h(x_i, \gamma) \right)}{n}. \tag{13}
$$

To see consistency, note that the regularity conditions imply that $n^{-1} \sum_{i=1}^n \rho_{\tau_2}\left(\frac{y_i}{\hat{\xi}_{i,1}} - h(x_i, \hat{\gamma}) \right)$ uniformly converges in probability to a finite constant. The uniform convergence theorem then guarantees that there exists a vector η such that $\hat{\gamma}$ converges in probability to η.

Suppose now that $\eta \neq \beta$. This would imply that $\exists \tilde{n}$ such that for $n > \tilde{n}$

$$
\sum_{i=1}^n \rho_{\tau_2}\left(\frac{y_i}{\hat{\xi}_{i,1}} - h(x_i, \beta) \right) > \sum_{i=1}^n \rho_{\tau_2}\left(\frac{y_i}{\hat{\xi}_{i,1}} - h(x_i, \hat{\gamma}) \right),
$$

which is not possible due to consistency of initial predictions, (12) and (13). Therefore, $\eta = \beta$, which completes the proof. □

Theorem 2 *Assume the usual regularity conditions for quantile regression models, and that $X^\top X/n$ is positive definite and converges to a positive definite matrix C. Assume initial predictions are consistent. Then, there exists a positive definite matrix Ω such that $\sqrt{n}(\hat{\gamma} - \beta)$ converges in distribution*

to a zero-centered multivariate Gaussian random variable with covariance matrix $\boldsymbol{\Omega}$.

Proof The statement follows from the fact that, for $j, k = 1, 2$ and $j \neq k$, $(\hat{\boldsymbol{\gamma}}(\tau_j), \hat{\xi}_{i,j})$ are asymptotically jointly normal conditional on $\hat{\xi}_{i,k}$, as argued in the previous section. Equality of $\hat{\boldsymbol{\gamma}}(\tau_1)$ and $\hat{\boldsymbol{\gamma}}(\tau_2)$ guarantees that $\hat{\boldsymbol{\gamma}}$ is asymptotically normal also unconditionally. □

As a general comment, it can be seen that theoretical properties of quantile ratio regression, including the rate of convergence, are very similar to those of classical quantile regression.

4 Simulation study

We conducted a simulation study to investigate the finite-sample properties of our methods. We considered the conditional quantile ratio model (2) with two response distributions: Weibull and lognormal. The former has quantile function $H_Y(\tau) = \sigma\{-\log(1-\tau)\}^{1/\alpha}$ for $Y \geq 0, \sigma > 0, \alpha > 0$, while the latter $H_Y(\tau) = \exp\{\mu + \sigma\Phi^{-1}(\tau)\}$ for $Y \geq 0$, $\mu \in \mathbb{R}, \sigma > 0$. Each Weibull response Y_i, $i = 1, \ldots, n$, was generated conditionally on $\eta_i = \beta_0 + \beta_1 x_i + \beta_2 w_i$, with $x_i \sim \text{Unif}(-0.5, 0.5)$ and $w_i \sim \text{Bin}(1, 0.4)$, by sampling a distribution with shape parameter α_i satisfying

$$\frac{H_{Y_i}(\tau_1)}{H_{Y_i}(\tau_2)} = \left\{\frac{\log(1-\tau_1)}{\log(1-\tau_2)}\right\}^{1/\alpha_i} = 1 + \exp(\eta_i);$$

while σ was set equal to 1. Analogously, the ith lognormal response was obtained with a value of σ_i so as to satisfy

$$\frac{H_{Y_i}(\tau_1)}{H_{Y_i}(\tau_2)} = \exp\left\{\sigma_i\left(\Phi^{-1}(\tau_1) - \Phi^{-1}(\tau_2)\right)\right\} = 1 + \exp(\eta_i),$$

with μ set equal to 0. Hence, since α_i and σ_i depend on η_i, the location, scale and shape of either distribution are affected by the covariates $\boldsymbol{x}_i$. We used two sets of values for the regression coefficients, $\boldsymbol{\beta} = (0.5, 0, 0)^\top$ and $\boldsymbol{\beta} = (0.5, -0.5, 0.5)^\top$, four quantile ratios, $0.8 : 0.2$, $0.9 : 0.1$, $0.9 : 0.4$ and $0.99 : 0.5$, and three sample sizes, $n \in \{100, 1000, 10000\}$, for a total of 24 combinations for each distribution. Both the Weibull and lognormal distribution are right-skewed for any of the values η_i resulting from these settings. Each combination was replicated 500 times.

The estimation algorithms discussed in Sects. 2.2 and 2.3 were assessed in terms of bias and mean squared error (MSE) of $\hat{\boldsymbol{\beta}}$, while the performance of the approximate standard error estimation approach of Sect. 2.4 was evaluated by calculating the ratio between the average estimated standard errors and the standard deviation of the estimates (empirical standard errors), along with the observed coverage of the

corresponding 95% confidence intervals. For comparison, we also estimated standard errors by means of ordinary bootstrap with 200 replications. Given the large number of results, we report detailed results obtained from the linearized algorithm in Sect. 2.3 for the quantile ratios $0.8 : 0.2$ and $0.9 : 0.4$ only, and briefly summarize the others.

Due to lack of direct competitors to our proposal, we considered two parametric alternatives as briefly introduced in Sect. 1, namely (i) a GLM-based estimation and (ii) a Naïve linear regression model applied to artificially grouped data. In the former case (i), the conditional quantiles $\hat{H}_{MLE}(\tau_1)$ and $\hat{H}_{MLE}(\tau_2)$ were obtained from the inversion of $\hat{F}_{MLE}$, a parametric distribution in the GLM family fitted by means of maximum likelihood. Then $\log(\hat{R}_{MLE} - 1)$, where $\hat{R}_{MLE} = \hat{H}_{MLE}(\tau_1)/\hat{H}_{MLE}(\tau_2)$, was regressed on $\boldsymbol{x}$ to estimate $\boldsymbol{\beta}_{MLE}$, which can be considered a parametric estimate of $\boldsymbol{\beta}$. In our simulation study, we used gamma and normal GLMs. In the second case (ii), by partitioning the observations into 20 groups of size $n/20$, each with index set $\mathcal{G}_g$, $g = 1, \ldots, 20$, the sample was aggregated so as to regress $\log(\hat{R}_g - 1)$ on $\bar{\boldsymbol{x}}_g$, where $\hat{R}_g = \hat{H}_g(\tau_1)/\hat{H}_g(\tau_2)$ is the ratio of the sample quantiles for group g and $\bar{\boldsymbol{x}}_g = 20\sum_{i\in\mathcal{G}_g}\boldsymbol{x}_i/n$. An ecological regression of this sort is clearly bound to be unsatisfactory, but it is arguably a 'quick and dirty' expedient one could put together to obtain an indirect estimate of $\boldsymbol{\beta}$, say $\boldsymbol{\beta}_g$. Note anyway that our real data example in Sect. 5 does have a grouping structure.

As shown in Tables 1 and 2, the bias and MSE of the QRR estimator under the Weibull distribution were reasonably small at the smallest sample size and practically null at higher sample sizes in all considered scenarios. Approximate-to-empirical average standard error ratios were also close to 1 (Table 3), with observed confidence interval coverage close to the nominal 95% or slightly conservative at most for the $0.8 : 0.2$ quantile ratio. Tables 4, 5 and 6 confirm the good behaviour of QRR also under the lognormal distribution, with the exception of the $0.9 : 0.4$ quantile ratio. For this scenario, the approximate standard errors were slightly below the empirical standard errors and, consequently, observed confidence interval coverage below the nominal level. This lesser performance at this quantile ratio is most likely due to the markedly lower density that the lognormal distribution has at its 90th percentile for a given η_i, compared to the density of the Weibull at its 90th percentile for the same value of η_i. On the other hand, the bootstrap standard errors were accurate estimates of the empirical standard errors, consistently under both distributions.

The results (not shown) for the quantile ratios $0.9 : 0.1$ and $0.99 : 0.5$ were in line with those for the other quantile ratios in terms of bias and MSE. However, the performance of the approximate estimation of the standard errors tended to deteriorate especially in the most extreme case $0.99 : 0.5$.

Table 1 Weibull—Bias for crude generalized linear model (gamma, GLMg, and normal, GLMn), naïve linear regression (LM), and quantile ratio regression (QRR)

Quantile ratio = 0.8 : 0.2

	$\beta_0 = 0.5$				$\beta_1 = 0$				$\beta_2 = 0$			
Sample size	GLMg	GLMn	LM	QRR	GLMg	GLMn	LM	QRR	GLMg	GLMn	LM	QRR
100	−0.01	−0.06	−0.53	−0.02	−0.00	−0.00	−0.02	0.01	0.00	0.00	−0.04	−0.02
1000	0.01	−0.07	−0.03	−0.00	0.00	0.00	0.10	−0.01	0.00	0.00	−0.06	−0.01
10,000	0.02	−0.07	−0.00	−0.00	0.00	−0.00	−0.07	−0.00	0.00	0.00	−0.01	−0.00

	$\beta_0 = 0.5$				$\beta_1 = -0.5$				$\beta_2 = 0.5$			
Sample size	GLMg	GLMn	LM	QRR	GLMg	GLMn	LM	QRR	GLMg	GLMn	LM	QRR
100	0.20	0.21	−0.53	−0.03	0.50	0.51	0.18	0.08	−0.50	−0.55	−0.09	−0.00
1000	0.23	0.19	−0.10	−0.00	0.50	0.54	−0.02	0.00	−0.50	−0.54	0.08	−0.01
10,000	0.24	0.20	−0.02	0.00	0.50	0.54	−0.09	−0.00	−0.50	−0.54	0.01	0.00

Quantile ratio = 0.9 : 0.4

	$\beta_0 = 0.5$				$\beta_1 = 0$				$\beta_2 = 0$			
Sample size	GLMg	GLMn	LM	QRR	GLMg	GLMn	LM	QRR	GLMg	GLMn	LM	QRR
100	0.05	−0.31	−0.52	−0.01	−0.00	0.02	0.01	−0.03	−0.00	0.01	−0.08	−0.04
1000	0.09	−0.30	−0.08	−0.00	−0.00	0.01	−0.04	−0.00	−0.00	−0.00	0.04	0.00
10,000	0.09	−0.30	0.01	0.00	0.00	−0.00	0.06	−0.00	0.00	−0.00	−0.03	−0.00

	$\beta_0 = 0.5$				$\beta_1 = -0.5$				$\beta_2 = 0.5$			
Sample size	GLMg	GLMn	LM	QRR	GLMg	GLMn	LM	QRR	GLMg	GLMn	LM	QRR
100	0.25	−0.11	−0.54	−0.02	0.50	0.56	−0.00	−0.02	−0.50	−0.55	−0.09	−0.04
1000	0.30	−0.09	−0.07	0.00	0.50	0.57	0.01	−0.00	−0.50	−0.58	−0.03	−0.00
10,000	0.30	−0.09	−0.02	0.00	0.50	0.57	0.02	0.00	−0.50	−0.57	−0.01	−0.00

Table 2 Weibull—Mean squared error for crude generalized linear model (gamma, GLMg, and normal, GLMn), naïve linear regression (LM), and quantile ratio regression (QRR)

Quantile ratio = 0.8 : 0.2

	$\beta_0 = 0.5$				$\beta_1 = 0$				$\beta_2 = 0$			
Sample size	GLMg	GLMn	LM	QRR	GLMg	GLMn	LM	QRR	GLMg	GLMn	LM	QRR
100	0.02	0.02	0.40	0.06	0.00	0.11	2.07	0.41	0.00	0.04	0.61	0.14
1000	0.00	0.01	0.13	0.01	0.00	0.01	2.35	0.04	0.00	0.00	0.81	0.01
10,000	0.00	0.01	0.13	0.00	0.00	0.00	2.30	0.00	0.00	0.00	0.83	0.00

	$\beta_0 = 0.5$				$\beta_1 = -0.5$				$\beta_2 = 0.5$			
Sample size	GLMg	GLMn	LM	QRR	GLMg	GLMn	LM	QRR	GLMg	GLMn	LM	QRR
100	0.06	0.10	0.41	0.06	0.25	0.50	1.81	0.46	0.25	0.39	0.66	0.17
1000	0.06	0.04	0.17	0.01	0.25	0.31	2.31	0.04	0.25	0.29	0.96	0.01
10,000	0.06	0.04	0.15	0.00	0.25	0.29	2.88	0.00	0.25	0.29	0.93	0.00

Quantile ratio = 0.9 : 0.4

	$\beta_0 = 0.5$				$\beta_1 = 0$				$\beta_2 = 0$			
Sample size	GLMg	GLMn	LM	QRR	GLMg	GLMn	LM	QRR	GLMg	GLMn	LM	QRR
100	0.02	0.11	0.42	0.06	0.00	0.08	2.07	0.43	0.00	0.03	0.71	0.16
1000	0.01	0.09	0.14	0.01	0.00	0.01	2.42	0.04	0.00	0.00	0.85	0.01

Table 2 continued

Quantile ratio = 0.9 : 0.4											
	$\beta_0 = 0.5$				$\beta_1 = 0$				$\beta_2 = 0$		
Sample size	GLMg	GLMn	LM	QRR	GLMg	GLMn	LM	QRR	GLMg	GLMn	LM	QRR
10,000	0.01	0.09	0.13	0.00	0.00	0.00	2.71	0.00	0.00	0.00	0.83	0.00

	$\beta_0 = 0.5$				$\beta_1 = -0.5$				$\beta_2 = 0.5$		
Sample size	GLMg	GLMn	LM	QRR	GLMg	GLMn	LM	QRR	GLMg	GLMn	LM	QRR
100	0.08	0.04	0.42	0.05	0.25	0.43	1.76	0.50	0.25	0.35	0.69	0.15
1000	0.09	0.01	0.19	0.01	0.25	0.33	2.65	0.05	0.25	0.34	1.13	0.02
10,000	0.09	0.01	0.16	0.00	0.25	0.33	2.94	0.00	0.25	0.33	0.99	0.00

Table 3 Weibull—Standard errors for quantile ratio regression (QRR)

Quantile ratio = 0.8 : 0.2											
$\beta = (0.5, 0, 0)$											
	ESE			ASE/ESE			BSE/ESE			Acov		
Sample size	β_0	β_1	β_2	β_0	β_1	β_2	β_0	β_1	β_2	β_0	β_1	β_2
100	0.24	0.64	0.37	1.19	1.21	1.22	1.12	1.19	1.17	0.98	0.99	0.98
1000	0.07	0.19	0.12	1.20	1.21	1.12	1.06	1.07	0.99	0.98	0.99	0.96
10,000	0.02	0.06	0.04	1.14	1.16	1.15	0.99	1.01	1.00	0.98	0.97	0.98
$\beta = (0.5, -0.5, 0.5)$											
	ESE			ASE/ESE			BSE/ESE			Acov		
Sample size	β_0	β_1	β_2	β_0	β_1	β_2	β_0	β_1	β_2	β_0	β_1	β_2
100	0.24	0.67	0.41	1.19	1.17	1.14	1.12	1.19	1.15	0.97	0.96	0.98
1000	0.07	0.21	0.12	1.19	1.14	1.17	1.05	1.06	1.09	0.98	0.97	0.98
10,000	0.02	0.07	0.04	1.16	1.10	1.11	1.00	0.98	1.00	0.97	0.97	0.97
Quantile ratio = 0.9 : 0.4											
$\beta = (0.5, 0, 0)$											
	ESE			ASE/ESE			BSE/ESE			Acov		
Sample size	β_0	β_1	β_2	β_0	β_1	β_2	β_0	β_1	β_2	β_0	β_1	β_2
100	0.25	0.66	0.40	1.01	1.03	1.00	1.12	1.20	1.14	0.95	0.95	0.95
1000	0.08	0.19	0.12	0.97	1.07	0.99	1.00	1.13	1.03	0.93	0.96	0.95
10,000	0.02	0.06	0.04	1.00	1.01	0.93	1.00	1.02	0.94	0.95	0.95	0.94
$\beta = (0.5, -0.5, 0.5)$											
	ESE			ASE/ESE			BSE/ESE			Acov		
Sample size	β_0	β_1	β_2	β_0	β_1	β_2	β_0	β_1	β_2	β_0	β_1	β_2
100	0.23	0.71	0.39	1.10	0.96	1.06	1.22	1.16	1.24	0.96	0.94	0.95
1000	0.07	0.21	0.12	1.04	0.97	1.00	1.09	1.05	1.10	0.95	0.93	0.95
10,000	0.02	0.06	0.04	0.96	1.02	0.94	0.97	1.06	1.00	0.94	0.96	0.93

ESE: empirical standard errors; ASE: approximate standard errors; BSE: bootstrap standard errors; Acov: coverage at nominal 95% for ASE

As expected, neither the GLM-based estimator nor the naïve linear regression estimator performed well, as evidenced by the much larger MSE values in all settings. It is worth noting that the apparent optimality of GLM-based estimation in some of the instances of the scenario with $\beta = (0.5, 0, 0)^\top$ is actually the result of parallel estimated quantiles $\hat{H}_{MLE}(\tau)$ that give constant ratios across the values of x. This behavior persists in the other scenario,

$\beta = (0.5, -0.5, 0.5)^\top$, resulting in an evident bias that does not give any sign to vanish with increasing sample sizes.

The nonlinear estimation procedure for QRR, as described in Sect. 2.2, yielded bias and MSE comparable to those reported above. As anticipated, the initial value $\hat{\xi}_2$, which was obtained using a standard linear model, resulted in an estimate of $\gamma(\tau_1)$ that was also a stationary point for the estimation of $\gamma(\tau_2)$ in 90–100% of the cases across scenarios.

Table 4 Lognormal—Bias for crude generalized linear model (gamma, GLMg, and normal, GLMn), naïve linear regression (LM), and quantile ratio regression (QRR)

Quantile ratio = 0.8 : 0.2

Sample size	$\beta_0 = 0.5$				$\beta_1 = 0$				$\beta_2 = 0$			
	GLMg	GLMn	LM	QRR	GLMg	GLMn	LM	QRR	GLMg	GLMn	LM	QRR
100	0.01	0.30	− 0.52	− 0.03	0.00	− 0.00	0.01	− 0.03	− 0.00	0.03	0.01	− 0.00
1000	0.04	0.32	− 0.07	− 0.01	− 0.00	0.01	− 0.04	− 0.00	0.00	− 0.00	0.02	0.00
10,000	0.04	0.32	− 0.00	− 0.00	0.00	0.00	0.02	− 0.00	− 0.00	− 0.00	− 0.01	0.00

Sample size	$\beta_0 = 0.5$				$\beta_1 = -0.5$				$\beta_2 = 0.5$			
	GLMg	GLMn	LM	QRR	GLMg	GLMn	LM	QRR	GLMg	GLMn	LM	QRR
100	0.24	1.15	− 0.48	− 0.02	0.50	1.05	− 0.01	− 0.04	− 0.50	− 1.00	− 0.11	− 0.01
1000	0.29	1.09	− 0.07	0.00	0.50	0.97	− 0.06	0.00	− 0.50	− 0.93	− 0.00	− 0.01
10,000	0.30	1.10	− 0.04	0.00	0.50	0.93	− 0.12	− 0.00	− 0.50	− 0.92	0.04	− 0.00

Quantile ratio = 0.9 : 0.4

Sample size	$\beta_0 = 0.5$				$\beta_1 = 0$				$\beta_2 = 0$			
	GLMg	GLMn	LM	QRR	GLMg	GLMn	LM	QRR	GLMg	GLMn	LM	QRR
100	− 0.15	− 0.25	− 0.63	− 0.03	− 0.00	0.03	0.00	0.01	− 0.00	0.01	0.06	− 0.03
1000	− 0.12	− 0.23	− 0.07	− 0.01	− 0.00	0.01	0.02	0.00	0.00	0.01	− 0.02	0.00
10,000	− 0.12	− 0.23	− 0.01	0.00	− 0.00	− 0.00	− 0.04	− 0.00	0.00	0.00	0.00	− 0.00

Sample size	$\beta_0 = 0.5$				$\beta_1 = -0.5$				$\beta_2 = 0.5$			
	GLMg	GLMn	LM	QRR	GLMg	GLMn	LM	QRR	GLMg	GLMn	LM	QRR
100	0.07	0.13	− 0.57	− 0.01	0.50	0.70	− 0.01	0.03	− 0.50	− 0.72	− 0.08	− 0.02
1000	0.10	0.19	− 0.09	− 0.00	0.50	0.72	− 0.03	0.00	− 0.50	− 0.73	0.04	− 0.01
10,000	0.10	0.19	− 0.01	− 0.00	0.50	0.72	− 0.08	0.00	− 0.50	− 0.72	− 0.01	0.00

Table 5 Lognormal—Mean squared error for crude generalized linear model (gamma, GLMg, and normal, GLMn), naïve linear regression (LM), and quantile ratio regression (QRR)

Quantile ratio = 0.8:0.2

Sample size	$\beta_0 = 0.5$				$\beta_1 = 0$				$\beta_2 = 0$			
	GLMg	GLMn	LM	QRR	GLMg	GLMn	LM	QRR	GLMg	GLMn	LM	QRR
100	0.02	0.18	0.38	0.04	0.00	0.34	1.66	0.34	0.00	0.10	0.55	0.11
1000	0.00	0.11	0.11	0.00	0.00	0.02	1.89	0.03	0.00	0.01	0.61	0.01
10,000	0.00	0.10	0.09	0.00	0.00	0.00	2.00	0.00	0.00	0.00	0.58	0.00

Sample size	$\beta_0 = 0.5$				$\beta_1 = -0.5$				$\beta_2 = 0.5$			
	GLMg	GLMn	LM	QRR	GLMg	GLMn	LM	QRR	GLMg	GLMn	LM	QRR
100	0.08	2.25	0.34	0.05	0.25	2.78	1.77	0.39	0.25	1.55	0.57	0.15
1000	0.09	1.34	0.11	0.00	0.25	1.12	1.93	0.03	0.25	0.93	0.64	0.01
10,000	0.09	1.22	0.12	0.00	0.25	0.88	2.26	0.00	0.25	0.85	0.74	0.00

Quantile ratio = 0.9:0.4

Sample size	$\beta_0 = 0.5$				$\beta_1 = 0$				$\beta_2 = 0$			
	GLMg	GLMn	LM	QRR	GLMg	GLMn	LM	QRR	GLMg	GLMn	LM	QRR
100	0.03	0.09	0.54	0.06	0.00	0.09	1.97	0.41	0.00	0.03	0.64	0.15
1000	0.02	0.06	0.15	0.01	0.00	0.01	2.54	0.04	0.00	0.00	0.86	0.01

Table 5 continued

Quantile ratio = 0.9:0.4											
Sample size	$\beta_0 = 0.5$				$\beta_1 = 0$				$\beta_2 = 0$			
	GLMg	GLMn	LM	QRR	GLMg	GLMn	LM	QRR	GLMg	GLMn	LM	QRR
10,000	0.01	0.05	0.16	0.00	0.00	0.00	2.76	0.00	0.00	0.00	0.98	0.00

Sample size	$\beta_0 = 0.5$				$\beta_1 = -0.5$				$\beta_2 = 0.5$			
	GLMg	GLMn	LM	QRR	GLMg	GLMn	LM	QRR	GLMg	GLMn	LM	QRR
100	0.02	0.14	0.48	0.07	0.25	0.72	2.21	0.52	0.25	0.59	0.78	0.19
1000	0.01	0.06	0.15	0.01	0.25	0.53	2.67	0.05	0.25	0.53	0.91	0.02
10,000	0.01	0.04	0.16	0.00	0.25	0.51	2.64	0.00	0.25	0.52	0.99	0.00

Table 6 Lognormal—standard errors for quantile ratio regression (QRR)

Quantile ratio = 0.8:0.2 $\beta = (0.5, 0, 0)$											
Sample size	ESE			ASE/ESE			BSE/ESE			Acov		
	β_0	β_1	β_2	β_0	β_1	β_2	β_0	β_1	β_2	β_0	β_1	β_2
100	0.19	0.58	0.33	1.30	1.18	1.22	1.24	1.18	1.19	0.99	0.98	0.98
1000	0.07	0.17	0.10	1.14	1.17	1.18	1.04	1.09	1.09	0.98	0.98	0.98
10,000	0.02	0.06	0.03	1.09	1.11	1.10	0.99	1.01	1.00	0.96	0.97	0.96
Quantile ratio = 0.8 : 0.2 $\beta = (0.5, -0.5, 0.5)$											
Sample size	ESE			ASE/ESE			BSE/ESE			Acov		
	β_0	β_1	β_2	β_0	β_1	β_2	β_0	β_1	β_2	β_0	β_1	β_2
100	0.22	0.62	0.38	1.14	1.12	1.07	1.13	1.17	1.11	0.97	0.98	0.98
1000	0.07	0.19	0.12	1.14	1.12	1.06	1.05	1.06	1.03	0.96	0.97	0.95
10,000	0.02	0.06	0.03	1.15	1.08	1.12	1.04	1.01	1.06	0.97	0.96	0.97
Quantile ratio = 0.9 : 0.4 $\beta = (0.5, 0, 0)$											
Sample size	ESE			ASE/ESE			BSE/ESE			Acov		
	β_0	β_1	β_2	β_0	β_1	β_2	β_0	β_1	β_2	β_0	β_1	β_2
100	0.24	0.64	0.39	0.92	0.94	0.92	1.17	1.24	1.18	0.91	0.95	0.92
1000	0.08	0.21	0.12	0.86	0.87	0.88	1.02	1.05	1.06	0.91	0.92	0.91
10,000	0.02	0.07	0.04	0.87	0.85	0.86	1.01	1.00	1.00	0.90	0.91	0.91
Quantile ratio = 0.9 : 0.4 $\beta = (0.5, -0.5, 0.5)$											
Sample size	ESE			ASE/ESE			BSE/ESE			Acov		
	β_0	β_1	β_2	β_0	β_1	β_2	β_0	β_1	β_2	β_0	β_1	β_2
100	0.26	0.72	0.43	0.86	0.85	0.85	1.10	1.16	1.13	0.92	0.89	0.90
1000	0.08	0.22	0.14	0.88	0.84	0.80	1.05	1.04	1.01	0.91	0.89	0.88
10,000	0.02	0.07	0.04	0.90	0.87	0.80	1.05	1.05	0.98	0.91	0.90	0.88

ESE: Empirical standard errors; ASE: approximate standard errors; BSE: bootstrap standard errors; Acov: coverage at nominal 95% for ASE

In less than 2% of the cases, a single refinement step was not sufficient.

In summary, QRR has little bias and MSE in all the scenarios and under all distributions considered in this simulation, whereas approximate standard error estimation as proposed in Sect. 2.4 performs best when the upper quantile of the ratio is not too high on the tail. In general, bootstrap standard errors are a valid alternative.

Table 7 Number of households (n), four sample quantiles of equivalized disposable income (EDI) and two quantile ratios by country for the EU-SILC data

Country	n	EDI sample quantile				Quantile ratio	
		0.2	0.4	0.8	0.9	0.8:0.2	0.9:0.4
Belgium	35,055	11,079	14,706	24,670	29,941	2.23	2.04
Croatia	48,735	2908	4288	8058	10,283	2.77	2.40
Czechia	110,901	4624	5865	9338	11,592	2.02	1.98
Denmark	81,757	19,464	24,871	38,289	45,943	1.97	1.85
France	45,301	14,990	19,590	31,780	40,470	2.12	2.07
Germany	172,393	12,340	16,733	28,341	35,189	2.30	2.10
Hungary	118,274	2993	3864	6212	7663	2.08	1.98
Ireland	73,110	12,257	16,199	31,397	40,000	2.56	2.47
Italy	288,014	9720	13,906	24,563	31,122	2.53	2.24
Netherlands	135,607	15,621	19,610	30,921	37,629	1.98	1.92
Norway	83,629	24,550	32,158	50,700	61,172	2.07	1.90
Poland	174,687	2625	3763	7017	9059	2.67	2.41
Portugal	80,677	4904	6858	13,250	18,093	2.70	2.64
Romania	83,436	1205	1815	3414	4352	2.83	2.40
Serbia	29,053	1218	1959	3906	5058	3.21	2.58
Slovak Republic	69,643	3518	4969	8287	10,117	2.36	2.04
Sweden	90,956	14,799	19,530	31,830	38,578	2.15	1.98
Switzerland	30,144	28,367	38,591	64,816	81,395	2.28	2.11
United Kingdom	124,995	11,774	16,100	29,800	38,234	2.53	2.37
Overall	1,876,367	4676	9756	26,395	35,327	5.65	3.62

5 Data analysis

We describe in this section an original application of our methods to economic inequality in Europe. We pooled data from several waves of the cross-sectional component of the European Union Survey on Income and Living Conditions (EU-SILC), where the equivalized disposable income (EDI) was collected at household level. For this analysis, we extracted observations on EDI for the 14-year time period going from 2004 to 2017. We then linked these data to several country-time specific macroeconomic indicators obtained from the World Bank data repository, namely unemployment rate, trade (sum of imports and exports values) as percentage of GDP; and from the Fraser Institute data repository, namely standardized indexes of hiring and firing regulations, and government transfers and subsides. The latter two variables range between zero and ten, with larger values corresponding to less regulations for hiring and dismissal of workers, and less government subsides, respectively. For details see Gwartney and Lawson (2003). There were several other variables available from these data sources that we considered, but that we did not include in the reported analysis, as they were either irrelevant to inequality or too strongly associated with other variables more readily interpretable. Moreover, we had to exclude some EU countries as they did not have income data (Cyprus, Estonia, Latvia, Lithuania, Luxembourg, Malta, Slovenia) or had income data limited to only

1 year (Greece). Finally, out of the remaining 19 countries included in the analysis, only 5 had data for the entire time period, while 5 had data for 8 years or less, thus making the assessment of temporal trajectories somehow limited.

It is important to remark that the data set is a repeated cross-sectional survey, since each household has been sampled only once (in a specific year). There is no repeated sampling, which makes the assumption that data are independent reasonable.

Number of households and selected sample quantiles for the outcome EDI, along with sample quantile ratios 0.8:0.2 (quintile ratio) and 0.9:0.4 (which, for practical purposes, we refer to as Palma ratio) are shown for each country and overall in Table 7. After exclusions, the sample consisted of 1,876,367 households in total. Croatia, Poland, Serbia and Portugal were among the countries with stronger inequality according to either quantile ratio; while Denmark, Norway, Netherlands and Czechia those with lower inequality. The mean (standard deviation) of unemployment rate, trade, hiring and dismissal regulations, and transfers and subsides were, 8.1 (3.4), 97.1 (40.4), 4.1 (1.4), and 4.5 (1.1), respectively.

We ran a QRR analysis on the quintile and Palma ratios using Eq. (2) with link function $g(u) = \log(u - 1)$. For each ratio, we fitted two models: Model 1, which included 18 dummy variables x_j, $j = 1, \ldots, 18$, for country (reference: Belgium) and a second degree polynomial for year t (cen-

Table 8 Estimates, standard errors (SE) and p-values from the 0.8:0.2 quantiles ratio regressions for the EU-SILC data

| | Model 1 | | | Model 2 | | |
	Estimate	SE	p-value	Estimate	SE	p-value
Intercept	0.240	0.010	< 0.001	0.388	0.030	< 0.001
Belgium (reference)	–	–	–	–	–	–
Croatia	0.466	0.013	< 0.001	0.245	0.023	< 0.001
Czechia	− 0.244	0.011	< 0.001	− 0.311	0.014	< 0.001
Denmark	− 0.193	0.011	< 0.001	− 0.445	0.019	< 0.001
France	− 0.010	0.013	0.448	− 0.209	0.022	< 0.001
Germany	0.111	0.011	< 0.001	− 0.050	0.017	0.002
Hungary	0.147	0.011	< 0.001	0.006	0.014	0.681
Ireland	0.290	0.011	< 0.001	0.288	0.017	< 0.001
Italy	0.259	0.010	< 0.001	0.033	0.022	0.131
Netherlands	− 0.169	0.010	< 0.001	− 0.211	0.011	< 0.001
Norway	− 0.096	0.011	< 0.001	− 0.335	0.020	< 0.001
Poland	0.274	0.011	< 0.001	0.057	0.020	0.004
Portugal	0.411	0.012	< 0.001	0.208	0.021	< 0.001
Romania	0.769	0.012	< 0.001	0.274	0.025	< 0.001
Serbia	0.689	0.016	< 0.001	0.476	0.023	< 0.001
Slovak Republic	− 0.198	0.012	< 0.001	− 0.174	0.015	< 0.001
Sweden	− 0.035	0.011	0.001	− 0.237	0.018	< 0.001
Switzerland	0.128	0.015	< 0.001	− 0.126	0.023	< 0.001
United Kingdom	0.288	0.011	< 0.001	− 0.039	0.025	0.115
Year	− 0.031	0.001	< 0.001	− 0.023	0.002	< 0.001
Year^2	0.002	0.000	< 0.001	0.001	0.000	< 0.001
Unemployment				0.004	0.001	< 0.001
Trade				− 0.002	0.000	< 0.001
Hiring and dismissal regulations				0.028	0.002	< 0.001
Transfers and subsides				0.017	0.004	< 0.001

tered at 2004); Model 2, which, in addition to the covariates of the first model, included the macroeconomic indicators introduced above, w_j, $j = 1, \ldots, 4$. Let $\boldsymbol{x}$ denote the vector collecting all the aforementioned variables. Then Model 2 can be written as

$$
\log\left\{ \frac{H_Y(\tau_1|\boldsymbol{x})}{H_Y(\tau_2|\boldsymbol{x})} - 1 \right\}
$$
$$
= \beta_0 + \sum_{j=1}^{18} x_j\beta_j + \beta_{19}t + \beta_{20}t^2 + w_1\beta_{21}
$$
$$
+ w_2\beta_{22} + w_3\beta_{23} + w_4\beta_{24},
$$

while Model 1 can be specified as above with $\beta_k = 0$, $k = 21, \ldots, 24$. The quantity $\exp(\beta_k)$, $k = 1, \ldots, 24$ thus represents the adjusted multiplicative effect of the kth covariate on the relative inequality $H_Y(\tau_1|\boldsymbol{x})/H_Y(\tau_2|\boldsymbol{x}) - 1$, while $\exp(\beta_0)$ is the baseline relative inequality. Standard errors were estimated as detailed in Sect. 2.4. (We also computed bootstrap standard errors and there were no appreciable dif-

ferences with approximate standard errors, probably due to the very large sample size.)

The results are reported in Tables 8 (quintile ratio) and 9 (Palma ratio). We will now make general statements about the results with no specific reference to either ratio, as the results are often similar. We will then point out and discuss differences.

There were clear differences between countries according to Model 1, with higher inequality in Croatia, Portugal, Romania, and Serbia, but lower inequality in Czechia, Netherlands, Slovak Republic and Nordic countries. Some differences persisted after adjusting for the macroeconomic indicators (e.g., Ireland), while other differences decreased in magnitude (e.g., Italy and Poland), or increased in magnitude (e.g., Denmark and Sweden). That is, differences in income inequality among countries are partly confounded by country-level fundamentals.

Unemployment is significantly associated with inequality, as one would expect. On the other hand, higher levels of trade seem to be associated with lower income inequality. Interestingly, higher levels of transfers and subsides by the

Table 9 Estimates, standard errors (SE) and p-values from the 0.9:0.4 quantiles ratio regressions for the EU-SILC data

	Model 1			Model 2		
	Estimate	SE	p-value	Estimate	SE	p-value
Intercept	0.117	0.009	< 0.001	-0.066	0.027	0.013
Belgium (reference)	–	–	–	–	–	–
Croatia	0.335	0.012	< 0.001	0.160	0.020	< 0.001
Czechia	-0.105	0.010	< 0.001	-0.184	0.013	< 0.001
Denmark	-0.227	0.010	< 0.001	-0.450	0.016	< 0.001
France	0.062	0.011	< 0.001	0.035	0.020	0.072
Germany	0.061	0.010	< 0.001	0.024	0.015	0.105
Hungary	0.093	0.010	0.409	-0.060	0.013	< 0.001
Ireland	0.359	0.010	< 0.001	0.236	0.016	< 0.001
Italy	0.167	0.009	< 0.001	0.058	0.019	0.003
Netherlands	-0.122	0.010	< 0.001	-0.119	0.010	< 0.001
Norway	-0.153	0.010	< 0.001	-0.330	0.019	< 0.001
Poland	0.230	0.010	< 0.001	0.031	0.018	0.082
Portugal	0.495	0.010	< 0.001	0.361	0.019	< 0.001
Romania	0.710	0.010	< 0.001	0.262	0.022	< 0.001
Serbia	0.466	0.014	< 0.001	0.310	0.019	< 0.001
Slovak Republic	-0.134	0.010	< 0.001	-0.167	0.013	< 0.001
Sweden	-0.105	0.010	< 0.001	-0.257	0.016	< 0.001
Switzerland	0.103	0.013	< 0.001	-0.187	0.020	< 0.001
United Kingdom	0.299	0.010	< 0.001	-0.003	0.022	0.886
Year	-0.026	0.001	< 0.001	-0.015	0.001	< 0.001
Year^2	0.001	0.000	< 0.001	0.001	0.000	< 0.001
Unemployment				0.005	0.001	< 0.001
Trade				-0.003	0.000	< 0.001
Hiring and dismissal regulations				0.033	0.002	< 0.001
Transfers and subsides				0.055	0.003	< 0.001

government and stricter regulations for hiring and dismissal of employees are both independently associated with lower income inequality.

As for the time trend, we note that both models suggest that inequality initially declined and then regained ground. To determine, approximately, at what point in time this occurred, we can simply consider that

$$\frac{\partial}{\partial t} \frac{H_Y(\tau_1|\boldsymbol{x})}{H_Y(\tau_2|\boldsymbol{x})} = (\beta_{19} + 2\beta_{20}t) \exp(\boldsymbol{x}^\top \boldsymbol{\beta}).$$

If we use the estimates from Model 1 in Table 8, the expression above is zero at $t = 7.75$, that is, towards the end of 2011. So while income inequality as measured by these quantile ratios had been improving in the EU area as of 2004, the conditions somewhat worsened after 2011, which we speculate to be partly explained by the lagged effects of the financial and economic crisis that started in 2007–2008.

6 Final remarks

To our knowledge, this work represents the first formal proposal for Quantile Ratio Regression (QRR). The methods developed herein are distribution free and can be scaled up to analyse large collections of data, as those obtained from multi-country surveys, for example. We believe that QRR is sufficiently versatile to find application and generate interest in many research areas.

We have proposed two strategies for estimation. The first, presented in Sect. 2.2, starts from the nonlinear specification of the model and works well for small to moderately sized data sets. An alternative strategy based on linearization is presented in Sect. 2.3 and is suitable for very large data sets (with a sample size in the order of millions). Both strategies performed well in a simulation study.

In our motivating application the data were obtained from a series of cross-sectional surveys, and the assumption that observations are independent is tenable. A possible venue for future research is the extension of our methods to panel data, as are the extensions to censored and discrete outcomes (e.g.,

Geraci and Farcomeni 2022). Given the recent interest and developments in causal inference for conditional quantiles (e.g., see Koenker et al. 2018), we note that instrumental variables (Chernozhukov and Hansen 2005, 2006) seem to us a natural candidate for conducting causal analysis in QRR, although we believe that the identification of causal effects in this highly non-linear setting might not be elementary.

We conclude by circling back to our opening statements, saying that another open route for further work is devising a methodology for performing *quantile share ratio* regression, where the summary statistic is not a ratio of quantiles (as in this work), but a ratio of shares above specific quantiles.

A Quantile ratio regression in R

In this appendix, we briefly illustrate the function `qrr` as provided in the `Qtools` package (Geraci 2016) in R. The reader is referred to function's documentation for all the details.

The `qrr` function requires the `formula` argument, as commonly specified with the response on the left of a '$\sim$' operator, and the terms, separated by '$+$' operators, on the right; the `data` argument, a data frame in which to interpret the variables named in the formula; and `taus`, a vector containing the two quantile indices for the quantile ratio of interest (the order of the quantiles is irrelevant). All other arguments are set by default and can be modified as appropriate. Estimation as in Sect. 2.2 is obtained with `linearize = FALSE`, whereas `linearize = TRUE` (the default) gives the large-sample algorithm discussed in Sect. 2.3. The `Qtools` package also provides S3 methods functions such as `summary`, `coef`, `predict` and `vcov` that are commonly used to summarize and extract information from fitted objects.

In the example further below, the data are generated according to the Weibull model in Sect. 4 with a single covariate x drawn from a uniform distribution. The formula $y \sim x$ in `qrr` specifies the model

$$\log\left\{\frac{H_Y(\tau_1|x)}{H_Y(\tau_2|x)} - 1\right\} = \beta_0 + \beta_1 x$$

that is to be fitted. Note that the only link function implemented in the package (currently version 1.5.9) is $g(u) = \log(u-1)$ but other transformations may be implemented in future versions of the package.

```
# generate data from a Weibull model
library(Qtools)
set.seed(123)
n <- 10000
x <- runif(n, -0.5, 0.5)
R <- 1 + exp(0.5 + x)
tau1 <- 0.8
tau2 <- 0.2
alpha <- 1/log(R)*log(log(1-tau1)
   /log(1-tau2))
y <- rweibull(n, shape = alpha, scale = 1)
dd <- data.frame(x = x, y = y)

# fit quintile ratio regression
fit <- qrr(y ~ x, data = dd,
    taus = c(tau1,tau2))
```

A call to `summary` yields the familiar regression output in R.

```
> summary(fit)
> Quantile ratio regression 0.8:0.2

> Coefficients linear predictor:
>             Estimate  Std.Err t value
        Pr(>|t|)
> (Intercept) 0.502843 0.020798
     24.177 < 2.2e-16 ***
> x           1.020963 0.071823
  14.215 < 2.2e-16 ***
> ---
> Signif. codes:  0 '***' 0.001 '**'
  0.01 '*' 0.05 '.' 0.1 ' ' 1

> Degrees of freedom: 10000 total; 9998
  residual
```

We conclude this tutorial by showing the poor results one obtains when using a GLM-based estimator (with normal family) on these data.

```
fit.glm <- glm(y ~ x, data = dd,
     family = gaussian)
m <- predict(fit.glm, type = "response")
sigma <- sqrt(fit.glm$deviance/fit.glm$df.
     residual)
Qhat1 <- qnorm(tau1, mean = m, sd = sigma)
Qhat2 <- qnorm(tau2, mean = m, sd = sigma)
z <- log(Qhat1/Qhat2 - 1)

> lm(z ~ x, data = dd)

> Coefficients:
> (Intercept)              x
>   0.45606        -0.08369
```

References

Beach, C.M., Davidson, R.: Distribution-free statistical inference with Lorenz curves and income shares. Rev. Econ. Stud. **50**(4), 723–735 (1983)

Biewen, M.: Bootstrap inference for inequality, mobility and poverty measurement. J. Economet. **108**, 317–342 (2002)

Bottai, M., Orsini, N., Geraci, M.: A gradient search maximization algorithm for the asymmetric Laplace likelihood. J. Stat. Comput. Simul. **85**, 1919–1925 (2015)

Buhmann, B., Rainwater, L., Schmaus, G., Smeeding, T.M.: Equivalence scales, well-being, inequality, and poverty: sensitivity estimates across ten counties using the Luxemburg Income Study (LIS) database. Rev. Income Wealth **34**, 115–142 (1988)

Chancel, L., Piketty, T.: Global income inequality, 1820–2020: the persistence and mutation of extreme inequality. J. Eur. Econ. Assoc. **19**, 3025–3062 (2021)

Chancel, L., Piketty, T., Saez, E., Zucman, G.: World Inequality Report 2022. Technical Report, World Inequality Lab, Paris (2022)

Cheng, C., Wu, J.: Interval estimation of quantile ratios applied to anti-cancer drug screening by xenograft experiments. Stat. Med. **29**(26), 2669–2678 (2010)

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

Chernozhukov, V., Hansen, C.: An IV model of quantile treatment effects. Econometrica **73**(1), 245–261 (2005)

Chernozhukov, V., Hansen, C.: Instrumental quantile regression inference for structural and treatment effect models. J. Economet. **132**(2), 491–525 (2006)

Cowell, F.: Measuring Inequality. Oxford University Press, Oxford (2011)

Das, N., Maiti, H.: Effect of size distribution of the starting powder on the pore size and its distribution of tape cast alumina microporous membranes. J. Eur. Ceram. Soc. **19**(3), 341–345 (1999)

Dominici, F., Cope, L., Naiman, D.Q., Zeger, S.L.: Smooth quantile ratio estimation. Biometrika **92**(3), 543–557 (2005)

Farcomeni, A., Geraci, M., Viroli, C.: Directional quantile classifiers. J. Comput. Graph. Stat. **31**, 907–916 (2022)

Feinberg, A.E., Porter, J., Saskin, R., Rangrej, J., Urbach, D.R.: Regional variation in the use of surgery in Ontario. CMAJ Open **3**, 310–316 (2015)

Feng, X., Liu, Q., Wang, C.: A lack-of-fit test for quantile regression process models. Stat. Probab. Lett. **192**, 109680 (2023)

Fernandes, M., Guerre, E., Horta, E.: Smoothing quantile regressions. J. Bus. Econ. Stat. **39**, 338–357 (2021)

Firpo, S., Fortin, N.M., Lemieux, T.: Decomposing wage distributions using recentered influence function regression. Econometrics **6**, 28 (2018)

Geraci, M.: Qtools: A collection of models and other tools for quantile inference. R J. **8**(2), 117–138 (2016)

Geraci, M., Farcomeni, A.: Mid-quantile regression for discrete responses. Stat. Methods Med. Res. **31**, 821–838 (2022)

Geraci, M., Jones, M.C.: Improved transformation-based quantile regression. Can. J. Stat. **43**(1), 118–132 (2015)

Goldman, M., Kaplan, D.M.: Non-parametric inference on (conditional) quantile differences and interquantile ranges, using L-statistics. Economet. J. **21**, 136–169 (2018)

Groeneveld, R.A.: A class of quantile measures for kurtosis. Am. Stat. **52**(4), 325–329 (1998)

Groeneveld, R.A., Meeden, G.: Measuring skewness and kurtosis. J. R. Stat. Soc. D **33**(4), 391–399 (1984)

Gwartney, J., Lawson, R.: The concept and measurement of economic freedom. Eur. J. Polit. Econ. **19**, 405–430 (2003)

Hahn, J.: Bayesian bootstrap of the quantile regression estimator: a large sample study. Int. Econ. Rev. **38**, 795–808 (1997)

He, X.: Quantile curves without crossing. Am. Stat. **51**(2), 186–192 (1997)

He, X., Pan, X., Tan, K.M., Zhou, W.-X.: Smoothed quantile regression with large-scale inference. J. Economet. **232**(2), 367–388 (2023)

Horowitz, J.L.: Bootstrap methods for median regression models. Econometrica **66**(6), 1327–1351 (1998)

Jones, M.C., Rosco, J.F., Pewsey, A.: Skewness-invariant measures of kurtosis. Am. Stat. **65**(2), 89–95 (2011)

Koenker, R.: Quantile Regression. Cambridge University Press, New York (2005)

Koenker, R.: Quantile regression: 40 years on. Annu. Rev. Econ. **9**, 155–176 (2017)

Koenker, R.: quantreg: Quantile regression. R Package Version **5**, 97 (2023)

Koenker, R., Bassett, G.: Regression quantiles. Econometrica **46**(1), 33–50 (1978)

Koenker, R., Chernozhukov, V., He, X., Peng, L.: Handbook of quantile regression. In: Handbook of Modern Statistical Methods. Chapman & Hall/CRC, Boca Raton (2018)

Koenker, R., Park, B.J.: An interior point algorithm for nonlinear quantile regression. J. Economet. **71**(1–2), 265–283 (1996)

Kumar, V., Cohen, J.T., van Klaveren, D., Soeteman, D.I., Wong, J.B., Neumann, P.J., Kent, D.M.: Risk-targeted lung cancer screening: a cost-effectiveness analysis. Ann. Intern. Med. **168**, 161–169 (2018)

Langel, M., Tillé, Y.: Statistical inference for the quintile share ratio. J. Stat. Plan. Inference **141**, 2976–2985 (2011)

Merlo, L., Maruotti, A., Petrella, L., Punzo, A.: Quantile hidden semi-Markov models for multivariate time series. Stat. Comput. **32**, 61 (2022)

Olkin, I., Yitzhaki, S.: Gini regression analysis. Int. Stat. Rev. **60**(2), 185–196 (1992)

Prendergast, L.A., Staudte, R.G.: When large *n* is not enough—distribution-free interval estimators for ratios of quantiles. J. Econ. Inequal. **15**, 277–293 (2017)

Prendergast, L.A., Staudte, R.G.: A simple and effective inequality measure. Am. Stat. **72**, 328–343 (2018)

Racine, J.S., Li, K.: Nonparametric conditional quantile estimation: a locally weighted quantile kernel approach. J. Economet. **201**, 72–94 (2017)

Redivo, E., Viroli, C., Farcomeni, A.: Quantile-distribution functions and their use for classification, with application to naïve Bayes classifiers. Stat. Comput. **33**, 55 (2023)

Rustomji, P., Bennett, N., Chiew, F.: Flood variability east of Australia's great dividing range. J. Hydrol. **374**(3), 196–208 (2009)

Sherwood, B., Li, S.: Quantile regression feature selection and estimation with grouped variables using Huber approximation. Stat. Comput. **32**, 75 (2022)

Waldmann, E.: Quantile regression: a short story on how and why. Stat. Model. **18**, 203–218 (2018)

Yu, K., Jones, M.C.: Local linear quantile regression. J. Am. Stat. Assoc. **93**, 228–237 (1998)

Springer