



A New Framework for Estimation of Unconditional Quantile Treatment Effects: The Residualized Quantile Regression (RQR) Model

Nicolai Borgen^{1,2} , Andreas Haupt³  and Øyvind Wiborg⁴

Abstract

The opportunities for understanding how treatment effects vary across different segments of the population have led to a rise in the use of quantile regressions for identifying unconditional quantile treatment effects (QTEs). However, existing quantile regression models fall into two categories: those that are unsuitable for identifying unconditional QTEs and those that often struggle with the complex data structures common in sociology and other social sciences. In particular, existing methods face difficulties with large data sets and high-dimensional fixed effects. The authors introduce a two-step approach to estimating unconditional QTEs, which is easy to use and aligns with the needs of sociologists. First, the treatment variable is decomposed into a systematic and random part, and then, the random variation in the treatment status is used as the sole independent variable in a quantile regression model. Through a series of simulations and three empirical applications, the authors provide strong evidence that the residualized quantile regression (RQR) approach provides approximately unbiased estimates of unconditional QTEs comparable with existing methods. Moreover, the RQR approach offers greater flexibility and enhances computational speed compared with existing models, and it can easily handle high-dimensional fixed effects. In sum, the RQR approach fills a pressing void in quantitative research methodology, offering a much-needed tool for studying treatment effect heterogeneity.

Keywords

quantile regression, quantile treatment effect, residual regression, residualized quantile regression, fixed effects

¹Center for Research on Equality in Education, University of Oslo, Oslo, Norway

²Department of Special Needs Education, University of Oslo, Oslo, Norway

³Department of Sociology and Computational Social Science, Karlsruhe Institute of Technology, Karlsruhe, Germany

⁴Department of Sociology and Human Geography, University of Oslo, Oslo, Norway

Corresponding Author:

Nicolai Borgen, Center for Research on Equality in Education, University of Oslo, PO Box 1140, Blindern, Oslo, 0318, Norway.

Email: n.t.borgen@isp.uio.no

Studying differences between groups has historically been akin to looking at differences in means. However, researchers increasingly turn to quantile regression models to get a complete view of how treatments affect outcomes (Koenker 2005). One advantage of quantile regression models over standard linear regression models is that we can study how differences between groups vary across the outcome variable's distribution, allowing researchers to explore new types of research questions. Generally, whereas linear regression models enable us to examine how the average of the outcome differs between groups, quantile regression models allow us to study how quantile values differ (Firpo 2007). Thus, quantile regression models can be used to analyze how individuals prone to have high outcomes react differently to treatment than do individuals with low propensity; this is called unconditional quantile treatment effects (QTEs). We can, for example, investigate whether the motherhood wage penalty is more pronounced for women at the upper end of the wage distribution compared with women at the lower end (Killewald and Bearak 2014).

Historically, the conditional quantile regression (CQR) model, which builds on Roger Koenker and colleagues' work in the mid-1970s, has been used to estimate quantile regression coefficients (Koenker 2017). CQR coefficients can be interpreted as unconditional QTEs whenever we do not need to include any control variables in our model (e.g., randomized treatment). However, unlike in linear regression models, including control variables changes the interpretation of the CQR coefficients, and they can no longer be interpreted as unconditional QTEs (Borgen, Haupt, and Wiborg 2023; Firpo 2007; Killewald and Bearak 2014; Rios-Avila and Maroto 2024; Wenz 2019). Therefore, solutions that allow the inclusion of control variables in quantile regression models while simultaneously preserving the coefficients' interpretation as unconditional QTEs are being developed (Firpo 2007; Frölich and Melly 2010; Powell 2020).

This article adds to this growing literature by offering a new quantile treatment estimation method, called residualized quantile regression (RQR), which complements existing approaches. In his seminal article, Firpo (2007) proposed an elegant solution to estimate unconditional QTEs with a single binary treatment variable using a propensity score matching framework. However, the propensity score framework cannot be used with nonbinary treatment variables, and including fixed effects is problematic in a propensity score framework. Recently, Powell (2020) developed the generalized quantile regression (GQR) model that allows for nonbinary treatment variables.¹ However, this method is computationally demanding, with computational issues growing with the model's complexity and the sample size. Thus, including high-dimensional fixed effects in large administrative data sets is challenging or practically impossible using the GQR model. As we will show, our estimation method can easily handle large data sets and complex model specifications.

The RQR model builds on the fact that a simple CQR model can estimate the unconditional QTE with a randomized treatment. In social sciences, assignment of the treatment status often rests on a selection process. However, under the assumption of selection on observables, the treatment assignment can be decomposed into a systematic part caused by the observed confounders and a random part. By using only the random part of the treatment assignment in a quantile regression model, we can straightforwardly estimate unconditional QTEs. More specifically, our approach involves a two-step method. First, we decompose the treatment variable into systematic and residual components. In most cases, a simple linear regression model, regressing the treatment on the observed confounders, will be sufficient, but other regression models can be used to account for the specific sorting patterns into treatment. The residuals derived from the first-step model can act as the as-if randomized treatment assignment. Next, the outcome variable is regressed on the residualized treatment variable using a CQR model with only the residualized treatment as the independent variable. Because the control variables

purge the treatment of confounding in the first step, they are redundant in the second step. Thus, our approach serves as a straightforward solution to estimate unconditional QTEs in the presence of covariates.

The RQR model has several advantages over current QTE approaches, making it a valuable addition to the quantile regression toolkit. Its most significant contribution lies in its ability to incorporate high-dimensional fixed effects into the estimation of unconditional QTEs, a critical feature for identifying the motherhood wage penalty and numerous other treatment effects. Additionally, the estimation procedure is computationally efficient, the model allows for both binary and nonbinary treatment variables, and it is straightforward to implement the RQR model in all software that provides a package for CQR or linear programming. Borgen, Haupt, and Wiborg (2021) introduce the `rq` and `rqplot` Stata commands that estimate and plot RQR coefficients.

The RQR model belongs to a different class of models than the popular unconditional quantile regression (UQR) model (Firpo, Fortin, and Lemieux 2009). The distinction between these models is discussed in detail by Rios-Avila and Maroto (2024) and Borgen et al. (2023). Here, we briefly clarify their difference. The UQR model approximates the marginal effects of independent variables on unconditional quantile values through a two-step approach. First, the quantile values of the outcome distribution are reduced to a fixed set of binary outcome variables via the recentered influence function. Then, these binary outcome variables are regressed on the independent variables using separate linear regression models. The UQR model is often used to identify QTEs, but it was developed to infer how independent variables influence overall quantile values and should be applied cautiously when examining QTEs of binary variables (Borgen et al. 2023). Within the UQR framework, Firpo and Pinto (2016) offer a solution to derive QTEs by reweighting the outcome distributions. This approach is succinctly outlined in Rios-Avila and Maroto (2024), and Rios-Avila (2020) provides tools to estimate QTEs.

In the following, we start by defining the unconditional QTE. We then describe the RQR model in more detail. Last, we showcase the RQR model's performance in data simulations and empirical applications on real data, comparing the RQR approach with other quantile regression approaches.

Unconditional QTEs

Ordinary least squares (OLS) and its estimation of average treatment effects (ATEs) is the main workhorse of quantitative empirical research, but scholars are increasingly turning to quantile regression models to estimate unconditional QTEs. The main attraction of unconditional QTEs is that it allows one to study treatment effect heterogeneity by individuals' overall propensity to have high or low outcomes.

Because of their close resemblance, let us briefly define ATE before turning to QTEs. In the potential outcomes framework (Morgan and Winship 2015), the causal effect of a treatment for a single unit is defined as

$$\delta_i = Y_{i,1} - Y_{i,0}, \quad (1)$$

where Y_{i1} is the value of Y for individual i when the treatment is set to 1, and Y_{i0} is the value of Y for the same individual when the treatment is set to 0. Observing the outcome in both treatment states is impossible in reality, and the unit-level causal effects are thus based on hypothetical, what-if states in a thought experiment. Using all observations within their two potential

states, we can calculate the commonly used ATE as the average difference between the potential outcomes:

$$\text{ATE} = E[Y_{i1}] - E[Y_{i0}]. \quad (2)$$

Likewise, if we know the whole distribution of the potential outcomes, $F_{Y_0}(y)$ and $F_{Y_1}(y)$, then we can define QTEs for the quantile τ in a similar manner:

$$\text{QTE}^\tau = Q_{Y_1}^\tau - Q_{Y_0}^\tau, \quad (3)$$

where $Q_{Y_1}^\tau$ and $Q_{Y_0}^\tau$ are the values of the quantile τ under the potential outcomes given the treatment and the nontreatment conditions (Frölich and Melly 2010). Consequently, QTE differs from ATE primarily by studying differences in quantile values rather than differences in means. Another difference is that QTEs potentially provide information about treatment effects for individuals at different parts of the outcome distribution. That is, assuming individuals maintain their ranks in the potential outcome distributions (i.e., rank preservation), QTE^τ can be interpreted as treatment effects for individuals at quantile τ of the potential outcome distribution in the absence of the treatment (Melly and Wüthrich 2017). To illustrate, rank preservation means that in a population of 100 people, the person with, for example, the 31st highest outcome under the treatment condition also has the 31st highest outcome under the nontreatment condition. The rank preservation assumption is strong and important; if it fails, differences between quantiles cannot be interpreted as individual-level treatment effects.

Whether unconditional QTEs differ from ATE depends on how the treatment influences the outcome (Hao and Naiman 2007). For example, suppose the treatment only shifts the outcome distribution's location left or right, as shown in Figure 1A. If a treatment only induces a location shift, it has a similar effect on all units across the potential outcome distribution. In that case, the difference between the 10th, 50th, 90th, or any other percentile under the treatment ($T = 1$) and no-treatment ($T = 0$) conditions is of equal size as the difference in means under the potential treatment states. In contrast, when the treatment induces both a location and a scale shift (Figure 1B), not all QTEs are identical to the difference in means. A scale shift represents heterogeneity in treatment effects across the ranks of the potential outcome distribution. Prior to using quantile regressions, one could check whether the treatment only affects the location of the outcome distribution, in which case a linear regression model is sufficient, or if it also affects the scale and skewness.

Estimating QTE in Cases with and without Randomization

Because counterfactual observations cannot be directly observed, estimation of unconditional QTEs must be based on the comparison of outcome distributions for the treated and untreated. Several assumptions must be met for this comparison of realized outcome distributions to provide an unbiased estimate of the unconditional QTEs, as we will detail when introducing the RQR model below. One of the key assumptions is the rank preservation assumption. Another is that the treated and untreated must differ only randomly concerning the composition of other factors that influence the outcome (i.e., the unconfoundedness assumption). If this assumption holds, then we can use the outcome distribution of the treatment groups as the what-if potential outcome distributions under different treatment conditions.

It follows from this that estimating unconditional QTEs is effortless in the rare case where treatment assignment is truly randomized. In that case, the treatment status is independent of other factors influencing the outcome, and therefore independent of the potential outcomes.

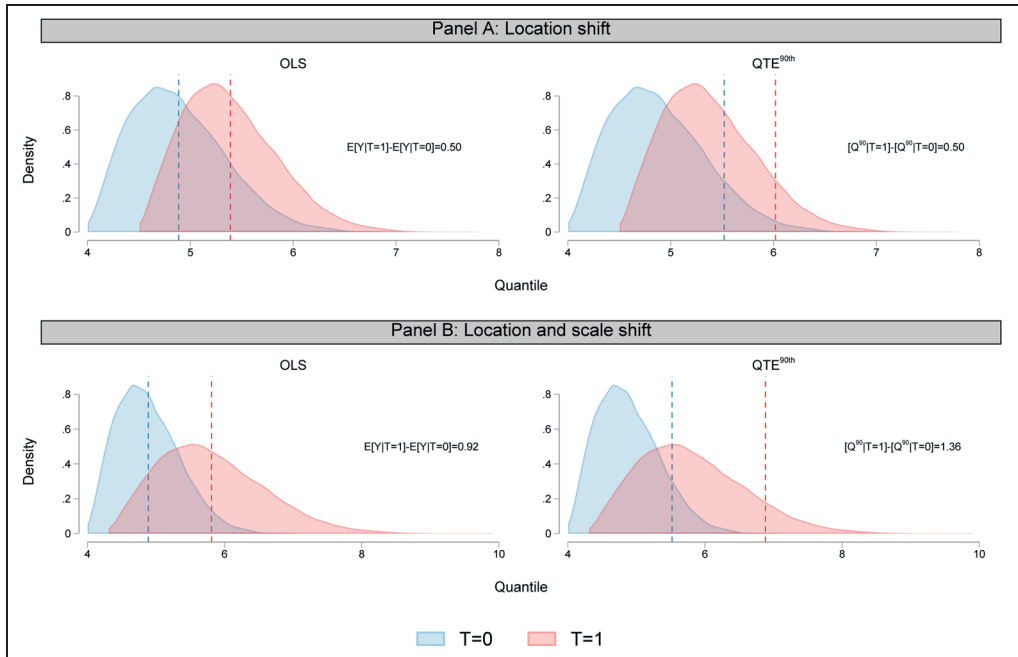


Figure 1. Illustrating the difference between OLS and QTE when the treatment variable induces only a location shift (A) and when the treatment induces both a location and a scale shift (B).

Note: The blue and red shaded areas represent kernel density estimates under the no-treatment ($T = 0$) and treatment ($T = 1$) conditions. The blue and red dotted lines show the mean (OLS) and 90th percentile ($\text{QTE}^{90\text{th}}$) in the potential outcome distributions. See part K in the online supplement for Stata code to replicate the graph. OLS = ordinary least squares; QTE = quantile treatment effect.

Thus, given a randomized treatment variable, we can ignore other influences and compare the distributions of the treated and untreated directly: one could calculate the quantile values among the treated and compare them with the corresponding quantile values among the untreated. The differences between these quantile values are the unconditional QTEs. Likewise, we can estimate these unconditional QTEs using a simple (conditional) quantile regression model without covariates (Koenker 2005).

However, treatment variables in the social sciences are typically not randomly assigned but involve selection processes. Therefore, in most cases, we want to add control variables or fixed effects to account for selection bias (Killewald and Bearak 2014). Even in randomized controlled trials, controls for baseline characteristics are often needed to ensure the estimates' internal validity. Unfortunately, estimating QTEs in the presence of control variables is more complicated than estimating treatment effects in the OLS model.

To illustrate the complexity of estimating unconditional QTEs with control variables, let us consider an example. This example also underscores the differences between linear regression and quantile regressions, as well as between quantile regression models estimating conditional and unconditional QTEs. We take the motherhood wage penalty as a starting point and, for pedagogical purposes, consider a simple model that assumes selection on observables, as is common in motherhood wage penalty studies (Cukrowska-Torzewska and Matysiak 2020). Consider the structural model where individual i 's wages y depend on motherhood status MS ,

years of education E , and other (unmeasured) causal factors e : $y_i = \beta_0 + \beta_1 MS_i + \beta_2 E_i + e_i$. A linear regression model can explore heterogeneity in the motherhood wage penalty by including interaction terms between motherhood status and other observed predictors, such as years of education (i.e., $\beta_3 [MS_i \times E_i]$).

In comparison, quantile regressions can be used to examine heterogeneity in terms of all factors influencing the outcome. Specifically, the CQR model estimates whether the effects of motherhood status differ according to unobserved factors that influence the outcome, that is, the rank of the residual e_i . This is sometimes called conditional QTEs (Frölich and Melly 2010), where conditional refers to the fact that the heterogeneity is estimated relative to unobserved factors that influence the outcome. Thus, the conditional QTE at, for instance, the 90th quantile is the motherhood penalty for women who have unobserved characteristics that put them at the 90th percentile of the residual e_i .

Studying conditional quantiles provides a unique way of estimating heterogeneity by unobserved factors influencing the outcome variable, making CQR a useful tool for examining heterogeneity by hard-to-observe characteristics. However, as highlighted in a number of studies (Borgen et al. 2023; Firpo 2007; Frölich and Melly 2010; Killewald and Bearak 2014; Wenz 2019), although comparing conditional quantile values efficiently accounts for selection on observables, it changes the interpretation of the coefficients. For example, instead of estimating treatment effects for units that, in the absence of the treatment, would have a high rank in the outcome distribution, the CQR model estimates effects on the basis of ranks conditional on the observed predictors. Women who have a high rank relative to others with the same education do not necessarily have an overall high rank. Furthermore, because the definition of high and low quantiles depends on the included independent variables, adding additional covariates will change the meaning of the estimated treatment effect.

Whereas CQR estimates effects on the basis of ranks conditional on the observed predictors, which means its coefficients have a conditional interpretation, scholars have developed approaches to translate CQR coefficients into their effect on overall marginal (or unconditional) quantiles. Notably, Machado and Santos Silva (2005) and Chernozhukov, Fernández-Val, and Melly (2013) provide methods that use CQR models to analyze counterfactual distributions. These methods, particularly Chernozhukov et al. (2013), aim to estimate and recover QTEs for settings involving binary treatments. This is achieved by linking CQR models to conditional distribution models, which are then integrated with covariate distributions to form counterfactual distributions.

Several approaches have also been developed to directly estimate unconditional QTEs in the presence of control variables (Firpo 2007; Frölich and Melly 2010; Powell 2020). Unlike the CQR model, QTE models attempt to estimate whether the effects of motherhood status differ according to all other factors that influence the outcome, both observed and unobserved. Let $e_i^* = \beta_2 E_i + e_i$ and rank e_i^* from the lowest to the highest value. The unconditional QTE at the 90th quantile, for instance, represents the motherhood penalty for women who, in the absence of motherhood, would be at the 90th percentile of the distribution of e_i^* . Thus, while linear regression and CQR allow the examination of heterogeneity based on *observed* (i.e., E_i) and *unobserved* (i.e., e_i) factors influencing the outcome, QTE models can be used to investigate heterogeneity by *overall* propensity to have high or low incomes (i.e., $\beta_2 E_i + e_i$), provided assumptions such as rank invariance and unconfoundedness are satisfied.

This article adds a new model to this class of QTE models. We claim this new RQR model has several practical advantages over existing ones, especially in terms of flexibility to add high-dimensional fixed effects and computational speed. In the following section, we present the framework, followed by a discussion of the advantages relative to existing models.

The RQR Model

We introduce a two-step approach to identify QTEs where the treatment variable is purged of confounding in the first step, followed by a QTE estimation in the second step using a quantile regression model. The core building block of this two-step approach is the decomposition of the treatment variable into a systematic piece explained by the observed control variables and a piece orthogonal to the controls. The next subsection elaborates on this decomposition before providing more details about the RQR approach.

Decomposition of the Treatment Variable

Assuming selection on observables, the variation in treatment variable values is a function of confounding variables and other random factors. Let C_i be such observed factors that influence both the treatment and the outcome (i.e., confounders) and E_i be all other factors influencing treatment status, including random noise:

$$T_i = f(C_i, E_i). \quad (4)$$

We want to decompose the treatment variable T_i into a piece explained by the observed confounding variables (C_i) and a residual piece (E_i). We can achieve this decomposition in several ways, with the appropriate approach depending on the specific selection processes. To illustrate our approach, we will begin with a correctly specified model in an undemanding confounding scenario where a linear regression model is sufficient, before turning to more challenging confounding structures and the consequences of misspecified models.

Consider the following data-generating process where T_i is the treatment variable, x_i is an observed confounder, and e_i captures all other causes of the treatment:

$$T_i = \delta_0 + \delta_1 x_i + e_i. \quad (5)$$

In this case, the confounder x_i induces only linear location shifts in the treatment T_i . Assuming this simple data-generating process, we can calculate the residuals \tilde{T}_i

$$\tilde{T}_i = T_i - \hat{T}_i = \delta_0 + \delta_1 x_i + e_i - \delta_0 - \delta_1 x_i = e_i \quad (6)$$

using a simple linear model.

For our purpose, the main takeaway point is the decomposition of the treatment variable into a piece explained by x_i ($E[T_i|x_i]$) and a residual piece orthogonal (in mean) to any function of x_i (\tilde{T}_i) (i.e., the conditional expectation function decomposition property; Cunningham 2020:56). This decomposition property is also applied in the classical Frisch-Waugh-Lovell theorem (Frisch and Waugh 1933; Lovell 1963, 2008), also known as the regression anatomy (Angrist and Pischke 2009:35–36; Filoso 2013) and double residual regression (Goldberger 1991:186). As proof that the residual piece is mean independent of x_i (assuming the model is correctly specified), consider the conditional expectation of the residuals \tilde{T}_i (from the regression of T_i on x_i) given the observed control variable x_i :

$$E[\tilde{T}_i|x_i] = E[T_i - E[T_i|x_i]|x_i] = E[T_i|x_i] - E[T_i|x_i] = 0. \quad (7)$$

Given a correctly specified model, the (treatment) residuals are mean independent of the confounder x_i and any function of x_i ($h[x_i]$): $E[h[x_i]\tilde{T}_i] = 0$. It follows that $\text{cov}(\tilde{T}_i, x_i) = 0$ as well. Furthermore, in this scenario, where the confounder merely shifts the mean of the

treatment, the residuals remain independent from the confounder, at the mean level and across all higher order moments. If confounders affect other moments of the treatment distribution (e.g., variance), the situation becomes more complex. Nonetheless, leveraging the decomposition and the treatment residual's orthogonality in mean to x_i ($E[\tilde{T}_i|x_i] = 0$), the next section presents the two-step estimator and its identification conditions.

A Two-Step Approach to Estimating QTEs

We propose to estimate QTEs using the following two-step approach:

Step 1: Decompose the treatment variable into a piece explained by the observed control variables and a residual piece.

Step 2: Regress the outcome variable on the residualized treatment variable using the CQR algorithm.

Several assumptions must be met for this two-step estimator to identify QTEs. The most important assumption is the unconfoundedness assumption, which loosely states that all variables affecting the treatment and the outcome (i.e., confounders) are observed and that the treatment selection is correctly specified. More specifically, we assume the potential outcomes are independent of the treatment T_i given the covariates (assumption 1). This assumption is similar to the conditional independence assumption of Powell (2020) and the unconfoundedness assumption of Firpo (2007).

Assumption 1 (unconfoundedness): $(Y_{i1}, Y_{i0}) \perp\!\!\!\perp T_i | x_i$.

This unconfoundedness assumption is fundamentally nonparametric, meaning it does not impose any specific functional form on the relationship between the potential outcomes, treatment, and observed control variables. Although nonparametric by nature, identification of treatment effects based on this assumption typically requires selection of a specific parametric model. Given that this parametric model adheres to the unconfoundedness assumption, a crucial point is that it implies the conditional independence of potential outcomes from the treatment residuals: $(Y_{i1}, Y_{i0}) \perp\!\!\!\perp \tilde{T}_i | x_i$. To understand this relationship, consider the decomposition of T_i into two components—a systematic part that can be predicted by the control variables $E[T_i|x_i]$, which is a deterministic function of x_i and a residual component \tilde{T}_i . This decomposition allows us to express the conditional independence assumption as $(Y_{i1}, Y_{i0}) \perp\!\!\!\perp (E[T_i|x_i] + \tilde{T}_i) | x_i$. Because $E[T_i|x_i]$ is a deterministic function of x_i , assumption 1 and the decomposition $T_i = E[T_i|x_i] + \tilde{T}_i$ implies $(Y_{i1}, Y_{i0}) \perp\!\!\!\perp \tilde{T}_i | x_i$. The stronger marginal statement $(Y_{i1}, Y_{i0}) \perp\!\!\!\perp \tilde{T}_i$ holds under additional structure (e.g., location-shift designs where $\tilde{T}_i \perp\!\!\!\perp x_i$).

In fact, the conditional independence of the residuals from the potential outcomes follows from the unconfoundedness assumption. That is, once the treatment is decomposed into $E[T_i|x_i] + \tilde{T}_i$, assuming $(Y_{i1}, Y_{i0}) \perp\!\!\!\perp (E[T_i|x_i] + \tilde{T}_i) | x_i$ implies that after adjusting for x_i , the systematic $E[T_i|x_i]$ and the residual part \tilde{T}_i are each conditionally individually independent of the potential outcomes. Thus, within the framework of the specified parametric model, the conditional independence $(Y_{i1}, Y_{i0}) \perp\!\!\!\perp \tilde{T}_i | x_i$ is not a separate condition but an implicit part of the unconfoundedness assumption. The marginal statement $(Y_{i1}, Y_{i0}) \perp\!\!\!\perp \tilde{T}_i$ applies in simple cases, such as location-shift designs where $\tilde{T}_i \perp\!\!\!\perp x_i$, but it generally fails for binary treatments.

However, it is crucial to note that assuming the potential outcomes are independent of the treatment given the covariates (i.e., assuming selection on observables) is a strong assumption, particularly within a parametric setting. Similar to Powell's (2020) GQR model and Firpo's (2007) propensity score approach, selection on unobservables would bias the estimates of QTEs.

The unconfoundedness assumption and the fact that the treatment is residualized before estimating the quantile regression model are the main departures from the CQR model. Thus, other conditions required in the CQR model, covered in detail by Koenker (2005), also apply to the RQR model, such as the assumption of no quantile crossing (He 1997) and a continuously distributed outcome variable (Machado and Santos Silva 2005). Let Y be a continuous random variable with the cumulative distribution function (CDF) of Y defined as $F_Y(y) = P(Y \leq y)$. If F_Y is strictly monotonically increasing, then the quantile function can be defined as $Q_Y(\tau) = F_Y^{-1}(\tau) = \inf\{y : F_Y(y) \geq \tau\}$.

Assumption 2 (CDF is strictly monotonically increasing): $Q_Y(p) = F_Y^{-1}(p)$.

Additionally, we invoke the rank invariance assumption for the treatment effects to be interpreted as individual-level QTEs. The rank invariance assumption is assumed in all quantile regression models that attempt to identify individual-level QTEs (Firpo 2007; Koenker 2005; Powell 2020), and prior work has developed approaches for testing rank invariance or rank similarity (see below) (Dong and Shen 2018; Frandsen and Lefgren 2018; Kim and Park 2022). Consider the binary treatment variable T_i and let Y_{i0} and Y_{i1} be the potential outcomes for $T_i = 0, 1$. Furthermore, let $r_i^0 \sim U[0, 1]$ and $r_i^1 \sim U[0, 1]$ be the potential ranks for $T_i = 0, 1$. We then assume the ranks in the potential outcome distributions are the same.

Assumption 3 (rank invariance assumption): $r_i^0 = r_i^1$.

The rank invariance assumption is strong, and a violation of this assumption prevents the estimates from being interpreted as individual-level QTEs. However, if the weaker rank similarity assumption holds, the RQR estimates may still have a meaningful interpretation. The rank similarity relaxes the rank invariance assumption by allowing for random deviations or "slippages" in ranks under the treatment and control conditions (Dong and Shen 2018). Specifically, rank invariance is the condition where an individual's potential rank with or without treatment remains the same, whereas rank similarity requires only that the potential ranks have the same conditional distribution when conditioned on specific observed and unobserved determinants of the common rank level. Formally, rank similarity is defined as the condition that $U_0|(X=x, V=v) \sim U_1|(X=x, V=v)$ for any (x, v) in its support, where U_0 and U_1 denote the potential rank in the outcome distributions under control and treatment, respectively, and X and V represent the observable and unobservable factors determining the common rank level (Dong and Shen 2018; Frandsen and Lefgren 2018). If this weaker rank similarity assumption holds, then the RQR estimates may be interpreted as differences in marginal outcome distributions.

Estimation of the RQR Model

The RQR model consists of a two-step approach, where the first step isolates as-if variation in the treatment (i.e., treatment residuals). Achieving this requires a correctly specified first-step model, including correctly specifying the control variables and choosing the appropriate link

function. Choosing the correct causal model, including what constitutes potential confounders, should be specified on the basis of know-how of the field (Morgan and Winship 2015; Pearl 2009). For example, not all covariates correlated with the treatment should be included, as highlighted by discussions on bias amplification of near-instrumental variables (Myers et al. 2011; Pearl 2011) and pretreatment collider variables (Elwert and Winship 2014). However, with a given set of covariates, and building on specification tests such as Ramsey's (1969) reset test and Pregibon's (1980) link test, we propose an F-test to jointly test whether there are unmodeled interactions or higher order terms among our covariates (see part C in the online supplement).

Concerning the choice of first-step estimator, the general framework developed here does not specify the approach used for decomposing the treatment variable in the first step. The first-step estimator should be chosen on the basis of whether it allows for satisfying the unconfoundedness assumption. For example, it is well known that with a binary treatment variable and common support problems, parametric models such as linear regression may perform poorly, and dropping observations out of support may be needed (Lechner and Strittmatter 2019). Thus, when the treatment is binary, there are benefits of using a logit or probit model to check common support and trim observations before performing the decomposition on the basis of the predicted probability of being treated from a logit model with common support. However, in most applications where we proceed under the assumption of selection on observables, a linear regression model offers a convenient first step to estimate $E[T_i|x_i]$ and construct treatment residuals. The following therefore considers a linear regression model as the first-step estimator.

As detailed above, under selection on observables and a correctly specified model, the first step purges the treatment variable of confounding, allowing us to use the residuals as an as-if randomized treatment variable in the second step. Thus, residualization of the treatment allows the exclusion of control variables in the second step under an additive conditional quantile condition, $Q_\tau(y_i|\tilde{T}_i, x_i) = q_0(\tau, x_i) + q_1(\tau)\tilde{T}_i$, which states that the quantile slope of the treatment residuals does not depend on x_i .

In the second step, linear programming methods are used to estimate coefficients.² The algorithm is only briefly described here; details can be found in Koenker (2005) and Hao and Naiman (2007). Known as the method of minimum absolute deviations, the QR estimator finds the coefficients that minimize the sum of weighted absolute residuals:

$$\sum_{i:y_i \geq x_i' \beta^{(\tau)}} \tau |y_i - x_i' \beta^{(\tau)}| + \sum_{i:y_i < x_i' \beta^{(\tau)}} (1 - \tau) |y_i - x_i' \beta^{(\tau)}|, \quad (8)$$

where $0 < \tau < 1$ and the superscript (τ) is included to clarify that the betas are allowed to differ by quantile. In the two-step RQR estimator, the coefficient matrix consists of only the residualized treatment variable and the constant. Thus, for the estimate of the p th quantile, we need to minimize the weighted absolute deviation of the coefficients of the residualized treatment variable $\beta_1^{(\tau)}$ and the constant term $\beta_0^{(\tau)}$:

$$\sum_{i:y_i \geq \beta_0^{(\tau)} + \beta_1^{(\tau)} \tilde{T}_i} \tau |y_i - \beta_0^{(\tau)} - \beta_1^{(\tau)} \tilde{T}_i| + \sum_{i:y_i < \beta_0^{(\tau)} + \beta_1^{(\tau)} \tilde{T}_i} (1 - \tau) |y_i - \beta_0^{(\tau)} - \beta_1^{(\tau)} \tilde{T}_i|. \quad (9)$$

Note that in this second-step regression, the constant term rarely has a sensible interpretation; it is the predicted quantile value of the outcome when the treatment residuals are zero.

Furthermore, by construction, there is a mechanical association between the residuals \tilde{T}_i and the (nonresidualized) treatment variable T_i , so regressing T_i on \tilde{T}_i trivially gives a coefficient of 1. Thus, when the residuals increase by one unit, the (nonresidualized) treatment, on average, increases by the same amount.³

To elucidate why $\beta_1^{(\tau)}$ from the second-step regression captures the unconditional QTE, one can view the QR model as a sorting algorithm. The second-step model estimates the coefficient of the treatment residuals by minimizing the quantile loss function:

$$L_\tau(\beta_0^{(\tau)}, \beta_1^{(\tau)}) = \sum_{i=1}^n [\tau \cdot u_i \cdot \mathbb{I}(u_i \geq 0) + (\tau - 1) \cdot u_i \cdot \mathbb{I}(u_i < 0)], \quad (10)$$

where quantile-specific predicted outcomes are defined as

$$Q_\tau(y_i) = \beta_0^{(\tau)} + \beta_1^{(\tau)} \tilde{T}_i \quad (11)$$

and the outcome residuals u_i as

$$u_i = y_i - Q_\tau(y_i). \quad (12)$$

Units are sorted based on whether their observed outcomes y_i fall above or below their quantile-specific predicted outcomes $Q_\tau(y_i)$, and the quantile regression optimization procedure penalizes these deviations according to their signs. By minimizing the loss function, the model optimizes the coefficients $\beta_0^{(\tau)}$ and $\beta_1^{(\tau)}$ to predict quantile values given the treatment variable.

Under assumption 1 and the additive conditional quantile condition, differences observed in the quantile-specific predicted outcome can be attributed to differences in treatment residuals:

$$\Delta Q_\tau(y_i) = \beta_1^{(\tau)} \Delta \tilde{T}_i. \quad (13)$$

That is, the mapping from $\Delta \tilde{T}$ to $\Delta Q_\tau(y_i)$ is invariant across x_i strata. Thus, the RQR model mirrors the core function of any QTE model: constructing outcome distributions that maintain consistent baseline characteristics but differ according to the treatment status of the units.

The adjustment-based two-step RQR approach provides asymptotically unbiased estimates of unconditional QTEs given that the unconfoundedness assumption holds. However, it is important to emphasize that the assumption of unconfoundedness is strong. Unobserved confounders or a misspecified model will bias RQR coefficients, similar to the case in other adjustment-based quantile regression models, including other QTE models, the UQR model, and the CQR model, as well as in linear regression models and propensity score matching approaches. An advantage of the RQR model is that we can include fixed effects to adjust for time-invariant unobserved confounders, in which case only failure to measure and correctly specify time-variant confounders is of concern (Allison 2009).

Inclusion of Fixed Effects

The inclusion of fixed effects constitutes a major challenge in nonlinear regression models, such as quantile regressions, because of the issue of the incidental parameter problem (Lancaster 2000). Several solutions to this problem have been developed within the CQR literature (Canay 2011; Koenker 2004; Machado and Santos Silva 2019; Powell 2022; Rios-Avila, Siles, and Canavire-Bacarreza 2024), including how to handle fixed effects without altering interpretation

of the estimates (Powell 2022), but there have been fewer developments concerning the inclusion of fixed effects within the class of estimators that identify unconditional QTEs. A notable exception is Rios-Avila (2020) and Rios-Avila and Maroto (2024), which suggest unconditional QTEs of binary treatment variables can be estimated by combining the UQR model with a reweighting strategy based on inverse probability weights. However, although fixed effects are easy to include in the standard UQR framework, the estimation of QTE via a reweighting of the UQR model with fixed effects remains an open question, as argued by Rios-Avila and Maroto (2024). We supplement this literature by suggesting an alternative approach to include fixed effects.

In the RQR model, fixed effects can be included in the first-step regression, along with the other control variables, to ensure the potential outcomes are independent of the treatment.⁴ Within the CQR literature, two-step estimators have been suggested to address the issue of the incidental parameter problem (e.g., Canay 2011; Machado and Santos Silva 2019; Rios-Avila et al. 2024). However, as these strategies aim to identify conditional QTEs, they differ from the approach suggested in the RQR model. For example, Canay (2011) assumes the fixed effects are location shifters and demean the outcome variable, before using CQR on the demeaned outcome variable. In contrast, the RQR model partials out the fixed effects from the treatment variable, and then uses a QR estimator with the demeaned treatment variable and the original (not demeaned) outcome variable. Two-step approaches, such as the one suggested in the RQR model, alleviate the incidental parameter problem but do not necessarily fully solve for it in all scenarios, particularly when the number of observations within each cluster is small and fixed (Canay 2011; Lancaster 2000; Machado and Santos Silva 2019; Rios-Avila et al. 2024). Note also that regarding the fixed effects in the RQR model, it is not the indicators themselves, but the unobserved confounders they represent, that contribute to satisfying the unconfoundedness assumption.

Inference

In various quantile regression models, estimation of standard errors using the bootstrap procedure (Hao and Naiman 2007; Mooney and Duval 1993) is preferred over the asymptotic procedure, including in the CQR model (Hao and Naiman 2007; Koenker and Hallock 2001), the propensity score QTE (PS-QTE) model (Firpo 2007), and the UQR model (Firpo et al. 2009). Estimating the asymptotic variance of quantile regression coefficients is notoriously challenging, as it depends on the unknown density of the dependent variable, and conclusions may be sensitive to the choice of kernel and bandwidths (Chernozhukov et al. 2013; Chernozhukov, Fernández-Val, and Melly 2022; Hagemann 2017; Koenker 2005; Powell 2020). Furthermore, simulation studies show that quantile regression inferences based on analytic standard errors perform worse than bootstrapping procedures (Chernozhukov et al. 2022; Hagemann 2017). Finally, providing analytic standard errors for testing differences of coefficients across quantiles is practically impossible (Hao and Naiman 2007), and bootstrapping or other simulation methods are therefore needed to test those types of hypotheses.

Bootstrapping within quantile regression models is computationally demanding (Fortin, Lemieux, and Firpo 2011), but this limitation can largely be circumvented by using the fast quantile regression algorithms provided by Chernozhukov, Fernández-Val, and Melly (2020). Thus, to provide valid inference, we propose to estimate standard errors in the RQR model using a bootstrap procedure, in which steps 1 and 2 described above are bootstrapped. Bootstrapping means randomly drawing M resamples of size N with replacements from the original data sample. In each resample, a decomposition of the treatment variable is followed

by an estimation of the quantile regression coefficients using the as-if randomized treatment variable. The standard deviations of the estimated coefficients are the standard errors of the RQR coefficients.

As noted above, the bootstrapping procedure also provides a solution to testing differences of coefficients across quantiles. With nonparametric quantile regression, we often want to know whether the effect of some independent variable differs across quantiles. Eyeballing overlapping confidence intervals (CIs) is often informally used for this purpose, but it provides a too-conservative test and increases the risk for type 2 error (Greenland et al. 2016). The bootstrapping procedure is a solution to this eyeballing fallacy in quantile regressions (Hao and Naiman 2007). Consider the case of comparing quantile regression coefficients at the 10th ($\beta^{[.10]}$) and 90th ($\beta^{[.90]}$) percentiles. In each bootstrap sample, the differences between the estimated quantile regression coefficients ($\beta^{[.90]} - \beta^{[.10]}$) is calculated, and the standard error of the difference is based on the standard deviation of these calculated differences. With multiple quantiles and comparisons, pairwise statistical significance across quantiles can, for example, be visualized with a heat plot of the p values (Brini, Borgen, and Borgen 2025). Note that pairwise tests can be misleading when comparing treatment effects at multiple quantiles, for example, with single significant differences reflecting chance findings. A Wald test can be used to test a joint null hypothesis that a specific set of quantile coefficients is equal (Hao and Naiman 2007).

Comparisons with Other Methods

The RQR model has several benefits compared with other quantile regression approaches. First, including high-dimensional fixed effects in the RQR model is straightforward. Like other control variables, fixed effects are included exclusively in the first step to account for confounding. After obtaining the treatment variable's residuals (first step), we can estimate the unconditional QTE using the quantile regression model with the residualized treatment and the outcome variable (second step). In contrast, including high-dimensional fixed effects in the propensity score framework derived in Firpo (2007) is not only computationally burdensome but also potentially problematic because of the logistic regression model in the first step. In Powell's (2020) GQR model, fixed effects can be included as information to construct the counterfactual distribution without treatment status (i.e., "proneess variables"); however, that computationally intensive solution is unfeasible with high-dimensional fixed effects.

Second, the RQR model is computationally efficient. The advent of big data and complex model specifications has made computational efficiency crucial in all regression models, but even more so in quantile regression models, where regressions are repeated at multiple quantiles and bootstrapping is needed for inference. The computational burden of quantile regressions may discourage researchers from using this approach (Chernozhukov et al. 2022; Fortin et al. 2011). When unconditional QTEs are estimated across multiple quantiles using the RQR model, one only needs to residualize the treatment variable once, making the estimation procedure more efficient and less computationally demanding. Furthermore, the computational time can be reduced considerably by applying newly developed quantile regression algorithms (Chernozhukov et al. 2020, 2022). As an illustration, consider the QTEs in Figure 3, discussed in detail below. Even with this undemanding model specification, the GQR model takes 20 times longer to run than the RQR model (for replication files, see part K in the online supplement).

Third, unlike the propensity score framework (Firpo 2007), the RQR approach extends to continuous treatment variables. Both binary and continuous treatment variables can be included in RQR without altering the coefficients' interpretation or changing the estimation procedure.

Although we can estimate QTEs of nonbinary treatment variables using the GQR model, that model cannot include fixed effects, as mentioned above.

Additionally, there are several minor benefits of the RQR approach. A fourth advantage is that the RQR framework can draw on the already extensive literature on CQR, including models to estimate parametric quantile regression (Frumento and Bottai 2016), methods to deal with count data (Machado and Santos Silva 2005), and fast algorithms (mentioned above). Finally, it is straightforward to implement the RQR estimator in all software that provides a package for CQR or linear programming, which makes it accessible to most researchers.⁵

However, like GQR, the RQR estimates lack meaningful interpretation when the rank similarity assumption is violated. Although a weaker condition than rank invariance, rank similarity is still a strong assumption, and the propensity score matching approach developed by Firpo (2007) has the advantage of being interpreted as differences between quantiles of the marginal distributions of potential outcomes, regardless of whether the rank invariance or rank similarity assumptions hold or not. Therefore, the propensity score matching approach is preferred over RQR (and GQR) if rank similarity is violated and researchers are interested in comparing marginal distributions rather than individual-level QTEs.

Data Simulations

Data-Generating Process

In this section we use Monte Carlo simulations to compare the RQR model's performance to other quantile regression approaches. We begin with fairly undemanding simulation scenarios before turning to more complex ones (results are presented Parts B–J in the online supplement). The data simulations in the main text consist of running 10,000 draws of $N = 2,000$ for two simulation scenarios. In both simulation scenarios, we have a continuous outcome variable y_i , a binary treatment variable t_i , and a binary control variable x_i . The two simulation scenarios differ concerning whether the control variable x_i affects only the outcome (scenario 1) or x_i affects both the outcome and the treatment variable (scenario 2). In simulation scenario 1, the treatment variable t_i is exogenous, and there is no need to include any control variables to identify the unconditional QTE. In contrast, controlling for the observed control variable x_i is necessary (and sufficient) to identify unconditional QTEs in simulation scenario 2.

More specifically, in scenario 1, we begin by randomly assigning 25 percent of the sample the value 1 on the binary control variable x_i and randomly assigning 10 percent of the sample the value 1 on the binary treatment variable t_i . Next, we define the potential outcome in the absence of the treatment y_i^0 as

$$y_i^0 = x_i \times 1 + \varepsilon_i, \text{ where } \varepsilon_i \sim N(0, 1). \quad (15)$$

Finally, we allow the strength of the treatment variable (t_i) to depend on individual i 's percentile rank r_i ($r_i \sim U[0, 1]$) in the distribution of potential outcomes in the absence of the treatment (y_i^0). The QTE monotonically increases across the outcome distribution.⁶

$$y_i = \left(\overbrace{r_i - 0.50}^{QTE} \right) \times t_i + y_i^0, r_i = \text{rank}(y_i^0). \quad (16)$$

The setup in scenario 2 is similar, except the conditional probability of being treated depends on x_i : $P(t_i = 1 | x_i = 0) = 0.067$ and $P(t_i = 1 | x_i = 1) = 0.20$. Thus, the likelihood of being

Table 1. Average Differences between Estimated Regression Coefficients and the True QTE ($\varphi^{[\tau]}$) and the Standard Deviation of the Differences ($\sigma_{\varphi}^{[\tau]}$) from Simulation Scenarios 1 and 2 for Selected quantiles (10,000 Draws of $N = 2,000$).

| | $Q^{.10}$ | | $Q^{.25}$ | | $Q^{.50}$ | | $Q^{.75}$ | | $Q^{.90}$ | |
|------------|-------------------|----------------------------|-------------------|----------------------------|-------------------|----------------------------|-------------------|----------------------------|-------------------|----------------------------|
| | $\varphi^{(.10)}$ | $\sigma_{\varphi}^{(.10)}$ | $\varphi^{(.25)}$ | $\sigma_{\varphi}^{(.25)}$ | $\varphi^{(.50)}$ | $\sigma_{\varphi}^{(.50)}$ | $\varphi^{(.75)}$ | $\sigma_{\varphi}^{(.75)}$ | $\varphi^{(.90)}$ | $\sigma_{\varphi}^{(.90)}$ |
| Scenario 1 | | | | | | | | | | |
| RQR | .003 | (.151) | .001 | (.133) | -.002 | (.128) | -.003 | (.134) | -.004 | (.154) |
| PS-QTE | .003 | (.151) | .002 | (.133) | -.002 | (.128) | -.003 | (.134) | -.005 | (.155) |
| GQR | .004 | (.151) | .003 | (.133) | -.001 | (.127) | -.003 | (.134) | -.003 | (.154) |
| UQR-QTE | .003 | (.152) | .001 | (.134) | -.002 | (.128) | -.002 | (.134) | -.005 | (.155) |
| CQR | .031 | (.149) | .027 | (.128) | .001 | (.121) | -.028 | (.129) | -.038 | (.148) |
| UQR | -.003 | (.164) | .037 | (.117) | -.002 | (.101) | -.038 | (.117) | .002 | (.165) |
| Scenario 2 | | | | | | | | | | |
| RQR | .007 | (.174) | .005 | (.147) | .004 | (.132) | .001 | (.129) | -.001 | (.142) |
| PS-QTE | .006 | (.183) | .003 | (.159) | .003 | (.145) | .000 | (.140) | -.002 | (.150) |
| GQR | .007 | (.176) | .007 | (.149) | .007 | (.132) | .002 | (.128) | -.001 | (.139) |
| UQR-QTE | .007 | (.184) | .003 | (.160) | .003 | (.146) | -.001 | (.141) | -.003 | (.151) |
| CQR | .084 | (.156) | .100 | (.134) | .081 | (.122) | .032 | (.126) | -.004 | (.145) |
| UQR | .080 | (.149) | .072 | (.108) | .003 | (.100) | -.011 | (.125) | .121 | (.195) |

Note: Data simulation is performed in Stata 16.0; files to replicate the results are available in part K in the online supplement. CQR is the conditional quantile regression model (Koenker 2005) estimated using the `qreg` command; RQR is the residualized quantile regression model introduced in this article; PS-QTE is the propensity score framework of Firpo (2007) estimated using the `ivqte` command (Frölich and Melly 2010); GQR is the generalized quantile regression (Powell 2020) estimated using the `genqreg` command; UQR-QTE is the reweighting approach of the UQR model suggested by Rios-Avila (2020); and UQR is the unconditional quantile regression model (Firpo et al. 2009) estimated using the `rifreg` command. QTE = quantile treatment effect.

treated is 13.3 percentage points higher for individuals with 1 on the observed control variable than for those with the value 0.

We estimate quantile regression coefficients using six different quantile regression models for each of the 10,000 draws and in each of the two simulation scenarios. Four of these estimation strategies identify unconditional QTEs, and their coefficients should, therefore, be nearly identical: the RQR method introduced in this article, Firpo's (2007) PS-QTE approach, Powell's (2020) GQR approach, and Rios-Avila's (2020) reweighted UQR model, which we call UQR-QTE. The last two strategies are the CQR model (Koenker 2005) and the UQR model (Firpo et al. 2009), which we include because they are popular quantile regression methods. The RQR coefficients may differ from the CQR and UQR coefficients, which identify conditional QTEs and unconditional partial effects. All models are run with the treatment variable t_i and the control variable x_i . The Monte Carlo Error is small and hovers around 0.001 to 0.002 for all models and quantiles, as shown in Figure A1 in the online supplement.

Main Simulation Results

The main simulation results are shown in Table 1 and Figure 2. The reported $\varphi^{(\tau)}$ quantities show the average difference between the estimated regression coefficient ($\hat{\beta}_j^{[\tau]}$) and the true QTE ($\beta^{[\tau]}$) at the quantile τ across the 10,000 independent draws j : $\varphi^{(\tau)} = E[\hat{\beta}_j^{[\tau]} - \beta^{[\tau]}]$. The $\varphi^{(\tau)}$ quantity should be interpreted as the estimated bias for the QTE models, where larger

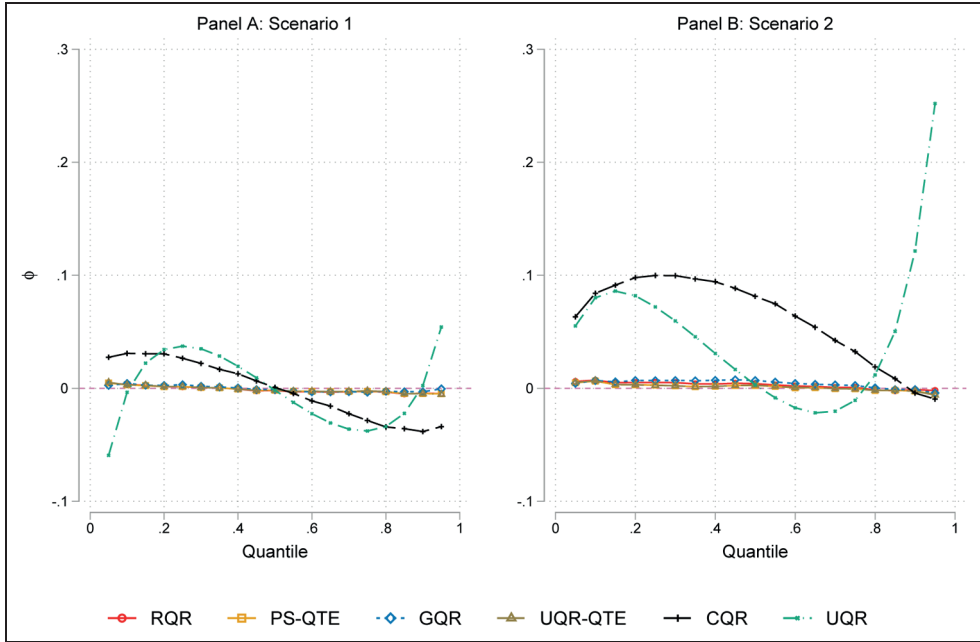


Figure 2. Average differences between estimated regression coefficients and the true QTE ($\phi^{[\tau]}$) from simulation scenarios 1 (A) and 2 (B).
 Note: See note in Table 1 for details. CQR = conditional quantile regression; GQR = generalized quantile regression; PS = propensity score; RQR = residualized quantile regression; QTE = quantile treatment effect; UQR = unconditional quantile regression.

values mean more bias. In contrast, because CQR and UQR models do not estimate unconditional QTEs, any differences between their coefficients and the true coefficient do not reflect a bias and should not be interpreted as such. Table 1 also reports the standard deviation across draws of the estimation errors, $\sigma_{\phi}^{(\tau)} = SD[\hat{\beta}_j^{(\tau)} - \beta^{(\tau)}]$. Table A1 in the online supplement presents simulation results using estimated regression coefficients, rather than differences between coefficients and true QTE.

As expected, all four QTE models yield coefficients that approximate the true QTE across the outcome distribution, with the bias ranging from 0.001 to 0.005 in most cases. The differences between simulation scenario 1 (where there is no confounding) and simulation scenario 2 (where conditioning on x_i is sufficient to account for confounding) is negligible. There is a slight tendency for the QTE models to overestimate the QTE at the bottom of the outcome distribution and underestimate it at the top. However, this bias amounts to only 1 percent to 3 percent of its standard deviation on the basis of the different draws, which is about 0.14 in the first scenario and slightly higher in the second scenario. Moreover, the bias is small relative to the size of the QTEs. For example, the estimated QTEs at the 10th percentile in scenario 1 are 0.75 percent off in the RQR, PS-QTE, and UQR-QTE models and 1.25 percent off in the GQR model (see Table A1). Overall, the similarity of the estimated effects across the four frameworks provides evidence that running a quantile regression model on a residualized treatment variable is a valuable approach to identifying unconditional QTEs.

Table 2. Average of the Absolute Values of the $\sigma_{\phi}^{(\tau)}$ and $\sigma_{\phi}^{(\tau)}$ Quantities across the Estimated Quantiles.

| | $ \phi^{(\tau)} $ | $ \sigma_{\phi}^{(\tau)} $ |
|------------|-------------------|----------------------------|
| Scenario 1 | | |
| RQR | .003 | .140 |
| PS-QTE | .003 | .140 |
| GQR | .002 | .140 |
| UQR-QTE | .003 | .141 |
| Scenario 2 | | |
| RQR | .003 | .144 |
| PS-QTE | .002 | .156 |
| GQR | .005 | .145 |
| UQR-QTE | .003 | .156 |

Note: In each of the 10,000 draws, the average coefficient ($\phi^{(\tau)}$) and its standard deviation ($\sigma_{\phi}^{(\tau)}$) are calculated (see Table 1 and Figure 2). This table shows the average of these quantities' absolute values at quantiles 5 to 95 at steps of 5 (19 observations for each scenario and method combination). GQR = generalized quantile regression; PS = propensity score; QTE = quantile treatment effect; RQR = residualized quantile regression; UQR = unconditional quantile regression.

This article's data simulation is not a comprehensive comparison of different QTE models. However, we note a few differences between the three QTE models in this setup. First, the bias is marginally lowest in the PS-QTE approach, followed by the RQR model, the UQR-QTE model, and finally, the GQR model (see Table 2). The average of the absolute value of the bias across the 19 quantiles is nearly identical at the third decimal in scenario 1. In scenario 2, the PS-QTE, RQR, and GQR models' biases are 0.002, 0.003, and 0.005. Second, the standard deviations across the draws are 7 to 8 percent lower in the RQR and GQR models than in the PS-QTE and UQR-QTE models in scenario 2 (and indistinguishable in scenario 1). All in all, the differences between the QTE models are trivial in this simulation setup.

The QTE models' coefficients differ from the CQR and UQR coefficients in both simulation scenarios. Let us start with the CQR results. As noted above, the difference between the CQR coefficients and the true QTE should not be interpreted as a bias, as the CQR model estimates conditional quantile value differences. As Figure 2 shows, the CQR coefficients differ in scenarios 1 and 2, despite the underlying unconditional QTE being the same in both scenarios. In fact, including the observed control variable x_i changes the interpretation of the conditional QTEs even in scenario 1, where the control variable x_i does not influence the treatment variable t_i (no-confounding scenario) (Wenz 2019). Contrary to this, excluding the control variable x_i does not change the estimated unconditional QTEs using RQR, PS-QTE, or GQR (not shown).

The unconditional partial effects, estimated using Firpo et al.'s (2009) UQR model, also differ considerably from the unconditional QTEs. Moreover, the UQR coefficients differ in scenarios 1 and 2, even though the underlying QTEs are the same. The differences between the QTE models (RQR, GQR, PS-QTE, and UQR-QTE) and the UQR model illustrate that the standard UQR model does not identify QTEs, although it is widely used in the literature for that purpose (Borgen et al. 2023). Thus, if the UQR model is used to examine unconditional QTEs of binary variables, one should use the reweighting approach (i.e., UQR-QTE).

Supplementary Simulation Scenarios

We supplement the simulation scenarios in the main text with a broader set of simulations in Parts B to J in the online supplement, including three simulation scenarios where the RQR model fails to identify the target estimand. We discuss these results briefly here, with more information included in the online supplement. First, part B in the supplement varies the distribution of the treatment, the structure of the QTE, and the skewness of the outcome distribution. Specifically, we simulate data with a binary treatment variable (as in the main text), a treatment variable with a uniformly distributed random component, and a treatment variable with a normally distributed random component. We also distinguish between three structures on the QTEs: constant, quadratic, and cubic. Finally, we allow the outcome's residual distribution to be either normal or right skewed. The results in all 18 simulation scenarios align with the results discussed above: RQR provides approximately unbiased estimates of the unconditional QTEs given a correctly specified first step.

Second, part C in the online supplement simulates data where confounders interact and have a quadratic effect on the treatment. We show that incorrectly specifying the treatment selection results in biased estimates of the QTE. However, such misspecification can, in some cases, be detected using a specification test, as illustrated by the example in the supplement. Moreover, when interactions and quadratic effects are correctly specified, the RQR model provides approximately unbiased estimates of the QTEs.

Third, part D in the online supplement provides an example where the confounder's effect on the treatment is allowed to vary randomly. For some individuals, the confounder strongly influences the treatment values, whereas it has no effect on others. (As a result, the variance of the treatment residuals increases as a function of the confounder.) We find that the RQR model provides approximately unbiased QTE estimates in this scenario.

Fourth, part E in the online supplement simulates data where the variance of the treatment is a function of the confounder. We find that the RQR model provides approximately unbiased estimates in a scenario where the variance of the treatment residuals is a function of the confounder.

Fifth, part F in the online supplement simulates data where the confounder has a curvilinear effect on the outcome and a scenario where the confounder has a QTE. Again, the RQR model provides approximately unbiased estimates in these simulation scenarios.

Sixth, part G in the supplement simulates panel data where there is a time-invariant unobserved confounder. Whereas other QTE approaches fail in such a scenario, we show that we can use the RQR model to identify approximately unbiased QTEs using a fixed effects estimator.

Finally, the last three supplements show simulations where the RQR model (and other QTE models) fails to identify QTEs. Part H simulates data where the rank invariance assumption is not met. Rank invariance means that individuals' ranks in the potential outcome distributions remain the same irrespective of their treatment. Without invoking this assumption, QTEs cannot be interpreted as treatment effects for individuals located in different parts of the distribution. The supplement shows that RQR provides biased individual-level QTE estimates when rank invariance is not met, as do the GQR, PS-QTE, and UQR-QTE models. Next, part I illustrates how the RQR model fails to identify QTEs in the presence of quantile crossing, along with the GQR model. Finally, the CQR algorithm requires the uniqueness of quantiles (Firpo 2007), which is not satisfied with a discrete outcome variable (Machado and Santos Silva 2005). Part J shows that the RQR model struggles with discrete outcome variables, and it struggles more than

the GQR model in the specific example included. However, the supplement also shows that the jittering approach of Machado and Santos Silva (2005) can be used to solve the problem.

Empirical Applications

After demonstrating the approximate unbiasedness of the RQR in various simulation scenarios where the identifying assumptions are met, we turn to real data applications. Our goal here is to benchmark RQR against established QTE estimators under comparable specifications, not to claim causal identification. In these applications, the true causal effects are unknown and identification is not guaranteed, because assumptions such as unconfoundedness and rank invariance may not hold and because of issues related to overcontrol. We compare estimated coefficients across methods to assess whether RQR recovers the same unconditional quantile contrasts in practice as established QTE estimators. Generally, similar coefficients support RQR's performance relative to established estimators, although coefficients may differ somewhat because of sampling variability and differences in specifications or assumptions.

Current Population Survey

This section illustrates the RQR model using the Outgoing Rotation Group supplement of the Current Population Survey to study the effects of union status on men's wages. We again highlight the quantile regression models' differences. Union wage effects served as a motivating example in Firpo et al.'s (2009) UQR article, which has become popular in studies of QTEs (Borgen et al. 2023). Firpo et al. (2009) demonstrate that UQR coefficients differ from CQR coefficients. Here, we use the union wage effects example to showcase that RQR coefficients differ from UQR (and CQR) coefficients but not GQR, PS-QTE, or UQR-QTE coefficients.⁷

Convincingly, differences between estimated union wage effects using QTE models are negligible; RQR, GQR, PS-QTE, and UQR-QTE coefficients decrease similarly, from about 0.40 at the 5th quantile to -0.10 at the 95th quantile (Figure 3). Concerning computational speed, less than a minute is needed to estimate the RQR coefficients on a standard laptop when using the fast quantile regression algorithms provided by Chernozhukov et al. (2020). The PS-QTE model takes more than five times longer to run, and the GQR model 20 times longer. The union wage effects are estimated with controls for age, education, marital status, and race. These control variables may not be sufficient to account for confounding; however, as potential bias is similar across models, the model specification allows us to compare RQR with the other quantile regression approaches.

Figure 4 compares RQR coefficients with UQR and CQR coefficients. For completeness, we also include the UQR-QTE coefficients. Two main results stand out in the cross-sectional analysis in Figure 4A. First, although the RQR and CQR models' estimated union wage effects are monotonically declining across the outcome distribution, the gradient is steeper in the RQR model. The RQR coefficients are larger at the bottom of the wage distribution, before falling and eventually reaching a substantial adverse effect at the top. This finding highlights the difference between conditional and unconditional QTEs, as others have previously discussed (see Firpo 2007; Killewald and Bearak 2014; Porter 2015; Wenz 2019). Second, the UQR coefficients differ from the RQR coefficients, especially in the bottom fifth of the wage distribution. The reason is that Firpo et al.'s (2009) UQR model identifies UQPE rather than unconditional QTE.

In Figure 4B, we estimate union wage effects with household fixed effects. Including high-dimensional fixed effects in the RQR model is straightforward; fixed effects are treated as any

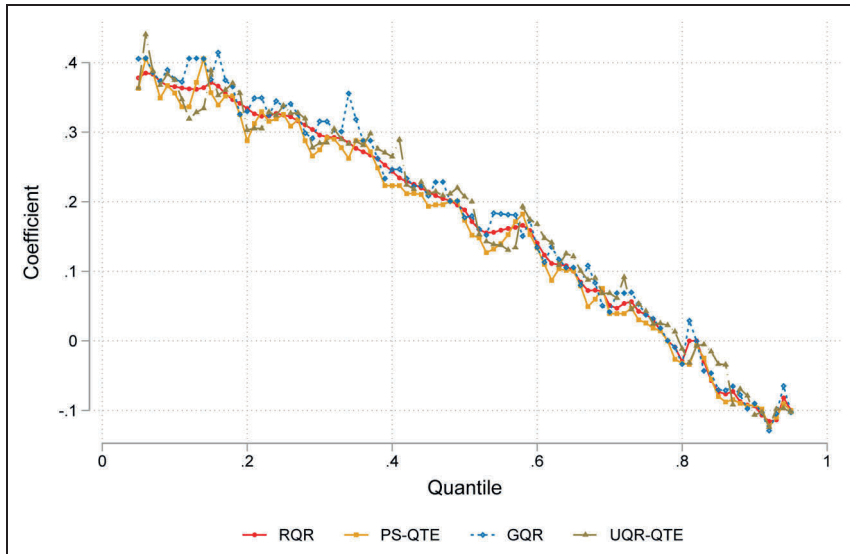


Figure 3. Cross-sectional effects of union status on log wages for full-time working men in the 1983 to 1986 Outgoing Rotation group supplement of the Current Population Survey ($N = 251,153$).

Note: The sample includes male household heads 16 to 64 years of age. The included control variables are respondents' age, five dummies for education level, a dummy variable for completed education, a dummy variable for married, and a dummy variable for nonwhite. PS-QTE is the weighting procedure of Firpo (2007) as implemented by the *ivqte* command in Stata (Frölich and Melly 2010). The GQR model is the generalized quantile regression model of Powell (2020). The UQR-QTE is the reweighting approach of unconditional quantile regression (UQR) suggested by Rios-Avila (2020). Data come from the National Bureau of Economic Research (<https://data.nber.org/cps/>). PS = propensity score; RQR = residualized quantile regression; QTE = quantile treatment effect.

other control variable and included only in the first step. After obtaining the treatment variable's residuals, unconditional QTEs can be estimated by regressing the outcome on the residualized treatment in a quantile regression model. Fixed effects can also, with ease, be included in the UQR model (Borgen 2016; Rios-Avila 2020). However, as it is challenging in the CQR model (for one approach, see Machado and Santos Silva 2019), and estimation of QTE via UQR with fixed effects remains an open question (Rios-Avila and Maroto 2024), Figure 4B compares the RQR and UQR-QTE coefficients solely with those of UQR. In the union example, the estimated within-household treatment effects are generally lower than the cross-sectional estimates, especially at the bottom of the wage distribution. However, the overall trend is similar in the cross-sectional estimate (Figure 4A) and the fixed effects estimate (Figure 4B).

Norwegian Register Data

We use Norwegian register data in the following example, which largely alleviates issues caused by imprecise coefficients in small samples. Figure 5 looks at associations between 8th-grade standardized test scores and birth order, birth month (i.e., relative school-starting age), immigrant background, gender, school socioeconomic status, and parental earnings, where all independent variables are estimated controlling for the other independent variables in the graph. We show coefficients of all variables to illustrate the method across variables with different types of distributions. The four different QTE models provide nearly identical estimates

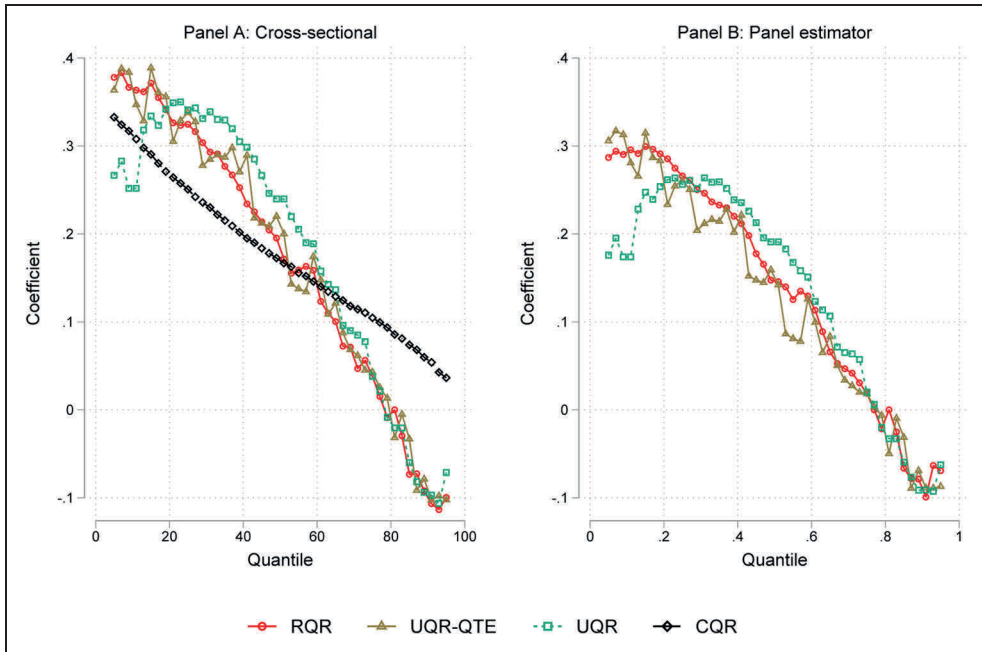


Figure 4. Effects of union status on log wages for full-time working men in the 1983 to 1986 Outgoing Rotation group supplement of the Current Population Survey ($N = 251,153$).

Note: Panel A shows cross-sectional estimates; panel B includes household fixed effects. See Figure 3 for details about data and variables. The UQR model is estimated using the `rifreg` (A) and `xtrifreg` (B) commands (Borgen 2016). CQR = conditional quantile regression; RQR = residualized quantile regression; QTE = quantile treatment effect; UQR = unconditional quantile regression.

for five of six variables. Importantly, this is the case despite the treatment effects having completely different effects across the outcome distribution. This result mirrors the findings in part B in the online supplement, where we use data simulations to show the RQR model can handle different treatment structures.

The exception is the variable *children of immigrants*, where the PS-QTE approach differs somewhat from the RQR, GQR, and UQR-QTE models. However, this difference does not warrant concern; rather, it showcases the well-known fact that regression and matching approaches may differ in some scenarios, despite identifying the same estimand. In this case, the inverse probability weighting matching approach also provides a different estimate of the ATE than does the classical OLS model (not shown).

National Longitudinal Survey of Young Working Women

As a final example, we use a subsample of the National Longitudinal Survey, containing panel data on 4,711 working women 14 to 26 years of age in 1968 from 1968 to 1988. The outcome is log wages, and as in Figure 5, we show coefficients for all independent variables. Although the smaller sample size results in noisier coefficients, we clearly see that the GQR, RQR, PS-QTE, and UQR-QTE coefficients resemble each other, irrespective of the type of independent

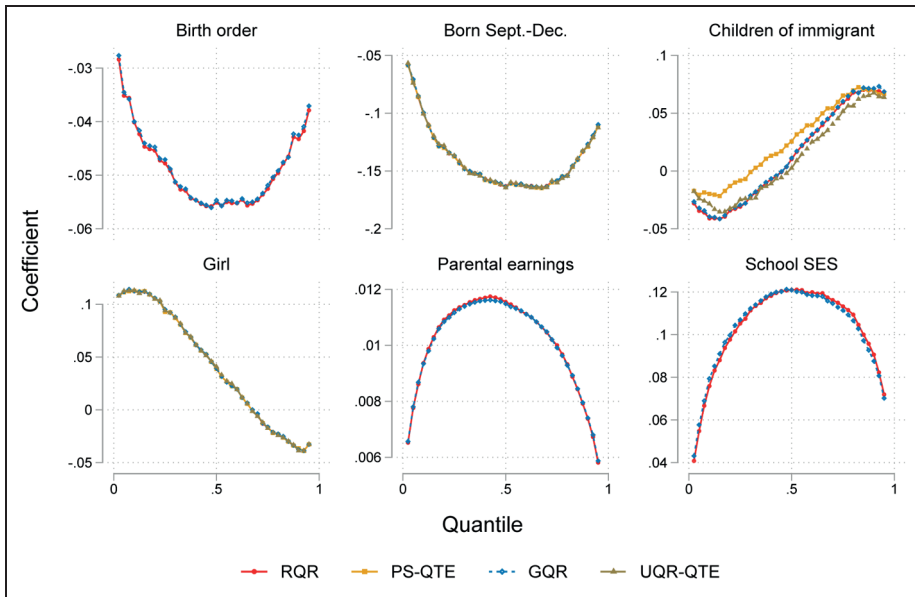


Figure 5. Comparing RQR, GQR, PS-QTE, and UQR-QTE coefficients on eighth grade standardized test scores in Norwegian register data ($N = 480,264$).

Note: The outcome variable is the average of eighth grade entry-test scores in reading and math for birth cohorts born in Norway between 1997 and 2005. Birth order is defined on the basis of both mother and father. Born Sept.–Dec. is a dummy variable for those born in the last quarter of the year; the school-starting cutoff is January 1. Children of immigrants are children born to at least one parent born outside Norway. Girl is a dummy for gender. Parental earnings measure parents' percentile earnings rank. School SES (socioeconomic status) is the average parental earnings rank among fifth graders at the same school. The residualized quantile regression (RQR) and generalized quantile regression (GQR) models are used for all variables; the PS-QTE and UQR-QTE models are only used for binary predictors. All coefficients are estimated controlling for the other independent variables in the graph. PS = propensity score; QTE = quantile treatment effect; UQR = unconditional quantile regression.

variable (binary or metric) and the structure of treatment effects across the distribution (see Figure 6).⁸

Finally, Figure 7 illustrates the bootstrapping procedure to test differences of coefficients across quantiles. The figure uses a heat plot to show pairwise statistical significance across quantiles for the age variable.

Discussion and Conclusions

Identification of unconditional QTEs has become increasingly popular within social sciences; however, contrary to popular belief, current methods to identify unconditional QTEs are incomplete. Firpo (2007) proposed a solution to estimate unconditional QTEs with a single binary treatment variable using a propensity-matching framework. Powell (2020) developed the GQR model that extends to nonbinary treatment variables. However, none of these approaches easily allows for high-dimensional (additive) fixed effects. Although a solution to estimate QTEs via UQR has been developed (Firpo and Pinto 2016; Rios-Avila 2020) and fixed effects are easy to include in the standard UQR framework (Borgen 2016; Rios-Avila 2020), estimation of QTE via UQR with fixed effects remains an open question (Rios-Avila and Maroto 2024).

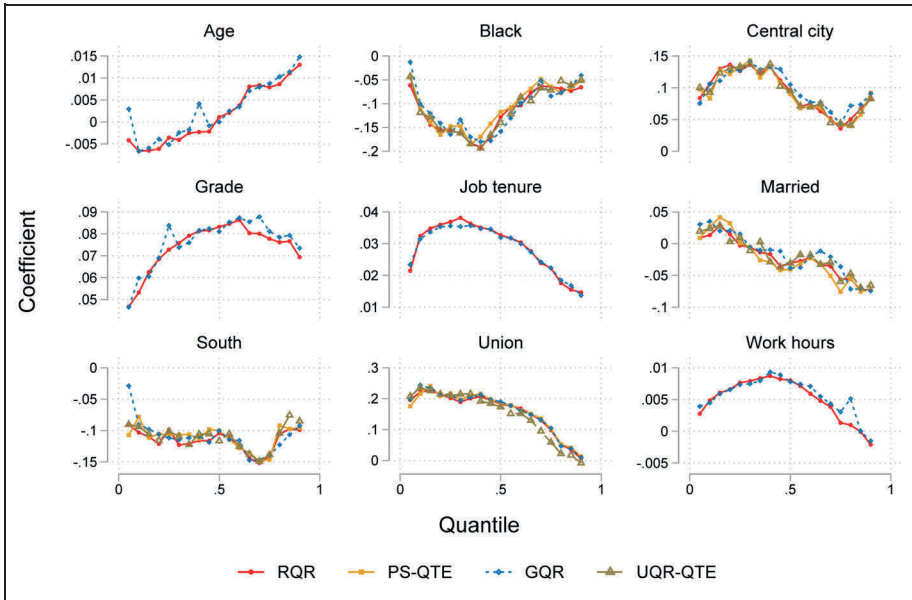


Figure 6. Comparing RQR, GQR, PS-QTE, and UQR-QTE coefficients on log wages in a subsample of the National Longitudinal Survey ($N = 3,956$).

Note: We keep observations born between 1943 and 1953 and that are between ages 25 and 45 during the interview years. Then we keep one observation per woman (the latest recorded one). The residualized quantile regression (RQR) and generalized quantile regression (GQR) models are used for all variables; the PS-QTE and UQR-QTE models are only used for binary predictors. All coefficients are estimated controlling for the other independent variables in the graph. PS = propensity score; QTE = quantile treatment effect; UQR = unconditional quantile regression.

This article fills a gap in the literature by introducing a straightforward framework to estimate unconditional QTEs of both continuous and binary treatment variables in the presence of covariates and high-dimensional fixed effects, called RQR. The main advantage of the RQR model compared with other QTE models is that fixed effects can be included, which is a common workhorse in causal modeling, as exemplified in the literature on motherhood and fatherhood wage penalties (Cooke 2014; England et al. 2016; Killewald and Bearak 2014). Although studies have highlighted that the fixed effects modeling strategy is not a panacea for causal inference (Imai and Kim 2021)—because of issues such as time-varying effects of time-invariant confounders (Ren and Allison 2025), potentially limited external validity (Hill et al. 2020; Lancaster 2000), and challenges related to variation in treatment timing in difference-in-difference models (Goodman-Bacon 2021)—it remains a powerful tool for identifying causal effects.

Furthermore, the RQR model can draw on the already extensive literature on CQR, is easy to implement in all software that provides packages for CQR or linear programming, and is computationally efficient compared with other quantile regression models. With better access to large-scale data and more complex model specifications, the computational burden may discourage researchers from using quantile regressions. The RQR model's computational time can be reduced considerably by applying newly developed quantile regression algorithms

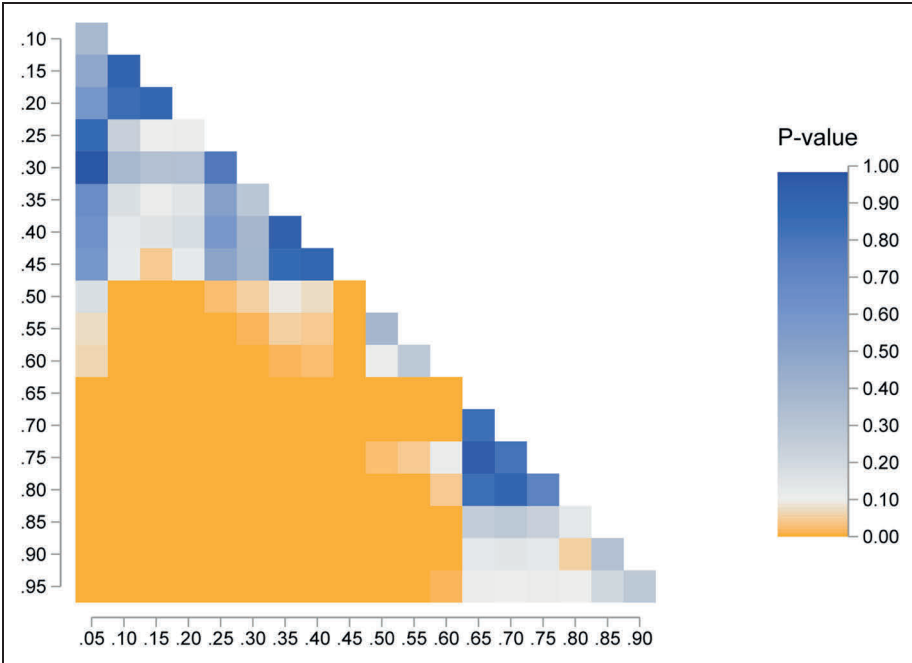


Figure 7. Heat plot of p values to compare differences of age coefficients across quantiles on the basis of residualized quantile regression coefficients from Figure 6.
Note: Differences are based on bootstrapping the difference in coefficients between quantiles (200 replications).

(Chernozhukov et al. 2020, 2022), making it substantially faster than current QTE approaches. Borgen et al. (2021) provide a Stata command that exploits these fast algorithms.

This article introduced a new framework for estimating unconditional QTEs and provided strong evidence that it produces coefficients similar to other QTE models. We suggested using OLS in the first-step regression to deconfound the treatment variable; however, future research should explore other first-step regressions, especially with binary treatment variables. Furthermore, statistical inference and estimation of RQR standard errors need more work. Bootstrapping the two-step procedure seemingly provides sound standard errors, as the 95 percent bootstrapped CIs in our data simulations have coverage rates of 94.34 percent (normal-approximation bootstrap CIs), 95.26 percent (percentile bootstrap CIs), and 94.52 percent (bias-corrected bootstrap CIs) (see Figure A4 in the online supplement). Future work should systematically test the performance of various bootstrap procedures in different simulation scenarios. Additionally, the interpretation of RQR coefficients as individual-level QTEs depends on the rank similarity assumption. Approaches for testing rank similarity have been developed (Dong and Shen 2018; Frandsen and Lefgren 2018; Kim and Park 2022), but more research is needed to test rank similarity within the RQR framework. With binary treatment variables, the propensity score matching approach developed by Firpo (2007) may be preferable, as the estimated treatment effects have a meaningful interpretation even when the rank similarity assumption does not hold. Finally, the current article only dealt with the case of selection on observables. The RQR framework should be developed to allow the identification of an endogenous treatment variable’s QTEs via instrumental variable estimation.

Authors' Note

The authors used Refine and GPT UiO (powered by OpenAI's GPT models) to assist with presubmission consistency checks. Grammarly and GPT UiO were used for language editing.

Acknowledgments

Earlier versions of this article were presented at the 2021 International Sociological Association RC28 Spring Meeting in Turku, Finland, the 2022 DAGStat Conference in Hamburg, Germany, the Social Inequalities and Population Dynamics group's seminar at the University of Oslo in 2021, and the EQOP seminar at the University of Oslo in 2021. We thank participants for their comments and suggestions.

ORCID iDs

Nicolai Borgen  <https://orcid.org/0000-0002-7638-3293>

Andreas Haupt  <https://orcid.org/0000-0002-6519-362X>

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The preparation of this article was supported by funding from the European Research Council (grants 818425 and 101115949) and was partially supported by the Research Council of Norway through its Centres of Excellence scheme (grant 331640). The register data were made available by Statistics Norway. Views and opinions expressed are those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency (ERCEA). Neither the European Union nor the granting authority can be held responsible for them.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Data Note

Data sets and Stata do-files to replicate the results are available at the Open Science Framework repository: https://osf.io/28ysu/overview?view_only=9b90041eccb7432cbfe38793cda58e5d. The replication folder includes code and data to reproduce all results, except Figure 5. Figure 5 used Norwegian administrative data provided by Statistics Norway. For confidentiality reasons, these data are not publicly available but can be accessed upon receiving relevant approvals.

Supplemental Material

Supplemental material for this article is available online.

Notes

1. We do not discuss Powell's (2022) nonadditive fixed effects panel estimator (quantile regression for panel data) in depth in this article. The motivation behind the RQR model and Powell's quantile regression for panel data model is similar—to identify unconditional QTEs in the presence of covariates—but the problem they address differs; Powell's model corrects for biases caused by unit specific trajectories.
2. This approach is similar in spirit to the PS-QTE developed by Firpo (2007), which estimates propensity scores, constructs inverse probability weights, and then minimizes a weighted quantile check loss with a binary treatment regressor.

3. To see this, let $T_i = \delta_0 + \delta_1 X_i + e_i$, where we without loss assume that the constant $\delta_0 = 0$. Then, consider that $\beta_{T_i, \tilde{T}_i} = \frac{\text{cov}(T_i, \tilde{T}_i)}{\text{var}(\tilde{T}_i)} = \frac{\text{cov}(E[T_i|X_i] + \tilde{T}_i, \tilde{T}_i)}{\text{var}(\tilde{T}_i)} = \frac{\text{cov}(E[T_i|X_i], \tilde{T}_i) + \text{cov}(\tilde{T}_i, \tilde{T}_i)}{\text{var}(\tilde{T}_i)} = 1$ as $\text{cov}(E[T_i|X_i], \tilde{T}_i) = 0$.
4. Because unobserved individual characteristics may lead to violation of rank similarity (Frandsen and Lefgren 2018), inclusion of fixed effects may also contribute to maintaining this assumption.
5. Including Stata (qreg), R (quantreg), Python (QuantReg), SAS (quantreg), SPSS (quantile regression), and more. For the time being, the GQR model is available only in Stata (genqreg) and the PS-QTE model is available in Stata (ivqte) and R (qte). The RQR model can be estimated in Stata using the rqr command (Borgen et al. 2021).
6. For example, the QTE is -0.40 at the 10th percentile, 0 at the 50th percentile, and 0.40 at the 90th percentile.
7. The supplementary data to Firpo et al.'s (2009) article did not include the household identification numbers necessary to include high-dimensional fixed effects. Figure A3 in the online supplement replicates the results of Figure 3 using the exact same data as Firpo et al. (2009).
8. Figure A6 in the online supplement shows the RQR coefficients with individual-level fixed effects.

References

- Allison, Paul D. 2009. *Fixed Effects Regression Models*. Thousand Oaks, CA: Sage.
- Angrist, Joshua D., and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press.
- Borgen, Nicolai T. 2016. "Fixed Effects in Unconditional Quantile Regression." *Stata Journal* 16(2): 403–15.
- Borgen, Nicolai T., Andreas Haupt, and Øyvind N. Wiborg. 2021. "Flexible and Fast Estimation of Quantile Treatment Effects: The rqr and rqrplot Commands." SocArXiv. <https://doi.org/10.31235/osf.io/4vquh>
- Borgen, Nicolai T., Andreas Haupt, and Øyvind Nicolay Wiborg. 2023. "Quantile Regression Estimands and Models: Revisiting the Motherhood Wage Penalty Debate." *European Sociological Review* 39(2): 317–31.
- Brini, Elisa, Solveig Topstad Borgen, and Nicolai T. Borgen. 2025. "Avoiding the Eyeballing Fallacy: Visualizing Statistical Differences between Estimates Using the Pheatplot Command." *Stata Journal* 25(1):77–96.
- Canay, Ivan A. 2011. "A Simple Approach to Quantile Regression for Panel Data." *Econometrics Journal* 14(3):368–86.
- Chernozhukov, Victor, Iván Fernández-Val, and Blaise Melly. 2013. "Inference on Counterfactual Distributions." *Econometrica* 81(6):2205–68.
- Chernozhukov, Victor, Iván Fernández-Val, and Blaise Melly. 2020. "Quantile and Distribution Regression in Stata: Algorithms, Pointwise and Functional Inference." Unpublished paper.
- Chernozhukov, Victor, Iván Fernández-Val, and Blaise Melly. 2022. "Fast Algorithms for the Quantile Regression Process." *Empirical Economics* 62: 7–33.
- Cooke, Lynn Prince. 2014. "Gendered Parenthood Penalties and Premiums across the Earnings Distribution in Australia, the United Kingdom, and the United States." *European Sociological Review* 30(3):360–72.
- Cukrowska-Torzewska, Ewa, and Anna Matysiak. 2020. "The Motherhood Wage Penalty: A Meta-Analysis." *Social Science Research* 88–89:102416.
- Cunningham, Scott. 2020. "Causal Inference: The Mixtape (V. 1.8)." Retrieved May 14, 2026. https://www.scunning.com/causalinference_norap.pdf.
- Dong, Yingying, and Shu Shen. 2018. "Testing for Rank Invariance or Similarity in Program Evaluation." *Review of Economics and Statistics* 100(1):78–85.
- Elwert, Felix, and Christopher Winship. 2014. "Endogenous Selection Bias: The Problem of Conditioning on a Collider Variable." *Annual Review of Sociology* 40:31–53.

- England, Paula, Jonathan Bearak, Michelle J. Budig, and Melissa J. Hodges. 2016. "Do Highly Paid, Highly Skilled Women Experience the Largest Motherhood Penalty?" *American Sociological Review* 81(6):1161–89.
- Filoso, Valerio. 2013. "Regression Anatomy, Revealed." *Stata Journal* 13(1):92–106.
- Firpo, Sergio. 2007. "Efficient Semiparametric Estimation of Quantile Treatment Effects." *Econometrica* 75(1):259–76.
- Firpo, Sergio, Nicole M. Fortin, and Thomas Lemieux. 2009. "Unconditional Quantile Regressions." *Econometrica* 77(3):953–73.
- Firpo, Sergio, and Cristine Pinto. 2016. "Identification and Estimation of Distributional Impacts of Interventions Using Changes in Inequality Measures." *Journal of Applied Econometrics* 31(3):457–86.
- Fortin, Nicole, Thomas Lemieux, and Sergio Firpo. 2011. "Decomposition Methods in Economics." pp. 1–102 in *Handbook of Labor Economics*, Vol. 4, edited by O. Ashenfelter, and D. Card. Amsterdam, the Netherlands: Elsevier.
- Frandsen, Brigham R., and Lars J. Lefgren. 2018. "Testing Rank Similarity." *Review of Economics and Statistics* 100(1):86–91.
- Frisch, Ragnar, and Frederick V. Waugh. 1933. "Partial Time Regressions as Compared with Individual Trends." *Econometrica* 1(4):387–401.
- Frölich, Markus, and Blaise Melly. 2010. "Estimation of Quantile Treatment Effects with Stata." *Stata Journal* 10(3):423–57.
- Frumento, Paolo, and Matteo Bottai. 2016. "Parametric Modeling of Quantile Regression Coefficient Functions." *Biometrics* 72(1):74–84.
- Goldberger, Arthur S. 1991. *A Course in Econometrics*. Cambridge, MA: Harvard University Press.
- Goodman-Bacon, Andrew. 2021. "Difference-in-Differences with Variation in Treatment Timing." *Journal of Econometrics* 225(2):254–77.
- Greenland, Sander, Stephen J. Senn, Kenneth J. Rothman, John B. Carlin, Charles Poole, Steven N. Goodman, and Douglas G. Altman. 2016. "Statistical Tests, P Values, Confidence Intervals, and Power: A Guide to Misinterpretations." *European Journal of Epidemiology* 31(4):337–50.
- Hagemann, Andreas. 2017. "Cluster-Robust Bootstrap Inference in Quantile Regression Models." *Journal of the American Statistical Association* 112(517):446–56.
- Hao, Lingxin, and Daniel Q. Naiman. 2007. *Quantile Regression*. Thousand Oaks, CA: Sage.
- He, Xuming. 1997. "Quantile Curves without Crossing." *American Statistician* 51(2):186–92.
- Hill, Terrence D., Andrew P. Davis, J. Micah Roos, and Michael T. French. 2020. "Limitations of Fixed-Effects Models for Panel Data." *Sociological Perspectives* 63(3):357–69.
- Imai, Kosuke, and In Song Kim. 2021. "On the Use of Two-Way Fixed Effects Regression Models for Causal Inference with Panel Data." *Political Analysis* 29(3):405–15.
- Killewald, Alexandra, and Jonathan Bearak. 2014. "Is the Motherhood Penalty Larger for Low-Wage Women? A Comment on Quantile Regression." *American Sociological Review* 79(2):350–57.
- Kim, Ju Hyun, and Byoung G. Park. 2022. "Testing Rank Similarity in the Local Average Treatment Effects Model." *Econometric Reviews* 41(10):1265–86.
- Koenker, Roger. 2004. "Quantile Regression for Longitudinal Data." *Journal of Multivariate Analysis* 91(1):74–89.
- Koenker, Roger. 2005. *Quantile Regression*. Cambridge, UK: Cambridge University Press.
- Koenker, Roger. 2017. "Quantile Regression: 40 Years On." *Annual Review of Economics* 9:155–76.
- Koenker, Roger, and Kevin Hallock. 2001. "Quantile Regression: An Introduction." *Journal of Economic Perspectives* 15(4):43–56.
- Lancaster, Tony. 2000. "The Incidental Parameter Problem since 1948." *Journal of Econometrics* 95(2):391–413.
- Lechner, Michael, and Anthony Strittmatter. 2019. "Practical Procedures to Deal with Common Support Problems in Matching Estimation." *Econometric Reviews* 38(2):193–207.
- Lovell, Michael C. 1963. "Seasonal Adjustment of Economic Time Series and Multiple Regression Analysis." *Journal of the American Statistical Association* 58(304):993–1010.

- Lovell, Michael C. 2008. "A Simple Proof of the FWL Theorem." *Journal of Economic Education* 39(1): 88–91.
- Machado, José A. F., and J.M.C. Santos Silva. 2005. "Quantiles for Counts." *Journal of the American Statistical Association* 100(472):1226–37.
- Machado, José A. F., and J. M. C. Santos Silva. 2019. "Quantiles via Moments." *Journal of Econometrics* 213(1):145–73.
- Melly, Blaise, and Kaspar Wüthrich. 2017. "Local Quantile Treatment Effects." Pp. 145–64 in *Handbook of Quantile Regression*, edited by R. Koenker, V. Chernozhukov, X. He, and L. Peng. Boca Raton, FL: Chapman & Hall/CRC.
- Mooney, Christopher Z., and Robert D. Duval. 1993. *Bootstrapping: A Nonparametric Approach to Statistical Inference*. Beverly Hills, CA: Sage.
- Morgan, Stephen L., and Christopher Winship. 2015. *Counterfactuals and Causal Inference*. Cambridge, UK: Cambridge University Press.
- Myers, Jessica A., Jeremy A. Rassen, Joshua J. Gagne, Krista F. Huybrechts, Sebastian Schneeweiss, Kenneth J. Rothman, Marshall M. Joffe, and Robert J. Glynn. 2011. "Effects of Adjusting for Instrumental Variables on Bias and Precision of Effect Estimates." *American Journal of Epidemiology* 174(11):1213–22.
- Pearl, Judea. 2009. *Causality: Models, Reasoning, and Inference*. Cambridge, UK: Cambridge University Press.
- Pearl, Judea. 2011. "Invited Commentary: Understanding Bias Amplification." *American Journal of Epidemiology* 174(11):1223–27.
- Porter, Stephen R. 2015. "Quantile Regression: Analyzing Changes in Distributions Instead of Means." Pp. 335–81 in *Higher Education: Handbook of Theory and Research*, edited by L. W. Perna. Cham, Switzerland: Springer.
- Powell, David. 2020. "Quantile Treatment Effects in the Presence of Covariates." *Review of Economics and Statistics* 102(5):994–1005.
- Powell, David. 2022. "Quantile Regression with Nonadditive Fixed Effects." *Empirical Economics* 63(5): 2675–91.
- Pregibon, Daryl. 1980. "Goodness of Link Tests for Generalized Linear Models." *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 29(1):15–14.
- Ramsey, James Bernard. 1969. "Tests for Specification Errors in Classical Linear Least-Squares Regression Analysis." *Journal of the Royal Statistical Society: Series B (Methodological)* 31(2):350–71.
- Ren, Chunhui, and Paul Allison. 2025. "Time-Invariant Variables' Time-Varying Effects: Misinterpretations of the Fixed-Effects Model in Sociological Research." *Sociology Compass* 19(9): e70113.
- Rios-Avila, Fernando. 2020. "Recentered Influence Functions (RIFs) in Stata: RIF Regression and RIF Decomposition." *Stata Journal* 20(1):51–94.
- Rios-Avila, Fernando, and Michelle Lee Maroto. 2024. "Moving beyond Linear Regression: Implementing and Interpreting Quantile Regression Models with Fixed Effects." *Sociological Methods & Research* 53:639–82.
- Rios-Avila, Fernando, Leonardo Siles, and Gustavo J. Canavire-Bacarreza. 2024. "Estimating Quantile Regressions with Multiple Fixed Effects through Method of Moments." *IZA Discussion Paper 17262*. Retrieved May 14, 2026. <https://docs.iza.org/dp17262.pdf>.
- Wenz, Sebastian E. 2019. "What Quantile Regression Does and Doesn't Do: A Commentary on Petscher and Logan (2014)." *Child Development* 90(4):1442–52.

Author Biographies

Nicolai Borgen is an associate professor at the Centre for Research on Equality in Education at the University of Oslo. His recent research focuses on social stratification, school and neighborhood effects, and quantitative methodology. His work has appeared in the *Proceedings of the National Academy of Sciences*, *Social Forces*, and the *European Sociological Review*.

Andreas Haupt is a professor of sociology at the Karlsruhe Institute of Technology. He likes stringent lines of thought, nonselective samples, and good coffee.

Øyvind Wiborg is a professor of sociology at the University of Oslo and currently a guest researcher at WZB Berlin and DIW Berlin. His research focuses on social inequality, including health disparities and intergenerational mobility in education, work, income, and wealth. He holds a PhD from the University of Oslo and was a visiting scholar at the University of California-Berkeley. His work appears in journals such as *Sociology of Education*, *European Societies*, *Demography*, the *British Journal of Sociology*, *European Sociological Review*, and *Lancet Regional Health*.