

## What Quantile Regression Does and Doesn't Do: A Commentary on Petscher and Logan (2014)

Wenz, Sebastian E.

Postprint / Postprint

Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

GESIS - Leibniz-Institut für Sozialwissenschaften

### Empfohlene Zitierung / Suggested Citation:

Wenz, S. E. (2019). What Quantile Regression Does and Doesn't Do: A Commentary on Petscher and Logan (2014). *Child development*, 90(4), 1442-1452. <https://doi.org/10.1111/cdev.13141>

### Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

### Terms of use:


This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

This is the peer reviewed version of an article in:  
**Child Development, Vol. 90, 2019, No. 4, pp. 1442-1452**  
which has been published in final form at <https://doi.org/10.1111/cdev.13141>.  
This article may be used for non-commercial purposes in accordance with Wiley Terms and  
Conditions for Use of Self-Archived Versions.  
Page numbers have been adjusted to the publishers version, whereby this postprint is fully quotable.

This commentary is a response to Petscher and Logan (2014): <https://doi.org/10.1111/cdev.12190>

## What Quantile Regression Does and Doesn't Do: A Commentary on Petscher and Logan (2014)

Sebastian E. Wenz 

GESIS—Leibniz Institute for the Social Sciences

Petscher and Logan's (2014) description of quantile regression (QR) might mislead readers to believe it would estimate the relation between an outcome,  $y$ , and one or more predictors,  $x$ , at different quantiles of the unconditional distribution of  $y$ . However, QR models the conditional quantile function of  $y$  given  $x$  just as linear regression models the conditional mean function. This article's contribution is twofold: First, it discusses potential consequences of methodological misconceptions and formulations of Petscher and Logan's (2014) presentation by contrasting features of QR and linear regression. Second, it reinforces the importance of correct understanding of QR in empirical research by illustrating similarities and differences in various QR estimators and linear regression using simulated data.

The article "Quantile Regression in the Study of Developmental Sciences" by Petscher and Logan (2014), henceforth PL, aims at familiarizing the readers of *Child Development* with quantile regression (QR), a method introduced by Koenker and Bassett (1978) and—until PL—only rarely used in developmental and educational psychology. I highly appreciate such an endeavor. However, PL's contribution suffers from several inadequacies, most of which are built on one fundamental misconception about the quantile regression model (QRM) and the QR estimator (Koenker & Bassett, 1978): PL claim that QR would provide "estimates of the relations between the predictor(s) and outcome, [. . .] across multiple points of the outcome's distribution" (PL, p. 861), "conditional on the performance on the outcome" (PL, p. 878), or "conditional on the score of the outcome" (p. 880). These and similar statements (e.g., pp. 870–872) are inaccurate and might mislead readers to believe that the QRM models the relation between an outcome,  $y$ , and one or more predictors,  $x$ , *at or conditional on* different quantiles of the *unconditional or marginal* distribution of  $y$  with distribution function  $F_Y(y)$ . I will show that the QR estimator may, under certain conditions, consistently estimate the effects of a treatment  $d$  on the quantiles of the marginal distribution conditional on counterfactual treatment states but unconditional on covariates,  $F_{Y|D}(y|d)$ . Typically, however, the inclusion of covariates,  $x$ , will be necessary and the effects estimated by the QR estimator will be effects on the quantiles of  $F_{Y|D,X}(y|d, x)$ —that is, quantiles of the conditional distribution of  $y$  as defined by all treatment and control variables included in the model.

Based on and related to this key misconception, PL provide several wrong and misleading expressions, descriptions, and explanations of which I think the following are the most consequential: PL suggest, the QR estimator would obtain estimates through a two-step process, in which the first step would "[identify] which sample scores for  $Y$  are associated with the selected quantile(s) of interest" (p. 864), and a separate, second step would estimate the coefficients. This

---

For helpful comments and suggestions on earlier versions of this commentary, I would like to thank Kerstin Hoenig, Christoph Homuth, Tobias Linberg, Steffen Schindler, and the participants in the weekly doctoral seminar at the Bamberg Graduate School of Social Sciences, University of Bamberg, Germany, as well as five anonymous reviewers. I appreciate helpful discussions with Britta Gauly, Dennis Köthemann, Anja Perry, Miriam Schmaus, and Reinhard Schunck. Last but not least, I would like to thank Jessica Logan and Yaacov Petscher for comments on an earlier version of the manuscript that helped me to be more precise in my critique of Petscher and Logan (2014).

Correspondence concerning this article should be addressed to Sebastian E. Wenz, GESIS—Leibniz Institute for the Social Sciences, Unter Sachsenhausen 6-8, 50667 Cologne, Germany. Electronic mail may be sent to [sewenz@gmail.com](mailto:sewenz@gmail.com).

is false. Also, neither the corresponding Equations—(2) and (3)—nor PL’s explanations of them (pp. 865, 866) make any sense, as I will show here.

PL’s discussion of the QRM’s assumptions (pp. 864, 866) may easily evoke the erroneous impression that the QRM would require no assumptions at all or completely different assumptions than the linear regression model (LRM) to arrive at certain desirable properties of the corresponding estimators. Also, PL focus on and misrepresent the role of the normality assumption for both QRM and LRM (pp. 864, 866). Next, I discuss those assumptions of QRM and LRM applied researchers will mainly worry about.

Not surprisingly, PL’s faulty notion, QR would provide estimates at or conditional on quantiles of the unconditional distribution of  $y$ , affects their interpretation of the corresponding estimates (e.g., on pp. 866, 870–871, 872). PL falsely claim that an estimated slope coefficient from a QRM for the  $r$ th quantile should be interpreted “as the gap in performance on  $Y$  at the [ $r$ th] quantile for individuals who were average on  $X$  compared to individuals who were 1  $SD$  above the mean on  $X$ ” (p. 866). PL extend this misconception to the LRM, whose estimates they interpret “at the mean of  $Y$ ” (p. 866; also see pp. 869–870, 873).

Additionally, PL’s inaccurate interpretations of QRM coefficients as associations or effects “for students who have very high or very low math scores” (p. 872; see pp. 861, 875 for similar wordings) might mislead readers to believe that QR estimates effects *for* those observations that lie *at* or around the respective quantiles of the unconditional distribution of  $y$ ,  $F_Y(y)$  or, maybe,  $F_{Y|D}(y|d)$ .

Based on their misconceptions, PL seem to believe that the QRM would be suited particularly well to model nonlinear relations and, thus, they focus on “nonlinear patterns” (p. 867) in the data they analyze and discuss the “curvilinear” (pp. 870, 875) shapes in plots of slope coefficients from the QRM. However, the QRM is not any better suited for or in any way superior to the LRM when modeling nonlinear relations.

Contributions citing PL (e.g., Law, Rush, King, Westrupp, & Reilly, 2018; McIlraith, 2018; Simzar, Martinez, Rutherford, Domina, & Conley, 2015; Tighe & Schatschneider, 2016) show that readers are indeed misled and copy PL’s misconceptions about the QRM. However, these misconceptions are by no means limited to PL and the related literature but seem widespread in the social sciences (see, e.g., Budig & Hodges, 2010, and the commentary by Killewald & Bearak, 2014).

My aim is to correct these mistakes by PL through this article that is divided into different parts: In the first part I address inaccurate depictions in PL that concern the general idea and goal of the QRM, the QR estimator and its estimation criterion, the assumptions that must hold to arrive at consistent point estimates and valid standard errors, as well as the interpretation of the QR estimates. Oftentimes I will contrast the QRM with the LRM that is also partly misrepresented by PL. In the second part, I discuss the results of Monte Carlo simulations for different scenarios to further facilitate the reader’s understanding of what the QRM does and how to interpret estimates provided by different estimators. I then briefly summarize and conclude.

### Major Misconceptions in PL

#### *What the QRM Does and Does Not Aim For*

In contrast to what PL claim, the QRM models the *conditional quantile function* of  $y$  given  $x$  just as the LRM models the *conditional mean function* of  $y$  given  $x$ . Thus, the QRM, as proposed by Koenker and Bassett (1978), is a *conditional quantile model* just as the LRM is a “conditional means model” (PL, p. 863). Both LRM and QRM assume that an outcome,  $y$ , can be expressed as a *linear* function of one or more predictors,  $x$ , and an error term,  $\varepsilon$ , that contains all predictors of  $y$  not in  $x$ . Of course, both LRM and QRM can be used to model nonlinear relations—for example, through the inclusion of polynomials of degree two or higher ( $x^k, k \geq 2$ ), or categorical predictors entering the model as multiple dummy variables.

While in the LRM, the relation of  $x$  and  $y$  for individual  $i$  in the population is often written as

$$y_i = \mathbf{x}_i\boldsymbol{\beta} + \varepsilon_i \quad (1)$$

and the population regression function for the mean or expectancy, respectively, of  $y$ , conditional on  $x$  is written as

$$E(y_i|\mathbf{x}_i) = \mathbf{x}_i\boldsymbol{\beta}, \quad (2)$$

the QRM assumes that in the population  $x$  and  $y$  can be related as follows:

$$y_i = \mathbf{x}_i\boldsymbol{\beta} + \varepsilon_{i\tau}, \quad (3)$$

where subscript  $\tau$ ,  $0 < \tau < 1$ , denotes a quantile of  $y$ . The QRM then seeks to model the population regression function for quantile  $\tau$  of  $y$  conditional on  $x$ :

$$Q_\tau(y_i|\mathbf{x}_i) = \mathbf{x}_i\boldsymbol{\beta}_\tau, \quad (4)$$

Interestingly, only a special case of Equation (1) is correctly given in PL, namely  $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$  (PL: 864, Equation 1). More importantly, there are no correct accounts of either Equations (3) or (4) in PL. PL write Equation (3) as  $Y_i = \beta_0^{(\tau)} + \beta_1^{(\tau)} + \varepsilon_i^{(\tau)}$  (PL: 866, Equation 3)—that is, the predictor(s),  $X_i$  or  $\mathbf{x}_i$ , are missing. Also, in this general form, Equation (4) is not given anywhere in PL.

#### *The Estimation Criterion of the QR Estimator*

A common way to estimate the parameters of the conditional mean function,  $E(y|\mathbf{x}) = \mathbf{x}\boldsymbol{\beta}$ , is to choose the ordinary least squares (OLS) estimator,  $\widehat{\boldsymbol{\beta}}$ , that minimizes the *sum of squared residuals*,  $\sum_{i=1}^N (y_i - \mathbf{x}_i\widehat{\boldsymbol{\beta}})^2$  or simply  $\sum_{i=1}^N \widehat{\varepsilon}_i^2$ . In a seemingly straightforward extension (Koenker, 2005, pp. 9–10; Koenker & Bassett, 1978), the parameters of the conditional quantile function,  $Q_\tau(y|\mathbf{x}) = \mathbf{x}\boldsymbol{\beta}_\tau$ , are estimated by the QR estimator,  $\widehat{\boldsymbol{\beta}}_\tau$ , that minimizes the *sum of weighted absolute residuals*,

$$\sum_{i: y_i \geq \mathbf{x}_i\widehat{\boldsymbol{\beta}}_\tau} \tau |y_i - \mathbf{x}_i\widehat{\boldsymbol{\beta}}_\tau| + \sum_{i: y_i < \mathbf{x}_i\widehat{\boldsymbol{\beta}}_\tau} (1 - \tau) |y_i - \mathbf{x}_i\widehat{\boldsymbol{\beta}}_\tau|, \quad (5)$$

where  $y_i - \mathbf{x}_i\widehat{\boldsymbol{\beta}}_\tau$  yields the residuals,  $\widehat{\varepsilon}_{i\tau}$ . PL (p. 865) correctly acknowledge what that means: The absolute values of positive residuals are weighted by  $\tau$ , whereas the absolute values of negative residuals are weighted by  $1 - \tau$ . However, contrast Equation (5) with Equation (2) in PL,  $\sum_{i \in \{i|Y_i \geq \xi_\tau\}} \tau |Y_i - \xi_\tau| + \sum_{i \in \{i|Y_i < \xi_\tau\}} (1 - \tau) |Y_i - \xi_\tau|$ , where  $Y_i$  is supposed to be “the vector of independent variables” and  $\xi_\tau$  “the dependent variable” (p. 865)—that is, the regression parameters are missing. Also, PL seem to suggest that their Equation (3),  $Y_i = \beta_0^{(\tau)} + \beta_1^{(\tau)} + \varepsilon_i^{(\tau)}$ , is the second step of the estimation process and expresses “the association between  $Y$  and  $X$  at a given quantile” (p. 866)—surprisingly, it does not contain any  $x$ -variables.

#### *Assumptions of the QRM*

To obtain unbiased and consistent (LRM) or at least consistent (QRM) estimators, very similar assumptions have to be met by QRM and LRM: The OLS estimator of the LRM is unbiased and consistent only if the model  $E(y|\mathbf{x}) = \mathbf{x}\boldsymbol{\beta}$  is correctly specified, that is, the *zero conditional mean assumption*,  $E(\varepsilon|\mathbf{x}) = 0$ , holds. Similarly, the QR estimator of the QRM yields consistent estimates only if the  $\tau$ th quantile of  $\varepsilon_\tau$  is zero at every point in the distribution of  $\mathbf{x}$  (Hao & Naiman, 2007, p. 29; Wooldridge, 2010, p. 454). We may call this assumption *zero conditional quantile assumption* and write:  $Q_\tau(\varepsilon_\tau|\mathbf{x}) = 0$ . Arguably, applied researchers will have to worry most about this assumption; PL do not discuss it at all.

Also, neither for LRM nor QRM, PL discuss the implications of meeting versus violating the assumptions of independent and identically distributed (i.i.d.) errors. That is, errors that do not correlate with each other,  $\text{cov}(\varepsilon_i, \varepsilon_j) = 0$ , for all  $i \neq j$ , and have constant variance conditional on the  $x$ -variables,  $\text{Var}(\varepsilon_i|\mathbf{x}_i) = \sigma^2$ , also known as homoscedastic error variance. An LRM with i.i.d. errors in which the zero conditional mean assumption holds, fulfills the Gauss–Markov assumptions under which OLS is the best linear unbiased estimator (BLUE). In the QRM, assuming the errors to be i.i.d. allows derivation of asymptotic standard errors that are computationally simple and obtained quickly (Koenker, 2005, pp. 73–74). However, both assumptions—i.i.d. errors—are likely to be violated in many settings in the social and developmental sciences. Luckily, standard errors that are robust toward clustering as well as heteroscedastic error variance are available for both LRM (Angrist & Pischke, 2009, pp. 293–325; Huber, 1967; White, 1980) and QRM (e.g., Koenker, 2005, pp. 68–113; Parente & Santos Silva, 2016). In non-i.i.d. settings, resampling techniques such as the bootstrap, permutation tests, and rank-based techniques provide useful alternatives for inference in the QRM (see Koenker, 2005, pp. 81–113; He, 2018, for overviews).

Note that violating the assumption of identically distributed, that is homoscedastic, errors does not affect the OLS point estimates:  $\widehat{\boldsymbol{\beta}}$  remains an unbiased and consistent estimator of  $\boldsymbol{\beta}$  as long as the zero conditional mean assumption holds. In the QRM, however, a violation of the homoscedasticity assumption *does* affect the point estimates. And, while heteroscedastic errors might lead to the undesirable phenomenon of some larger quantile having lower predicted values than some lower quantile for the same values on  $\mathbf{x}$ —a phenomenon known as *quantile crossing* (Koenker, 2005, p. 55)—modeling heteroscedastic errors is a key reason for turning to the QRM, as I will show in more detail next. Remarkably, the words homoscedasticity or homoskedasticity, heteroscedasticity or heteroskedasticity, or related terms do not appear once in PL.

In their discussion, PL focus on the normality assumption (pp. 863–864, 866), an assumption that

is—in the context of the LRM—typically considered to be “much stronger” (Wooldridge, 2009, p. 118) than all other assumptions, but also the “least important” (Gelman & Hill, 2007, p. 46). Furthermore, the normality assumption only applies to  $\varepsilon$ , not to the unconditional distributions of  $y$  or the  $x$ -variables (cf. PL, p. 864). If the Gauss–Markov assumptions hold and the errors are normally distributed in the population,  $\varepsilon_i \sim N(0, \sigma^2)$ , the OLS estimator is BLUE and we can rely on exact statistical inference for significance tests and confidence intervals (Wooldridge, 2009, pp. 118, 805–807). However, even if the normality assumption is violated, the usual significance tests for the LRM are valid in large enough samples due to the asymptotic normality of OLS (Wooldridge, 2009, pp. 172–175). Only if neither normality assumption nor large-sample approximation hold, the usual  $t$  statistics and, thus,  $p$ -values are invalid. Note that the OLS estimator remains unbiased and consistent if both the normality assumption and the assumption of homoscedastic errors are violated, but the other assumptions hold. In the QRM, the normality assumption is dropped entirely.

### *Interpretation of QRM Coefficients*

Unlike PL suggest, in the QRM, a slope coefficient for a predictor  $x_1$ ,  $\beta_{1\tau}$ , indicates the amount of change in the *conditional quantile*  $\tau$  of  $y$ ,  $Q_\tau(y|x)$ , associated with a one-unit change in  $x_1$ —whether this change equals a standard deviation or not (cf. PL, p. 866). The intercept or constant in a QRM is interpreted as the predicted value for quantile  $\tau$  of  $y$  given  $x = 0$ , that is,  $\beta_{0\tau} \equiv Q_\tau(y|x = 0)$ .

To say that “the traditional linear regression interpretation of a slope” (p. 866) could not be applied to the QRM, since through increasing  $x$  by one unit, “an individual would no longer be at the [particular quantile modeled]” (p. 866) of  $y$ , ignores some and misrepresents other features of the QRM: Since the QRM is a linear model, the interpretation of  $\beta_{1\tau}$  holds regardless of the values of the covariates,  $x$ , and, unless we explicitly model nonlinear relations, it also holds regardless of the initial value of  $x_1$ . Most importantly, the QRM is concerned with quantiles of the distribution of  $y$  conditional on all covariates in the model,  $x$ ,  $F_{y|x}(y|x)$ , not—as PL suggest—the unconditional distribution of  $y$ ,  $F_Y(y)$ . Also, changing  $x_1$  does not move *individuals* away from that conditional quantile. In contrast, changing  $x_1$  by one unit moves the *distribution* of  $y$  conditional on  $x$  so that the *value* of the  $\tau$ th quantile of this distribution changes by the amount of  $\beta_{1\tau}$ . Contrary to PL’s conception, the QRM alone does not allow to interpret  $\beta_{1\tau}$  as affecting individuals at a particular—conditional or unconditional—quantile. This is possible only if individuals either exactly keep their ranks regardless of the values on  $x_1$  and  $x$ , that is, if the so-called *rank-invariance assumption* or *rank-preservation assumption*, respectively, holds, or if individuals at least do not systematically change their ranks, that is, if the *rank-similarity assumption* holds (Angrist & Pischke, 2009, p. 281; Chernozhukov & Hansen, 2005, p. 250; Koenker, 2005, pp. 26–31).

But even then, in contrast to the LRM, the interpretation remains conditional on  $x$ : While applying the law of iterated expectations to  $E(y|x) = x\beta$  yields  $E(y) = E(x)\beta$ , quantiles do not feature analogous properties. Thus, while LRM coefficients may be interpreted as changes not only in the *conditional* mean of  $y$ ,  $E(y|x)$ , but also in the *unconditional* mean of  $y$ ,  $E(y)$ , QRM coefficients cannot readily be interpreted as changes in unconditional quantiles of  $y$  (see Angrist & Pischke, 2009, pp. 31–33, 282; Wooldridge, 2009, pp. 735–736; Wooldridge, 2010, pp. 18–22, 453). Methods for investigating effects on quantiles of the unconditional distribution of  $y$  have been developed more recently (e.g., Chernozhukov, Fernández-Val, & Melly, 2013; Firpo, 2007; Firpo, Fortin, & Lemieux, 2009; Frölich & Melly, 2013; Gosling, Machin, & Meghir, 2000; Machado & Mata, 2005; Powell, 2016). While a detailed discussion of these methods is beyond the scope of this commentary, I will return to this issue in the next section.

### **Simulation Studies**

The simulation studies in this section are meant to illustrate how LRM and QRM behave in different scenarios. I imply that the QRM is estimated by the traditional conditional QR estimator (Koenker & Bassett, 1978) unless stated otherwise. All simulations were conducted using Stata 15.1 (College Station, TX: StataCorp LLC). For every scenario I simulated 10,000 LRMs—estimated using OLS—and 10,000 QRMs per quantile ( $\tau = .10, .25, .50, .75, .90$ ), each simulation with  $N = 1,000$  (scenarios in Table 1) or  $N = 2,000$  (scenarios in Table 2) observations. Figures 1–3 show data from one draw each. I chose scenarios that are illustrative with regard to the misconceptions and limitations in PL’s discussion. Do-files with Stata Syntax for all simulation studies and additional material are available at <https://osf.io/79gpf/>.

Table 1

Results for Intercept ( $\beta_0$ ) and Slope ( $\beta_1, \beta_2$ ) Coefficients of Simulated LRMs, Estimated Via OLS, and QRMs (Koenker & Bassett, 1978) for Quantiles .10, .25, .50, .75, and .90

Scenario	Coef	LRM $\beta_j$ (SE)	Quantile regressions				
			Q. <sub>.10</sub> $\beta_j$ (SE)	Q. <sub>.25</sub> $\beta_j$ (SE)	Q. <sub>.50</sub> $\beta_j$ (SE)	Q. <sub>.75</sub> $\beta_j$ (SE)	Q. <sub>.90</sub> $\beta_j$ (SE)
1a	$\beta_0$	0.00 (.16)	-6.40 (.28)	-3.37 (.22)	0.00 (.20)	3.37 (.22)	6.40 (.28)
	$\beta_1$	3.00 (.17)	1.73 (.25)	2.33 (.20)	3.00 (.18)	3.67 (.20)	4.28 (.25)
1b	$\beta_0$	1.00 (.04)	-0.27 (.08)	0.33 (.06)	1.00 (.06)	1.67 (.06)	2.27 (.08)
	$\beta_1$	-1.00 (.10)	-2.30 (.18)	-1.67 (.14)	-1.00 (.13)	-0.33 (.14)	0.30 (.18)
2a	$\beta_0$	0.00 (.11)	-3.06 (.23)	-1.45 (.15)	0.14 (.13)	1.65 (.14)	2.95 (.17)
	$\beta_1$	4.00 (.08)	4.00 (.16)	4.00 (.11)	4.00 (.09)	4.00 (.10)	4.00 (.12)
	$\beta_2$						
2b	$\beta_0$	0.00 (.09)	-2.56 (.16)	-1.35 (.12)	0.00 (.11)	1.35 (.12)	2.56 (.16)
	$\beta_1$	6.00 (.11)	6.00 (.19)	6.00 (.15)	6.00 (.14)	6.00 (.15)	6.00 (.19)
	$\beta_2$	-1.00 (.05)	-1.00 (.08)	-1.00 (.06)	-1.00 (.06)	-1.00 (.06)	-1.00 (.08)
3a	$\beta_0$	0.06 (.14)	-3.59 (.23)	-1.90 (.15)	-0.23 (.12)	1.33 (.12)	2.68 (.15)
	$\beta_1$	3.96 (.14)	3.38 (.20)	3.81 (.14)	4.35 (.11)	4.93 (.12)	5.48 (.14)
	$\beta_2$						
3b	$\beta_0$	0.00 (.10)	-2.57 (.17)	-1.35 (.14)	0.00 (.12)	1.35 (.14)	2.57 (.17)
	$\beta_1$	6.00 (.14)	4.74 (.14)	5.34 (.10)	6.00 (.09)	6.67 (.10)	7.26 (.14)
	$\beta_2$	-1.00 (.09)	-1.01 (.10)	-1.00 (.08)	-1.00 (.07)	-1.00 (.08)	-0.99 (.10)

Note. Scenarios 2a and 2b feature simple asymptotic SEs, Scenarios 1a, 1b, 3a, and 3b feature heteroscedasticity robust SEs. Each simulation performs 10,000 replications. Each replication samples  $N = 1,000$  observations. The scenarios' data generating processes are given in the captions of Figures 1 (Scenario 1a & 1b), 2 (Scenario 2a & 2b), and 3 (Scenario 3a & 3b). LRM = linear regression model; QRM = quantile regression model.

Table 2

Biases of Different Estimators for ATE and QTEs Compared to True ATE and QTEs

Scenario	ATE	QTEs					
		Q. <sub>.10</sub>	Q. <sub>.25</sub>	Q. <sub>.50</sub>	Q. <sub>.75</sub>	Q. <sub>.90</sub>	
True ATE/QTEs	.50	.17	.31	.48	.68	.87	
Bias of naïve estimates	.30	Koenker and Bassett (1978)	.42	.39	.31	.23	.18
		Firpo (2007)	.42	.39	.31	.23	.18
		Firpo et al. (2009)	.42	.41	.34	.26	.16
Bias when controlling for $z$	.00	Koenker and Bassett (1978)	-.07	-.02	.02	.04	.04
		Firpo (2007)	.00	.00	.00	.00	.00
		Firpo et al. (2009)	-.01	.00	.02	.03	-.01
Bias when controlling for $z, u$	.00	Koenker and Bassett (1978)	-.20	-.09	.02	.10	.16
		Firpo (2007)	.00	.00	.00	.00	.00
		Firpo et al. (2009)	-.01	.00	.02	.04	-.01

Note. Estimates for ATEs are obtained via OLS. Stata's `qreg` implements the traditional QR estimator (Koenker & Bassett, 1978), `ivqte` (Frölich & Melly, 2010) implements Firpo's (2007) estimator and `rifreg` implements Firpo et al.'s (2009) estimator. Each simulation performs 10,000 replications. Each replication samples  $N = 2,000$  observations:  $n = 1,500$  if  $z = 0$ ,  $n = 500$  if  $z = 1$ . Data generating process for all scenarios except for the true values is  $y = w + u$ , where  $w \sim N(2,1)$  if  $z = 0 \cap d = 0$ ,  $w \sim N(1,1)$  if  $z = 1 \cap d = 0$ ,  $w \sim N(2.5, \sqrt{2})$  if  $z = 0 \cap d = 1$ ,  $w \sim N(1.5, \sqrt{2})$  if  $z = 1 \cap d = 1$ ,  $u \sim N(1,1)$  if  $z = 0$ ,  $u \sim N(0,1)$  if  $z = 1$ ,  $d \sim B(1, .6)$  if  $z = 0$ , and  $d \sim B(1, .4)$  if  $z = 1$ . True ATE/QTEs are obtained using randomized data, where  $d \sim B(1, .5)$  for both  $z = 0$  and  $z = 1$ . QRM = quantile regression model; QTE = quantile treatment effects; ATE = average treatment effect.

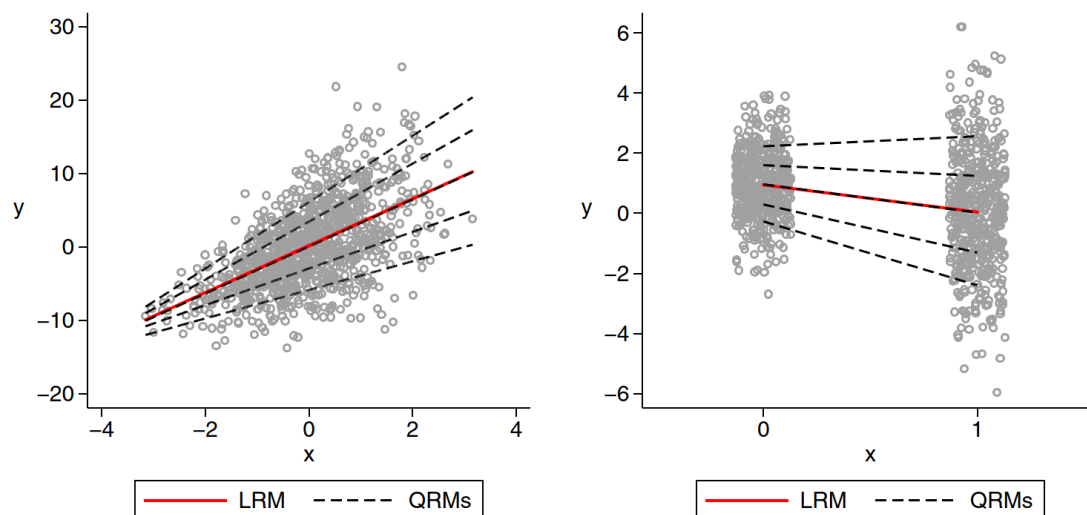


Figure 1. Linear relation with heteroscedastic errors in linear regression model (LRM) and quantile regression model (QRM): Scatterplot of  $y$  against  $x$  with fit lines from LRM (solid red line) and QRMs (dashed lines) for quantiles .10, .25, .50, .75, and .90 from bottom to top.

Source: Simulated data,  $N = 1,000$  randomly drawn observations. Data generating process (dgp) for left panel (Scenario 1a):  $y = 3x + \varepsilon$ , where  $x \sim N(0,1)$  and  $\varepsilon \sim N(0,5+x)$ ; dgp for right panel (Scenario 1b):  $y \sim N(1,1)$  for all  $x = 0$  and  $y \sim N(0,2)$  for all  $x = 1$ .

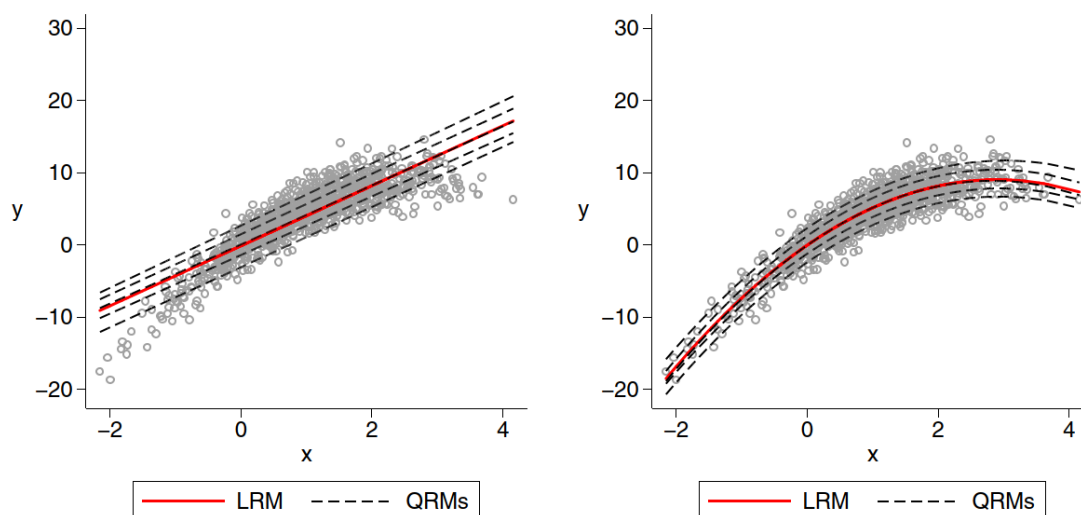


Figure 2. Nonlinear relation with homoscedastic errors in linear regression model (LRM) and quantile regression model (QRM). Scatterplot of  $y$  against  $x$  with fit lines from LRM (solid red line) and QRMs (dashed lines) for quantiles .10, .25, .50, .75, and .90 from bottom to top.

Source: Simulated data,  $N = 1,000$  randomly drawn observations. Data generating process for both panels:  $y = 6x - x^2 + \varepsilon$ , where  $x \sim N(1,1)$  and  $\varepsilon \sim N(0,2)$ . Left panel shows fit lines for incorrectly specified models leaving out  $x^2$  (Scenario 2a); right panel shows fit lines for correctly specified models including  $x^2$  (Scenario 2b).

### Unconditional Versus Conditional Quantiles

Scenarios 1a and 1b (see Figure 1, Table 1, and—for the data generating processes—the caption of Figure 1) should help to understand that the QRM does not model effects *at* quantiles of the *unconditional* distribution of  $y$  in the sense of  $F_Y(y)$ . The following calculation illustrates the differences between conditional and unconditional quantiles of  $y$ : Using the data from one draw of  $N = 1,000$  observations for Scenario 1a (Figure 1, left panel), the QRM for  $\tau = .90$  yields:  $Q_{.90}(y|x) = 6.08 + 4.52x$  (see Table 1 for averages over 10,000 draws). For  $x = -2$ , we get  $Q_{.90}(y|x = -2) = -2.96$ , the predicted .90th quantile of  $y$  at  $x = -2$ . However,  $-2.96$  roughly corresponds to the .33rd quantile of the *unconditional* distribution of  $y$ .

Now, the slope estimate,  $\hat{\beta}_{1(Q_{.90})} = 4.52$ , tells us that a unit-change in  $x$  is associated with a change of 4.52 units in the .90th quantile of  $y$ —conditional on  $x$ . Put differently, the distribution of  $y$  conditional on  $x$  changes in such a way that its .90th quantile moves up by 4.52 units:  $Q_{.90}(y|x = -2) + \hat{\beta}_{1(Q_{.90})} \equiv Q_{.90}(y|x = -1) = 1.56$ , a value that roughly corresponds to the .64th quantile of  $y$ . If the rank-preservation assumption or at least the

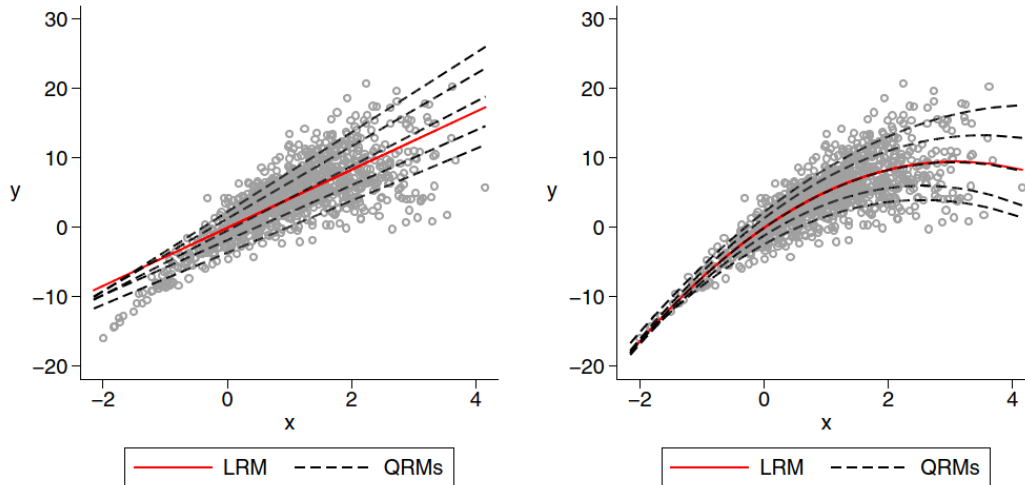


Figure 3. Nonlinear relation with heteroscedastic errors in linear regression model (LRM) and quantile regression model (QRM): Scatterplot of  $y$  against  $x$  with fit lines from LRM (solid red line) and QRMs (dashed lines) for quantiles .10, .25, .50, .75, and .90 from bottom to top.

Source: Simulated data,  $N = 1,000$  randomly drawn observations. Data generating process for both panels:  $y = 6x - x^2 + \varepsilon$ , where  $x \sim N(1,1)$  and  $\varepsilon \sim N(0,2 + x)$ . Left panel shows fit lines for incorrectly specified models leaving out  $x^2$  (Scenario 3a); right panel shows fit lines for correctly specified models including  $x^2$  (Scenario 3b).

rank-similarity assumption holds,  $\hat{\beta}_{1(Q_{.90})}$  can be interpreted as the effect for those observations that lie at the .90th quantile of  $y$  conditional on  $x$ —for example, those at  $(x, y) = (-2, -2.96)$  or  $(x, y) = (-1, 1.56)$  or at any other point on the fit line  $Q_{.90}(y|x) = 6.08 + 4.52x$ . However, observations that lie at the .90th quantile of the unconditional distribution of  $y$ ,  $Q_{.90}(y) = 8$ , may lie on fit lines from QRMs for other conditional quantiles—for example,  $Q_{.50}(y|x = 2.49) = .04 + 3.20 \times 2.49 = 8$ —that feature different slopes and, thus, different effects.

The problem is virtually the same in Scenario 1b, where no covariates are present, too, and  $x$  is dichotomous: Just like OLS estimates the mean difference between the groups defined by  $x = 0$  and  $x = 1$  in  $y$ , the QR for quantile  $\tau$  yields an estimate of the difference between the  $\tau$ th quantile of  $y|x = 0$  and the  $\tau$ th quantile of  $y|x = 1$ . A closer look at the fit lines in the right panel of Figure 1 suffices to see that they connect observations that differ on  $y$ . So, even if rank-preservation holds in a bivariate scenario (cf. Killewald & Bearak, 2014, p. 351), referring to group differences as differences or associations at quantiles of the unconditional distribution of  $y$ ,  $F_{Y(y)}$ , does not make a lot of sense, since the difference between observations from two different groups that lie at the same quantile of the unconditional distribution of  $y$  is, of course, 0.

### Quantile Treatment Effects

While estimated effects from a QRM never refer to  $F_Y(y)$  but always to  $F_{Y|x}(y|x)$ , where  $x$  contains all right hand side variables from the respective regression Equation in (4), it can be helpful to distinguish between a treatment variable,  $d$ , and control variables,  $x$ , to define  $F_{Y|D,X}(y|d, x)$ . Indeed, the recent literature on quantile treatment effects (QTEs) focuses on *unconditional* QTEs that are conceptualized as the difference between the quantile functions of the marginal potential distributions of treatment and control, formally  $F_{Y|D=1}^{-1}(\tau) - F_{Y|D=0}^{-1}(\tau)$  (Koenker, 2017, p. 158) or, using a more common notation as in Melly and Wüthrich (2018, p. 146),

$$\Delta(\tau) = Q_{Y_1}(\tau) - Q_{Y_0}(\tau) \quad (6)$$

That is, the counterfactual or potential distributions,  $Y_1$  and  $Y_0$ , are conditional on the potential or counterfactual causal states  $D = d$  but unconditional on covariates and, thus, concern  $F_{Y|D}(y|d)$  but not  $F_{Y|D,X}(y|d, x)$ . Therefore, for the special case of randomized controlled trials without any covariates, the QR estimator (Koenker & Bassett, 1978) will identify and estimate the QTE from Equation (6), since the distributions of  $y$  conditional on causal states can be seen as two separate marginal, that is, unconditional, distributions representing different potential states of the world (e.g., Djebbari & Smith, 2008, p. 69; Frölich & Melly, 2013, p. 346; Heckman, Smith, & Clements, 1997, p. 495). When covariates are included—for example, to address noncompliance or increase the precision of estimates—this interpretation is lost even for experimental data.

As an example, take Scenario 1b (right panel of Figure 1) and imagine an experimental study



examining the causal effect of a training program,  $x$ , on test anxiety,  $y$ . Here,  $y|x = 0$  would be the unconditional distribution of test anxiety for the untreated control group,  $Y_0$ , and  $y|x = 1$  would be the unconditional distribution of test anxiety for the treatment group,  $Y_1$ . If rank-invariance or at least rank-similarity holds, the slopes from the QRMs for different quantiles imply that the program reduced test anxiety most effectively for students at the .10th quantile but increased it for students at the .90th quantile of  $F_{Y|x}(y|x)$ —both before and after treatment. To learn about such heterogeneity in treatment effects should be of substantive interest to researchers, policy makers, and practitioners.

#### *Heteroscedasticity, Higher Moments, and Beyond*

The heterogeneous treatment effect in the example above is due to the data generating process of Scenario 1b that creates heteroscedastic error variance. Scenario 1a also features heteroscedasticity. However, from the single lines of both LRM and QRM we merely learn whether a change in  $x$  is associated with a change in the mean of  $y$  and the conditional .10th, .25th, .50th, .75th, and .90th quantile of  $y$ . It is the simultaneous estimation of QRMs for different quantiles that enables researchers to detect and describe the change in shape of the distribution of  $y$  conditional on  $x$  and, thus, heteroscedasticity but also higher moments such as skewness and kurtosis and how they change with  $x$ . In fact, QR is even more powerful, since “quantiles are inherently local” (Koenker, 2017, p. 156) and can be used to detect and describe the smallest and most local changes in a distributional mass. That the QRM is particularly well suited to detect and describe heteroscedasticity (Koenker & Bassett, 1982) is not discussed by PL at all. A discussion of the QRM’s capacity to model higher moments is also not provided by PL and beyond the scope of this commentary (but see Hao & Naiman, 2007, pp. 12–14).

#### *Nonlinear Relations*

In contrast to what PL suggest by focusing on “nonlinear patterns” (p. 867) in analyzing their data, the QRM does not capture nonlinear relations unless explicitly modeled. Scenarios 2a and 2b feature a common data generating process in which  $y$  and  $x$  have a nonlinear relation, but the errors are homoscedastic. The models, LRM and QRMs, of Scenario 2a (Table 1, left panel of Figure 2) were misspecified on purpose by leaving out  $x^2$ .

Following PL’s diction, we should nevertheless expect larger slopes,  $\hat{\beta}_{1\tau}$ , for lower quantiles than for higher quantiles of  $F_Y(y)$ , since, as is obvious from Figure 2, for or at lower values of  $y$  a one-unit change in  $x$  results in a larger increase in  $y$  than for or at higher values of  $y$ . However, the lines from both LRM and all QRMs run in parallel and  $\hat{\beta}_1$  does not differ between QRMs for different quantiles. The models of Scenario 2b (Table 1, right panel of Figure 2) were correctly specified by including  $x^2$ . The lines still run in parallel but now fit the data much better, since the nonlinear or curvilinear relation of  $x$  and  $y$  is modeled properly by including  $x^2$ —not by running QRMs for different quantiles. Figure 3 shows what happens if heteroscedastic error variance is added to the nonlinear relation.

Again, I misspecified the models of Scenario 3a (see left panel of Figure 3) by excluding  $x^2$ , but correctly specified the models of Scenario 3b (see right panel of Figure 3) by including  $x^2$ . Now that the variance of  $y$  increases with  $x$  in both scenarios, the slopes for  $x_1$  differ—only that, following PL, we would probably expect the opposite pattern again, namely steeper slopes for lower quantiles than for higher quantiles. Using correct specifications such as in Scenario 3b enables researchers to simultaneously model a nonlinear location shift and a shape shift.

#### *Multiple Regression: Conditional Versus Unconditional QR*

Relying on their inaccurate conceptualizations, PL erroneously interpret the estimates of multiple QRMs as effects, differences, or “gaps” (p. 873) at different quantiles or “at different points of the outcome” (p. 875), by which they obviously refer to different quantiles of the unconditional distribution of  $y$  (e.g., “moderate math achievement scores,” “those students with higher or lower scores,” p. 875). PL (p. 873) correctly state that the logic underlying a multiple QRM is the same as underlying a multiple LRM: Including multiple covariates allows estimating the association of  $x$  and  $y$  net of the other covariate(s). However, since PL miss the marked differences between conditional and unconditional quantiles and the importance of this distinction for understanding the QRM, PL do not see that adding covariates typically exacerbates the differences between conditional and unconditional quantiles and that the slopes of a multiple QRM cannot be interpreted as effects on the quantiles of the unconditional distribution of  $y$ —be it in the sense of  $F_Y(y)$  or  $F_{Y|D}(y|d)$ . This is because the QR estimator

(Koenker & Bassett, 1978) assesses changes in the conditional distribution,  $F_{Y|D,X}(y|d, \mathbf{x})$ , and because—in contrast to the mean—in case of quantiles it depends on the distribution of the covariates,  $\mathbf{x}$ , how these changes affect  $F_{Y|D}(y|d)$ .  $F_Y(y)$ , of course, is affected by the distribution of both  $\mathbf{x}$  and  $d$  (Powell, 2013, 2016).

That the inclusion of covariates may also serve the purpose of statistically and substantively explaining heteroscedasticity and, hence, heterogenous treatment effects, is also not discussed by PL.

Table 2 presents results of simulations that illustrate the consequences of adding covariates for the OLS estimator, the traditional QR estimator (Koenker & Bassett, 1978), the estimator suggested by Firpo (2007), and the one suggested by Firpo et al. (2009). In contrast to the traditional conditional QR estimator, the latter seek to provide estimates for effects on quantiles of the unconditional distribution of  $y$ , generally speaking. Firpo (2007) uses a semiparametric two-step approach in which the first step is estimating a propensity score that is then used to calculate a reweighed version of the traditional QR estimator (Koenker & Bassett, 1978). Firpo et al. (2009) propose to regress what they call the *recentered influence function*, of the unconditional quantile of the outcome on the predictors.

Imagine we are interested in the average treatment effect (ATE) and the QTEs of an extracurricular activity,  $d$ , on a math achievement score,  $y$ , in the sense of Equation (6),  $\Delta(.10, .25, .50, .75, .90)$ . In the observational data at hand, naïve estimates of ATE and QTEs will be biased (Table 2, “Bias of naïve estimates”), since ethnic majority students,  $z = 0$ , are more likely to take part in the activity than ethnic minority students,  $z = 1$ , and have a higher mean on two covariates,  $w$  and  $u$ , that both predict  $y$ , and of which only  $w$  is affected by the treatment (see caption of Table 2 for detailed data generating processes).

So, controlling for  $z$  is necessary and sufficient to identify the ATE and QTEs. Once  $z$  is controlled (Table 2, “Bias when controlling for  $z$ ”), OLS provides an unbiased estimate of the ATE as does Firpo’s (2007) estimator for the QTEs. The “first-order approximation” (Chernozhukov et al., 2013, p. 2213) to the QTE in Equation (6) provided by Firpo et al. (2009) yields slightly biased estimates except for  $\Delta(.25)$ . However, the traditional QR estimator (Koenker & Bassett, 1978) yields biased estimates for all QTEs—it overestimates the effects on the .50th, .75th, and .90th quantile but underestimates the effects on the .25th and .10th quantile of  $F_{Y|D}(y|d)$ . Including  $u$  in the model is not necessary for identification but decreases the standard errors of all estimates (not shown in Table 2; see Supporting Information at <https://osf.io/79gpf/>). However, it *increases* the bias of the QR estimator (Koenker & Bassett, 1978). While the estimator by Firpo (2007) provides unbiased estimates again, the biases of estimates provided by the estimator from Firpo et al. (2009) remain virtually unchanged.

To conclude, if interest lies in effects on unconditional quantiles of  $y$  and it is necessary to control for covariates, researchers cannot rely on the traditional QR estimator (Koenker & Bassett, 1978), but need to turn to alternative *unconditional* QR estimators. Again, OLS provides unbiased estimates of the ATE, but misses the heterogeneous effect of the treatment. Finally, a note of caution: Results from this example with a dichotomous treatment cannot readily be generalized to other scenarios. This applies to performance and bias of all estimators evaluated.

### Summary and Conclusion

PL have called the attention of researchers in the developmental and educational sciences to QR (Koenker, 2005; Koenker & Bassett, 1978), a method largely unknown and only rarely used in these fields. Unfortunately, PL’s presentation conveys a fundamental misconception about the QRM, namely that it would model the relation of one or more predictors,  $\mathbf{x}$ , and an outcome,  $y$ , at different quantiles of the unconditional distribution of  $y$ .

In contrast, the QRM models the *conditional quantile function* of  $y$  given  $\mathbf{x}$ . Estimates of the QR estimator (Koenker & Bassett, 1978) should be interpreted as *differences between* or *effects on* quantiles of the *conditional* distribution of  $y$  given  $\mathbf{x}$ . Only under rank-invariance or rank-similarity the QRM may allow statements about observations. Also, to examine effects on the unconditional distribution of  $y$ , more recent approaches (e.g., Chernozhukov et al., 2013; Firpo, 2007; Firpo et al., 2009; Powell, 2016) need to be applied. Porter (2015) and Killewald and Bearak (2014), for instance, provide accessible discussions of one of these approaches—the *Unconditional Quantile Regression Model* (Firpo et al., 2009). More advanced contributions that also cover other unconditional QR estimators include Frölich and Melly (2010, 2013), Melly and Wüthrich (2018), and Powell (2013, 2016). Furthermore, the QRM is not suited particularly well to model nonlinear relations but indeed for detecting and describing heteroscedasticity. In contrast to PL, I have discussed the most important assumptions of the

QRM and showed that the differences to the LRM are less pronounced than PL suggest.

The numerous accounts of inaccurate and misleading expressions, descriptions, and explanations, as well as the considerable amount of important information missing in PL might, and most likely will, lead readers to misconceive what QRM and QR estimator (Koenker & Bassett, 1978) can and cannot do. Therefore, I suggest avoiding PL as a source of learning about QR entirely. Instead I recommend Koenker and Bassett (1978), Koenker (2005, 2017), and Koenker, Chernozhukov, He, and Peng (2018) to the mathematically inclined readers of *Child Development*. Introductory treatments that may be more accessible to most applied researchers are provided, among others, by Angrist and Pischke (2009), Cade & Noon (2003), Fitzenberger and Wilke (2015), Hao and Naiman (2007), Koenker and Hallock (2001), and Porter (2015). Understood and used correctly, conditional and unconditional forms of QR may help to answer old research questions in a new way or inspire researchers to ask and answer research questions they have not yet asked.

## References

- Angrist, J. D., & Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton, NJ: Princeton University Press.
- Budig, M. J., & Hodges, M. J. (2010). Differences in disadvantage variation in the motherhood penalty across white women's earnings distribution. *American Sociological Review*, *75*, 705–728. <https://doi.org/10.1177/0003122410381593>
- Cade, B. S., & Noon, B. R. (2003). A gentle introduction to quantile regression for ecologists. *Frontiers in Ecology and the Environment*, *1*, 412–420. [https://doi.org/10.1890/1540-9295\(2003\)001\[0412:AGITQR\]2.0.CO;2](https://doi.org/10.1890/1540-9295(2003)001[0412:AGITQR]2.0.CO;2)
- Chernozhukov, V., Fernández-Val, I., & Melly, B. (2013). Inference on counterfactual distributions. *Econometrica*, *81*, 2205–2268. <https://doi.org/10.3982/ECTA10582>
- Chernozhukov, V., & Hansen, C. (2005). An IV model of quantile treatment effects. *Econometrica*, *73*, 245–261. <https://doi.org/10.1111/j.1468-0262.2005.00570.x>
- Djebbari, H., & Smith, J. (2008). Heterogeneous impacts in PROGRESA. *Journal of Econometrics*, *145*, 64–80. <https://doi.org/10.1016/j.jeconom.2008.05.012>
- Firpo, S. (2007). Efficient semiparametric estimation of quantile treatment effects. *Econometrica*, *75*, 259–276. <https://doi.org/10.1111/j.1468-0262.2007.00738.x>
- Firpo, S., Fortin, N. M., & Lemieux, T. (2009). Unconditional quantile regressions. *Econometrica*, *77*, 953–973. <https://doi.org/10.3982/ECTA6822>
- Fitzenberger, B., & Wilke, R. A. (2015). Quantile Regression Methods. In R. Scott, & S. Kosslyn (Eds.), *Emerging trends in the social and behavioral sciences* (pp. 1–18). Wiley. <https://doi.org/10.1002/9781118900772.etrds0269>
- Frölich, M., & Melly, B. (2010). Estimation of quantile treatment effects with Stata. *Stata Journal*, *10*, 423–457. Retrieved from <https://www.stata-journal.com/article.html?article=st0203>
- Frölich, M., & Melly, B. (2013). Unconditional quantile treatment effects under endogeneity. *Journal of Business & Economic Statistics*, *31*, 346–357. <https://doi.org/10.1080/07350015.2013.803869>
- Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge, MA: Cambridge University Press.
- Gosling, A., Machin, S., & Meghir, C. (2000). The changing distribution of male wages in the UK. *The Review of Economic Studies*, *67*, 635–666. <https://doi.org/10.1111/1467-937X.00148>
- Hao, L., & Naiman, D. Q. (2007). *Quantile regression*. Thousand Oaks, CA: Sage. <https://doi.org/10.4135/9781412985550>
- He, X. (2018). Resampling methods for quantile regression. In R. Koenker, V. Chernozhukov, X. He, & L. Peng (Eds.), *Handbook of quantile regression* (pp. 7–19). New York, NY: Chapman & Hall/CRC.
- Heckman, J. J., Smith, J., & Clements, N. (1997). Making the most out of programme evaluations and social experiments: Accounting for heterogeneity in programme impacts. *The Review of Economic Studies*, *64*, 487–535. <https://doi.org/10.2307/2971729>
- Huber, P. J. (1967). The behavior of maximum likelihood estimates under nonstandard conditions. In M. Luncien, L. Cam, & J. Neyman (Eds.), *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability* (Vol. 1, pp. 221–233). Berkeley, CA: University of California Press.
- Killewald, A., & Bearak, J. (2014). Is the motherhood penalty larger for low-wage women? A comment on quantile regression. *American Sociological Review*, *79*, 350–357. <https://doi.org/10.1177/0003122414524574>
- Koenker, R. (2005). *Quantile regression*. Cambridge, MA: Cambridge University Press. <https://doi.org/10.1017/CBO9780511754098>
- Koenker, R. (2017). Quantile regression: 40 years on. *Annual Review of Economics*, *9*, 155–176. <https://doi.org/10.1146/annurev-economics-063016-103651>
- Koenker, R., & Bassett, Jr., G. (1978). Regression quantiles. *Econometrica*, *46*, 33–50. <https://doi.org/10.2307/1913643>
- Koenker, R., & Bassett, Jr., G. (1982). Robust tests for heteroscedasticity based on regression quantiles. *Econometrica*, *50*, 43–61. <https://doi.org/10.2307/1912528>
- Koenker, R., Chernozhukov, V., He, X., & Peng, L. (Eds.). (2018). *Handbook of quantile regression*. Boca Raton, FL: Chapman & Hall.
- Koenker, R., & Hallock, K. F. (2001). Quantile regression. *Journal of Economic Perspectives*, *15*, 143–156. Retrieved from <http://www.jstor.org/stable/2696522>
- Law, J., Rush, R., King, T., Westrupp, E., & Reilly, S. (2018). Early home activities and oral language skills in

- middle childhood: A quantile analysis. *Child Development*, 89, 295–309. <https://doi.org/10.1111/cdev.12727>
- Machado, J. A. F., & Mata, J. (2005). Counterfactual decomposition of changes in wage distributions using quantile regression. *Journal of Applied Econometrics*, 20, 445–465. <https://doi.org/10.1002/jae.788>
- McIlraith, A. L., & Language and Reading Research Consortium. (2018). Predicting word reading ability: A quantile regression study. *Journal of Research in Reading*, 41, 79–96. <https://doi.org/10.1111/1467-9817.12089>
- Melly, B., & Wüthrich, K. (2018). Local quantile treatment effects. In R. Koenker, V. Chernozhukov, X. He, & L. Peng (Eds.), *Handbook of quantile regression* (pp. 145–164). Boca Raton, FL: Chapman & Hall.
- Parente, P. M. D. C., & Santos Silva, J. M. C. (2016). Quantile regression with clustered data. *Journal of Econometric Methods*, 5, 1–15. <https://doi.org/10.1515/jem-2014-0011>
- Petscher, Y., & Logan, J. A. R. (2014). Quantile regression in the study of developmental sciences. *Child Development*, 85, 861–881. <https://doi.org/10.1111/cdev.12190>
- Porter, S. R. (2015). Quantile regression: Analyzing changes in distributions instead of means. In M. B. Paulsen (Ed.), *Higher education: Handbook of theory and research* (pp. 335–381). Cham, Switzerland: Springer International. [http://link.springer.com/chapter/10.1007/978-3-319-12835-1\\_8](http://link.springer.com/chapter/10.1007/978-3-319-12835-1_8)
- Porter, D. (2013). *A new framework for estimation of quantile treatment effects*. Retrieved from [http://www.rand.org/pubs/working\\_papers/WR824-1.html](http://www.rand.org/pubs/working_papers/WR824-1.html)
- Powell, D. (2016). *Quantile treatment effects in the presence of covariates*. Retrieved from [https://works.bepress.com/david\\_powell/4/](https://works.bepress.com/david_powell/4/)
- Simzar, R. M., Martinez, M., Rutherford, T., Domina, T., & Conley, A. M. (2015). Raising the stakes: How students' motivation for mathematics associates with high- and low-stakes test achievement. *Learning and Individual Differences*, 39, 49–63. <https://doi.org/10.1016/j.lindif.2015.03.002>
- Tighe, E. L., & Schatschneider, C. (2016). A quantile regression approach to understanding the relations among morphological awareness, vocabulary, and reading comprehension in adult basic education students. *Journal of Learning Disabilities*, 49, 424–436. <https://doi.org/10.1177/0022219414556771>
- White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, 48, 817–838. <https://doi.org/10.2307/1912934>
- Wooldridge, J. (2009). *Introductory econometrics: A modern approach* (4th ed.). Mason, OH: Cengage Learning.
- Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data* (2nd ed.). Cambridge, MA: The MIT Press.

### Supporting Information

Additional supporting information may be found in the online version of this article at the publisher's website:

**Appendix S1.** Instructions for Replication

**Code S1.** Stata Syntax File: Sets Up Stata for Replication

**Code S2.** Stata Syntax File: Replicates Figures 1-3

**Code S3.** Stata Syntax File: Replicates Simulation Results in Table 1

**Code S4.** Stata Syntax File: Replicates Simulation Results in Table 2