1

Quantile Regression Applied to RETs

This document considers how to apply quantile regression to an RET using limited information SEM. I first describe the numerical example I use and then apply conditional quantile regression to the data. I then discuss unconditional quantile regression. I assume you have read the material on quantile regression in Chapters 6 and 8 and the document on my website for Chapter 6, *Quantile Regression*.

THE NUMERICAL EXAMPLE

The example I use is a three mediator RET where the outcome variable is a person's behavioral commitment to living a "green" (environmentally respectful) lifestyle. It was measured on a 0 to 100 multi-item scale three months after the intervention and reflected a multivariate profile of "green" behaviors the individual had performed in the past three months. The baseline mean and median scores for the measure were about 32. The program targeted three mediators, (1) increasing knowledge about the short term consequences of the climate/environmental crisis, (2) increasing knowledge about the long term consequences of the climate/environmental crisis, and (3) increasing knowledge about "green" technologies to deal with the climate crisis. Each of the mediators was scored on a knowledge test with 100 points possible. The influence diagram is in Figure 1, absent covariates to reduce clutter.

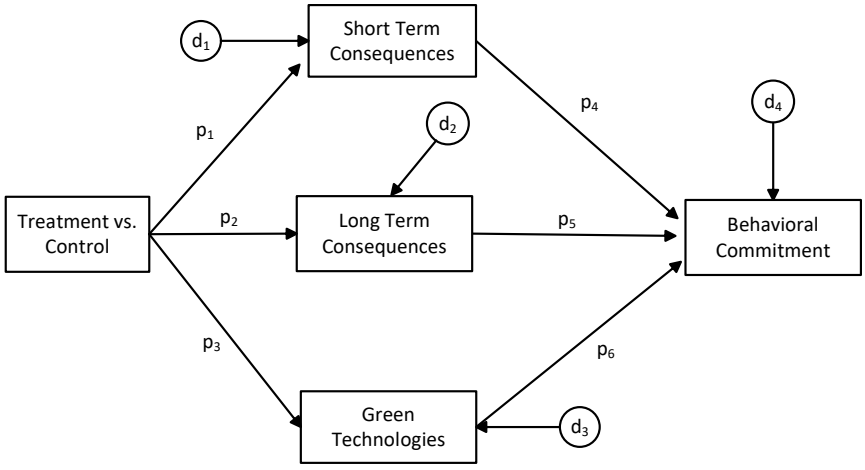


FIGURE 1. Numerical example

The study also obtained baseline measures of the mediators and the outcome as well as a covariate reflecting education, scored continuously from 1 to 10 with higher scores indicating greater educational achievement. In practice, an RET likely would include more covariates but I restrict the number here for pedagogical reasons. The control group was an active control that received educational materials on healthy eating. The core equations for the model, including the covariates, are (note: TREAT = the treatment condition, COV is the covariate of educational background, M1 is the mediator for knowledge about short term consequences; M2 is the mediator for knowledge about long term consequences; M3 is the mediator for technology knowledge; Y is the outcome; if a variable name ends with B, then it is the baseline version of the referent variable):

$$M1 = a_1 + p_1\ TREAT + b_1\ M1B + b_2\ COV + d_1 \quad [1]$$

$$M2 = a_2 + p_2\ TREAT + b_3\ M2B + b_4\ COV + d_2 \quad [2]$$

$$M3 = a_3 + p_3\ TREAT + b_5\ M3B + b_6\ COV + d_3 \quad [3]$$

$$Y = a_4 + p_4\ M1 + p_5\ M2 + p_6\ M3 + b_7\ YB + b_8\ COV + d_4 \quad [4]$$

These equations include the baseline measure of the endogenous outcome and COV as covariates, but researchers might choose not to, an issue I discuss in more detail shortly.

CONDITIONAL QUANTILE REGRESSION

Preliminary Analyses

The first step is to conduct preliminary analyses to determine if quantile regression is viable and is likely to provide useful insights relative to a more traditional OLS regression. Quantile regression assumes linearity for continuous predictors and outcomes, that the model is correctly specified, additivity among the predictors, independence of observations, and reasonably continuous endogenous variables. It does not require population normality for the disturbances but it does assume homoscedasticity. See Hao & Naiman (2007) and Wentz (2018) for more elaboration of the assumptions.

I use the program on my website called *Running interval smoother* to gain perspectives on linearity. The program creates a scatter plot between a continuous predictor and an outcome and plots the smooth for any specified quantile. For example, Figure 2 presents the plot for Y as a function of M1 with the smooth for the 0.20 quantile. The smooth is reasonably linear and this also was the case when I explored the other continuous predictors and the outcome for their 20^{th}, 50^{th}, and 80^{th} quantiles, which are the quantiles I focus my primary analysis on. The linearity assumption seems viable.

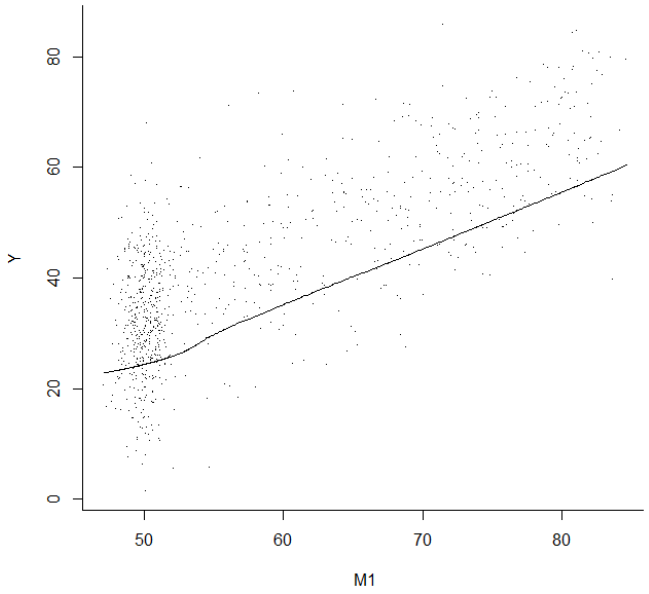
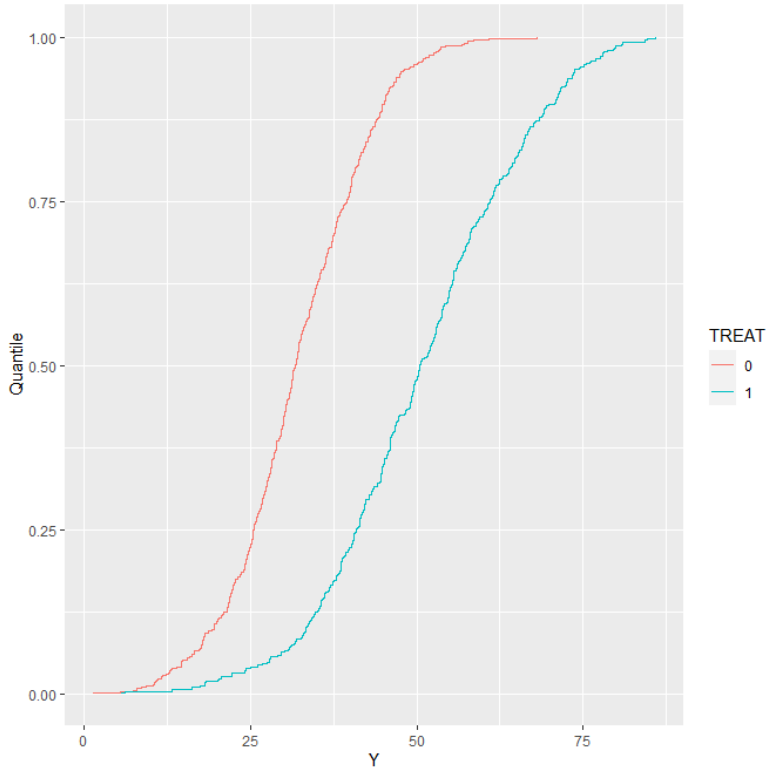


FIGURE 2. Running interval smoother

Examination of the frequency distributions for M1, M2, M3, and Y (the endogenous variables in the model) confirmed they were many-valued and functionally continuous with few ties. Quantile regression is robust to outliers but not necessarily to unusual leverages. I used the program on my website called *Leverage analysis* to identify individuals with unusual leverages in the predictor space and found six individuals to have unusual predictor profiles. For the primary quantile regression analyses I ultimately conducted, I did so both with and without the six cases included. The conclusions were comparable in the two analyses so I report the results for the total sample.

I also conducted preliminary analyses to gain perspectives on whether quantile regression is needed in place of OLS or classic maximum likelihood analyses. This involved performing a rough graphical analysis of the quantile treatment effect with no covariates. I used the program on my website called *Quantile plot* to plot the cumulative distribution functions for the outcome variable as a function of the treatment and control groups to evaluate the raw quantile treatment effect (see Chapter 8 for a description of the logic of these plots). If the treatment effects at the different quantiles are decidedly not uniform, then this suggests that quantile regression may be useful for revealing the lack of uniform effects across the distribution. Here is the plot for the current data:



Key is the amount of horizontal spread between the two lines at any given quantile. In this plot, you can see that the quantile difference between the treatment and control groups (i.e., the horizontal spread between lines) tends to be wider at the median than it is at the 10th quantile and that the spread is even larger at the 80th quantile. This suggests the traditional analysis of mean differences between the two groups will not tell the whole story about the effects of the intervention, which differs depending on where in the distribution we look. Here are the values for the deciles that the *Quantile plot* program produced that documents the decile differences algebraically:

Decile	Treatment	Control	Difference
0.1	33.34	19.22	14.12
0.2	38.85	24.18	14.66
0.3	42.95	26.98	15.97
0.4	46.71	29.58	17.14
0.5	50.77	31.75	19.02
0.6	54.48	34.38	20.09
0.7	58.23	37.54	20.69
0.8	63.89	40.90	22.98
0.9	70.32	45.10	25.22

The increase in behavioral commitment (the outcome) as a function of the intervention is about 14 units at the 10th decile; at the median the increase is about 19 units; and at the 90th decile, the increase is about 25 units. The question becomes are these differences statistically significantly different from one another, a question I answer in the formal analyses using the conditional quantile regression analyses reported later. However, it appears that the quantile analysis will indeed be meaningful.

Another issue I check during preliminary analyses are the covariates to include in the model. For questions about the total effect of the intervention on the outcome and the effect of the intervention on the mediators, some argue that we do not need covariates because of the use of random assignment to the treatment condition. Such assignment, the argument goes, guarantees that the treatment dummy variable will be uncorrelated or functionally uncorrelated with the baseline covariates rendering confounding dynamics for them moot. In traditional OLS and maximum likelihood based linear regression focused on intervention effects in RCTs, we typically include covariates for three reasons, (1) to increase statistical power of the tests (i.e., to increase the efficiency of our estimates), (2) to address bias that may be caused by missing data mechanisms or treatment dropouts (see Chapters 26 and 27), and/or (3) to address sample imbalance that can occur even in the face of random assignment (see Chapter 4).

With larger N and proper execution of random assignment, issues of sample imbalance usually are of lesser concern, per my discussion in Chapter 4. The traditional method for evaluating sample imbalance is to regress each covariate onto a dummy variable that represents the treatment versus control conditions and then evaluate the magnitude of the imbalance between the groups that has occurred as reflected by the regression coefficient associated with the dummy variable. For example, for the covariate COV (education) the mean for the control condition was 3.85 and for the intervention condition it was 4.17, the difference of which is 0.32. The squared R from this analysis was 0.026. These indices document the magnitude of the imbalance which then can be judged in terms of whether it likely will create problems. Some researchers perform formal statistical tests of the mean differences, but given valid random assignment to conditions, such tests usually are not recommended (see Chapter 4). Instead, the focus is on evaluating the *magnitude* of the imbalance. The quantile regression counterpart of this analysis is to compare treatment and control quantile values on the covariates across the deciles of the outcome. Here are the results for COV covariate when I used the program on my website called *Deciles and MAD* to compare the deciles for the treatment and control groups (note: if you want to estimate pseudo R squares for these analyses, you need to conduct the contrasts using the *Quantile regression* program on my website using a dummy variable; the program reports pseudo R squares):

Decile	Treatment	Control	Difference
0.1	2.94	2.52	0.42
0.2	3.31	3.06	0.25
0.3	3.61	3.33	0.28
0.4	3.88	3.57	0.31
0.5	4.18	3.84	0.34
0.6	4.45	4.12	0.33
0.7	4.70	4.35	0.35
0.8	4.97	4.65	0.32
0.9	5.47	5.11	0.36

Given my knowledge of this scale from past research, I judge that none of these differences are notable and that imbalance on COV is unlikely to be an issue (see my discussion in Chapter 4). If I decide to include COV as a covariate in the primary analyses, the observed sample imbalances in COV will be statistically adjusted for.

The current example has no missing data, so the need for covariates to address missing data mechanisms are moot. If missing data are present and the data are not missing completely at random, then I might need to include covariates to adjust for the bias that the missing data can introduce.

Finally, including baseline covariates that are relatively uncorrelated with the treatment condition (which should be the case given random assignment) and that are meaningful determinants of the outcome has the potential to increase statistical power and parameter estimation precision, just as it does in OLS regression. The degree to which this occurs depends, in part, on your sample size; if the sample size is large to begin with, as in the present case, the incremental power due to adding covariates will be smaller because you have high power to begin with. This is the case for the behavioral commitment example where the sample size is 902. Also, if the covariate is only weakly associated with the outcome, the incremental statistical power and precision will be less, everything else being equal.

As discussed in Chapter 2, many methodologists argue that the choice of covariates should not just be based on empirics but also on theory, past research, and common sense. Indeed, many feel theory and past research should take priority when making such decisions. In conditional quantile regression, an additional consideration is that the inclusion of covariates changes the targeted quantiles because they become *conditional quantiles* defined by the additional predictors and covariates in the model (Porter, 2015; see also my discussion in Chapter 6). Suppose the median is my target quantile and I include both a treatment dummy variable and COV (education) in an equation predicting

knowledge about the short term consequences of the environmental/climate crisis, i.e., M1. In this case, the coefficient for the treatment condition is not the estimated effect of the intervention on the median of M1. Rather, the coefficient is a weighted average of the effect of the intervention on the median of M1 at each of the different levels of education, i.e., it is conditional on COV. This is not the same as the effect of the treatment on the median of the overall distribution of M1 ignoring education. The parameter being estimated by the treatment variable coefficient is different with and without the education covariate in the equation. Without the covariate, you are estimating the effect of the intervention for individuals who are at the middle of the M1 distribution per se. With the covariate in the equation, this is no longer true; you are now estimating (and then averaging) the effect of the intervention for individuals who are at the middle of each of the different educational groups defined by the covariate. These individuals can be located anywhere on the unconditional M1 distribution.

Choosing covariates for the equation linking the mediators to the outcome in an RET is even more complicated because now there is no random assignment at work. The equation is purely "observational." This means that in addition to the above issues we also need to be concerned about confounds and the need to statistically control for them. In the current example, I use the baseline Y (i.e., YB) and COV to control for such confounds when evaluating the effects of the various M on Y but, in reality, my list of covariates likely would be longer. I keep the list short here for pedagogical purposes.

Finally, the *Quantile regression* program on my website reports bootstrapped standard errors, p values, and confidence intervals for sensitivity analyses to compare with their more traditional counterparts in conditional quantile regression. If the conclusions are comparable, then issues like heteroscedasticity likely are non-problematic.

In sum, the above discussion highlights key issues to think about as you approach the quantile regression analyses for RET data. Without getting too sidetracked on these issues, I now move forward to the major analyses.

Model Fit

As discussed in Chapter 8, limited information structural equation modeling (LISEM) evaluates model fit but it does so using different approaches than traditional full information SEM (FISEM). One strategy that maps onto the examination of modification indices in FISEM is to identify omitted paths in the overarching model and then test if the omission of those paths represents consequential model/specification error. Theoretically meaningful omitted paths are a major source of poor global fit values in FISEM. For the model in Figure 1, one obvious omitted path that could be influential is the direct effect

between the treatment condition and the outcome. The model in its current form assumes that all of the effects of the intervention on the outcome are captured by the intervention's effects on the three mediators, i.e., it assumes the dashed path in the reproduction of Figure 1 in Figure 3 equals zero.

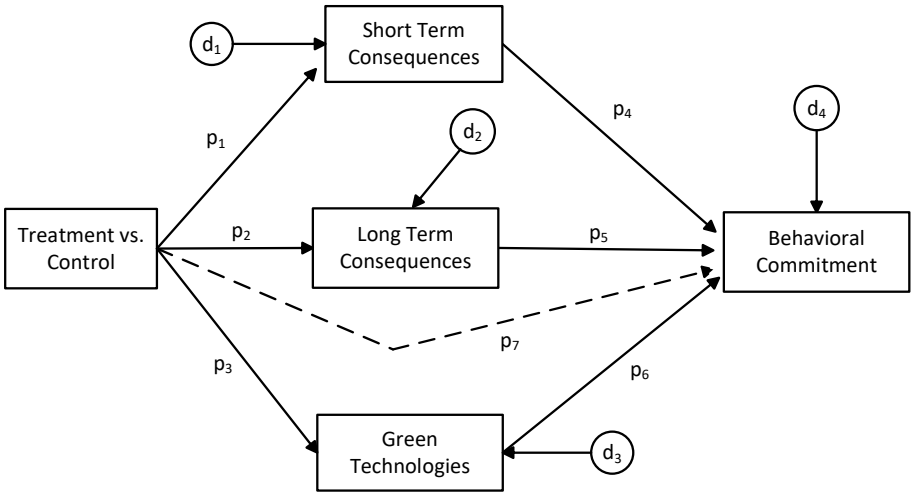


FIGURE 3. Omitted direct effect of treatment on outcome

Equation 4 for the original model in Figure 1 is as follows:

$$Y = a_4 + p_4\,M1 + p_5\,M2 + p_6\,M3 + b_7\,YB + b_8\,COV + d_4 \qquad [4]$$

The equation is obtained by regressing the endogenous variable Y onto all variables with an arrow pointing directly to it. If there is a direct effect of the treatment on Y over and above the mediators and covariates, per Figure 3, Equation 4 would be modified as follows:

$$Y = a_4 + p_4\,M1 + p_5\,M2 + p_6\,M3 + b_7\,YB + b_8\,COV + p_7\,TREAT + d_4 \qquad [4a]$$

If I execute the above equation in a quantile regression model, a statistically significant value of p_7 would be consistent with a statistically significant modification index in FISEM and would suggest a misspecified model (see Chapter 8 for elaboration).

I used the *Quantile regression* program on my website to test the coefficients in Equation 4a at the 20^{th} quantile, the 50^{th} quantile, and the 80^{th} quantile, my target quantiles. For the 20^{th} quantile, the TREAT coefficient was 0.23 (Critical Ratio (CR) =

0.12, p = 0.90), for the 50th quantile, the TREAT coefficient was 2.12 (CR = 1.55, p = 0.12), and for the 80th quantile, the TREAT coefficient was -1.21 (CR = 0.95, p = 0.34). None of these coefficients were statistically significant, which is consistent with the hypothesized model in Figure 1 that omits the p_7.

Another omitted path is a causal path from M1 to M2, per the dashed arrow in Figure 4. If this path makes no theoretical sense, then there is no need to test for its presence. If it does make theoretical sense, then I use the quantile regression program to regress M2 onto M1, M2B, COV and TREAT, i.e., all the variables with arrows pointing directly to M2 plus covariates that I think are important to control for confounds. For the 20th quantile, the M1→M2 coefficient in this regression analysis was 0.02 (CR = 0.42, p = 0.68), for the 50th quantile the M1→M2 coefficient was 0.03 (CR = 0.50, p = 0.62), and for the 80th quantile the M1→M2 coefficient was 0.05 (CR = 0.67, p = 0.50). None of these coefficients were statistically significant, which is consistent with the original hypothesized model in Figure 1.

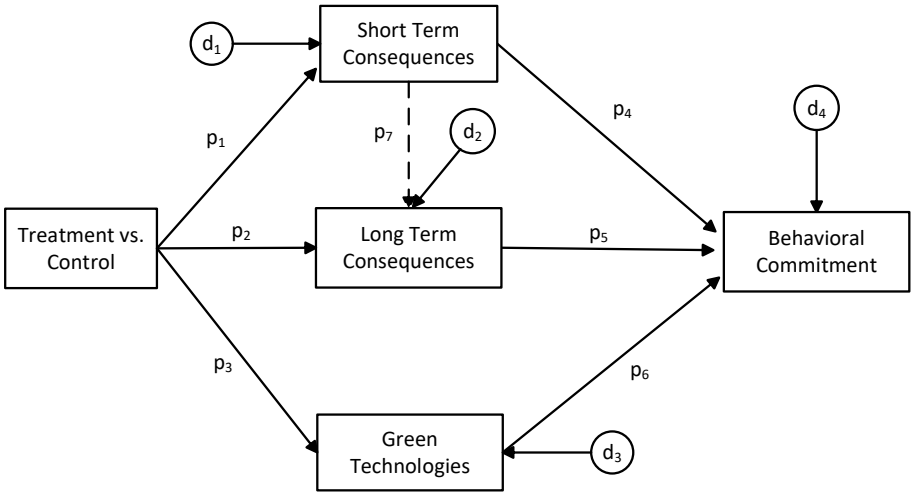


FIGURE 4. Omitted causal path from M1 to M2

Having said that, I do not think that a causal path between M1 and M2 makes theoretical sense nor does one between M1 and M3 or M2 and M3, in either direction. I simply do not see compelling logic for how one type of knowledge in this model casually influences another type of knowledge.[1] What seems more plausible instead of causal

[1] A difference between LISEM modification indices and FISEM modification indices is that the former are calculated at the discretion of the analyst because the analyst thinks they may be theoretically coherent. FISEM, by

links between the different types of knowledge is the possibility of correlated disturbances between the mediators, namely d_1 with d_2, d_1 with d_3, and/or d_2 with d_3. Such correlations would result if there are unmeasured common causes to the three types of knowledge, much like educational background (COV) is a common cause of them but which I measured and statistically controlled for. For example, one possible unmeasured common cause might be having close friends who are knowledgeable about "being green" or having close family relatives who are environmentally conscious.

In FISEM, accurately modeling these disturbances is important. Failing to do so can permeate bias to other coefficients in the model that rely on having accurate representations of the sample correlations between the mediators, such as the coefficients that are estimated in Equation 4 ($Y = a_4 + p_4\ M1 + p_5\ M2 + p_6\ M3 + b_7\ YB + b_8\ COV + d_4$). In LISEM, because it works with one equation at a time, the correlations between the mediators in Equation 4 are taken as givens and the sample correlations are directly used in the estimation process. In this sense, LISEM is less susceptible to specification error than FISEM.[2]

In sum, although LISEM based quantile regression does not yield the same type of global fit statistics that traditional FISEM does, model fit can be addressed using the adaptive modification index approach. In the current case, the model seems to provide satisfactory fit to the data.

The Quantile Regression Analysis

I seek to answer three questions in the RET, (1) what is the total effect of the intervention on the outcome of behavioral commitment to a green lifestyle, (2) what is the effect of the intervention on the three mediators that the program targets, and (3) what are the effects of the targeted mediators on the outcome of behavioral commitment to a "green" lifestyle.

For the first question I use the *Quantile regression* program on my website to regress behavioral commitment (Y) onto the treatment dummy variable, the baseline Y, and COV. I decide to focus my analysis on the 20^{th}, 50^{th} and 80^{th} quantiles and compute the results for the treatment coefficients under two scenarios. The first scenario includes covariates in the equation to improve efficiency and adjust for sampling imbalance. It analyzes conditional quantiles adjusting for the covariates. The second scenario excludes the covariates, which is justified given random assignment to the treatment conditions

contrast, calculates indices for all omitted parameters whether they are coherent or not.

[2] LISEM, however, ignores the correlations between disturbances for M1, M2 and M3 which, in turn, can lead to less efficient (but still unbiased) standard errors for them when the effect of the treatment condition on each mediator is estimated (via a dynamic known as seemingly unrelated regressions; Wooldridge, 2010).

and given the absence of missing data. It turns out that with a single predictor and no covariates, the second result converges to that for an unconditional quantile treatment effect because the distributions for the treatment and control groups can be viewed as two separate unconditional distributions representing different states of the world (see Wentz, 1997). Here are the results:[3]

Model	20th Quantile	50th Quantile	80th Quantile
Treatment with covariates	14.30* ±2.16	18.83* ±2.26	22.60* ±2.36
Treatment without covariates	14.44* ±2.20	18.72* ±2.18	23.31* ±2.40

(Notes: * $p < 0.05$; margins of error (the half widths of the 95% confidence interval) are reported after the ± sign)

The 20th quantile for the intervention group was about 14 units higher than the 20th quantile for the control group; the median for the intervention group was about 18 units higher for the intervention group than the control group; and the 80th quantile for the intervention group was about 23 units higher than the 80th quantile for the control group. Each of these differences were statistically significant, $p < 0.05$. The quantile regression program also conducted significance tests comparing the magnitude of the treatment coefficients at the different quantiles, considered pairwise. The tests revealed that the treatment coefficient at the 80th quantile was statistically significantly larger than the treatment coefficient at the 50th quantile and that the treatment coefficient at the 50th quantile was statistically significantly larger than the treatment coefficient at the 20th quantile. This was true with or without familywise error corrections. Note that if I conduct a traditional analysis comparing mean Y values for the treatment and control conditions, the mean difference is 18.67 ($p < 0.05$). This result masks the different effect magnitudes of the intervention at the lower and upper quantiles of the commitment distribution.

The second RET question asks if the intervention affects each of the three mediators that it targets for change. I conducted three separate quantile regressions using the three mediator-based equations (Equations 1 to 3) that map onto this question. I again evaluated intervention effects at the 20th, 50th and 80th quantiles. Here are the key results for the analysis that included the covariates:

[3] The video for the quantile regression program on my website walks you through output for the current example.

Dependent Variable	20th Quantile	50th Quantile	80th Quantile
M1: Short term consequences	6.85^{*} ±1.30	16.62^{*} ±1.62	25.16^{*} ±1.26
M2: Long term consequences	9.88^{*a} ±1.84	9.56^{*a} ±1.74	9.70^{*a} ±2.02
M3: Green technologies	-0.21^{a} ±1.86	-1.14^{a} ±1.52	-0.36^{a} ±2.30

(Notes: * $p < 0.05$; 95% margins of error are reported after the ± sign; within a row, coefficients with common letter superscripts are not statistically significantly different from one another)

Here are the corresponding results without the covariates, which are reasonable to examine given that I randomly assigned individuals to the treatment condition:

Dependent Variable	20th Quantile	50th Quantile	80th Quantile
M1: Short term consequences	6.84^{*} ±1.28	16.65^{*} ±1.62	25.13^{*} ±1.24
M2: Long term consequences	9.22^{*a} ±1.86	9.59^{*a} ±1.68	9.86^{*a} ±1.88
M3: Green technologies	0.23^{a} ±2.00	-1.20^{a} ±1.50	-0.33^{a} ±2.18

The intervention yielded a statistically significant effect on the first mediator, knowledge about the short term consequences of the environmental/climate crisis. This was true at the 20th quantile, the 50th quantile, and the 80th quantile. However, there were notable differences in the magnitude of the effect of the intervention relative to the M1 distribution. A person ranked at the 80th quantile in a population of people who all had been exposed to the intervention would score 25.13 units higher on M1 as compared to a person ranked at the 80th quantile in a population of control individuals who were not exposed to the intervention. By contrast, a person ranked at the 20th quantile in a population of people who all had been exposed to the intervention would score only 6.84 units higher on M1 as compared to a person ranked at the 20th quantile in a population of control individuals who were not exposed to the intervention. Framed another way, the M1 knowledge score defining the cutoff for the upper 20% of individuals (defined by the 80th quantile) is estimated to be 25.13 units higher than the score that is the cutoff for the top 20% scorers in the control group. By contrast, the cut off score to define the lower 20% of individuals in the intervention group (defined by the 20th quantile) is estimated to be 6.84 units higher than the corresponding cutoff for individuals in the control group.

Clearly, the intervention has a bigger effect on the higher quantiles than the lower quantiles for M1. The question for the program evaluator then becomes why is this the case and how can we make the intervention more effective for the lower quantiles? Perhaps the intervention is less memorable or less interesting to those in the lower

quantiles as compared to those in the higher quantiles such that something needs to be done to generate greater interest for the former. Note also that the traditional analysis of mean M1 scores would mask these differential effects

Knowledge about the long term consequences of the environmental/climate crisis (M2) also were affected by the treatment condition with the intervention tending to increase quantile or cutoff values (both conditional and otherwise) by about 9 or 10 points relative to the control group. Note that in contrast to M1, these effects tended to be uniform across the different quantiles.

Finally, for "green" technology knowledge, the intervention did not have statistically significant effects on the M3 quantiles and this was true at all the quantiles examined. Program designers obviously thought they had structured activities in their program to bring about meaningful change in M3 but those activities failed to do so. The strategies they used need to be revisited.

The third question addresses the effects of the mediators on the outcome. In this case, I need to include the covariates because there is no operative random assignment. I used the quantile regression program on my website and regressed Y onto M1, M2, M3, COV and YB for the 20th, 50th and 80th quantiles. Here are the results:

Predictor	20th Quantile	50th Quantile	80th Quantile
M1: Short term consequences	1.02*a ±0.10	1.04*a ±0.09	1.04*a ±0.08
M2: Long term consequences	0.18*a ±0.09	0.19*a ±0.08	0.18*a ±0.09
M3: Green technologies	-0.02a ±0.10	0.02a ±0.09	-0.05a ±0.10

Knowledge about the short term consequences of the environmental/climate crisis (M1) yielded statistically significant effects on behavioral commitment that were about equal in magnitude for all three (conditional) quantiles: For every one unit that M1 increases, the value of the 20th quantile is predicted to increase by about one unit; for every one unit that M1 increases, the median (50th quantile) is predicted to increase by about one unit, and for every one unit that M1 increases, the value of the 80th quantile is predicted to increase by about one unit. The same pattern is true for M2, but the magnitude of the coefficient is lower, namely about one fifth of a unit (i.e., about 0.18). Finally, M3 is not statistically significantly related to the outcome at any of the specified quantiles. Although the program designers thought this mediator would be relevant to behavioral commitments to a "green lifestyle," the data suggest the designers were wrong.

In sum, across the analyses for the three questions, we learn useful information about program dynamics, more information than we would learn from traditional OLS

regression and the analysis of means. Although I believe them to be of secondary importance, omnibus mediation effects for each mediator can be evaluated using the classic joint significance test (JST), which I discuss in depth in Chapter 10. With the JST, the mediational properties of a variable for the effect of the treatment on the outcome is said to be non-zero (i.e., statistically significant) if all of the links in its mediational chain yield statistically significant coefficients. For M1, the coefficient for the TREAT→M1 link was statistically significant for the 20th quantile as was the coefficient for the M1→Y link. M1 thus mediates some of the effect of TREAT on the 20th quantile of Y. This also is true for the 50th and 80th quantiles of Y. For M2, the coefficient for the TREAT→M2 link was statistically significant for the 20th quantile as was the coefficient for the M2→Y link. M2 thus mediates some of the effect of TREAT on the 20th quantile of Y as well as at the 50th and 80th quantiles of Y. For M3, none of the TREAT→M3 links were statistically significant nor were any of the links from M3→Y. We therefore conclude there is not sufficient support to conclude that M3 mediates some of the effect of TREAT on Y.

Here is a heuristic for interpreting coefficients in conditional quantile regression. In OLS regression we frame coefficients as how much a unit increase in X is predicted to change the mean of Y controlling for the other predictors. In quantile regression using the 50th quantile, we frame coefficients as how much a unit increase in X is predicted to change the median of Y controlling for the other predictors. For quantiles other than the 50th, substitute the phrase "change the Qth quantile" for "change the median" in the prior sentence, where Q is the value of the quantile that you target (expressed from possible values of 1 to 99). Also be explicit that in the presence of covariates, the quantiles you focus on are conditional quantiles.

Some Technical Points

In this section, I consider four technical issues that you may find helpful to consider when applying quantile regression to RETs. These include the estimation of unconditional quantile treatment effects, the use of log transforms in quantile regression, profile analysis in quantile regression, and the inclusion of baseline outcome measures.

Unconditional Quantile Treatment Effects

An important property of conditional quantile regression is its ability to evaluate differential effects of a predictor on the outcome across conditional distributions. Sometimes researchers instead want to identify such effects on the *unconditional* distribution of the outcome. This was illustrated in the prior analyses of treatment effects both with and without covariates in the prediction equation in which I took advantage of

randomization to the treatment condition. By evaluating the treatment effect with no covariates in the model, the quantile regression model essentially focused on the unconditional outcome distribution, allowing me to estimate the treatment effect for it. However, omitting covariates in randomized trials usually is not advisable if non-trivial bias creeps into randomization through mechanisms like missing data, non-compliance, and/or treatment dropouts. In such cases, we may still want to document intervention effects on the unconditional outcome distribution but now we must do so by controlling for confounds that can cause bias using a method other than conditional quantile regression. This is because conditional quantile regression shifts quantile definitions by virtue of covariate conditioning, per my discussion in Chapter 6.

A method for obtaining such estimates has been developed using an analytic framework suggested by Firpo (2007) and that has been incorporated into an R package called QTE. I make the analysis available on my webpage in the program called *Quantile treatment effects*. The approach can only be used with a binary treatment variable (e.g., intervention versus control) but the covariates can be binary, nominal, or continuous. The method proposed by Firpo (2007; see also Firpo & Pinto, 2016; Firpo, Fortin & Lemieux, 2018) uses a two-step approach. At the first step, inverse probability weights (IPWs) are used to equalize the distributions of the covariates in the two groups that define the treatment variable. These weights are defined using propensity scoring that predict group membership from the covariates. At the second step, the weights are used in weighted quantile regression to estimate the treatment effect but where the two groups now have similar distributions on the covariates thereby negating their confounding nature (Firpo 2007). When I applied this method to the numerical example to evaluate the quantile treatment effect for the intervention on the behavioral commitment to a "green" lifestyle controlling for YB and COV, the QTE for the 20^{th} quantile was 14.26 ±2.20, for the 50^{th} quantile it was 18.45 ±2.33, and for the 80^{th} quantile it was 23.32 ±2.61. These values are close to their population values for the unconditional outcome distribution.

A QTE estimation method known as the Residualized Quantile Regression approach (RQR) has been developed by Borgen, Haupt and Wiborg (2022) to estimate unconditional quantile treatment effects but for cases where the treatment or target variable has two or more values. This method is also available in the *Quantile treatment effects* program on my webpage.

In sum, if you are interested in analyzing treatment effects on distributions by comparing the outcome distributions of two different groups that share the same distribution of characteristics, then unconditional quantile treatment effects are reasonable to estimate. However, if you want to compare two groups on conditional

quantiles after adjusting for a set of covariates, the conditional treatment effect as estimated in conditional quantile regression is appropriate.

Log Transformations in Quantile Regression

You may encounter studies using quantile regression where researchers transform the outcome by calculating the natural log of it and then pursue the analysis after the transformation. The purpose of such transformations is to model the effect of a predictor on the outcome in relative terms rather than absolute terms. Consider a quantile regression model that focuses on the median of Y where Y is the hourly rate of pay for workers and it is predicted from the race of the worker (White = 1 versus non-White = 0). Suppose I find the following result:

$$\text{Wage} = 14.42 + 5.57\ D_{WHITE}$$

The intercept is the hourly wage for non-Whites and the regression coefficient for D_{WHITE} is the median difference between the hourly wage for Whites minus the hourly wage for non-Whites, i.e., 19.99 – 14.42 = 5.57. Using the program on my website called *Deciles and MADs*, here are the results for each decile:

Decile	Grp 1 Value	Grp 0 Value	Difference
0.1	14.85002	10.85044	3.999576
0.2	16.56285	11.85584	4.707012
0.3	17.71505	12.89243	4.822619
0.4	18.88349	13.71598	5.167505
0.5	19.98580	14.41711	5.568687
0.6	21.12313	15.53446	5.588675
0.7	22.26987	16.31975	5.950118
0.8	23.50127	17.11757	6.383697
0.9	25.23604	18.52013	6.715909

Note that the wage differential favors Whites at each decile but that the disparity in wages is larger in the upper end of the wage distribution than in the lower end of the distribution. For example, Whites at the 80th decile are paid about $6.40 more than Non-Whites whereas at the 20th decile, the difference is $4.70.

Another way of documenting wage disparities uses proportion or percentages. For example, I can ask by what percentage are Whites paid more than Non-Whites. If I divide the White wage at the 20th decile by the non-White wage at the 20th decile, I obtain

$$19.99/14.42 = 1.38$$

Subtracting 1.0 from this number and multiplying the result by 100 I find that Whites are

paid 38% more than non-Whites. It turns out that when analyzing the log of the outcome, we effectively analyze such relative differences rather than the absolute differences that I reported in the prior analysis. Consider the numbers X1 and X2. It turns out that

$$\log(X1) - \log(X2) = \log(X1/X2)$$

so that with log transforms, the coefficient for the dummy variable D_{WHITE} in the prior regression analysis will equal the log of the 20th quantile for Whites divided by the 20th quantile for non-Whites. If I calculate the exponent of this value to remove the log, I obtain 1.38. Here are the results for the different deciles when I analyze the log wages and calculate the exponents of the resulting coefficients:

Decile	Exp of Diff
0.1	1.368514
0.2	1.397056
0.3	1.373890
0.4	1.376802
0.5	1.386252
0.6	1.359799
0.7	1.364576
0.8	1.372847
0.9	1.362740

What is striking in these results is that the percentage by which Whites are paid more than non-Whites is roughly the same in each decile, namely about 38%. This portrays a somewhat different picture than when I found when I analyzed the absolute differences in wages. In that analysis, I found that the wage differential was greater in the higher end of the distribution than the lower end of the distribution. Both results are correct; they just come at the wage disparity from different vantage points. Parenthetically, for a continuous predictor, the exponent of the coefficient indicates the multiplicative factor by which the outcome increases given a one unit increase in the predictor, much like the multiplicative factor an odds in logistic regression.

There are cautions to keep in mind when working with log transformed outcomes. First, as I discuss in Chapter 10, the percentage value you obtain will differ depending on which group on a dummy variable is your reference group. For example, if I use Whites as the reference group (scored 0) and score Non-Whites using the value 1. Here are the results I obtain for exponent of the dummy variable coefficient:

Decile	Exp of Diff
0.1	0.7307195
0.2	0.7157908
0.3	0.7278601

```
0.4         0.7263207
0.5         0.7213697
0.6         0.7354028
0.7         0.7328286
0.8         0.7284131
0.9         0.7338158
```

In this case, the percentage by which non-Whites are underpaid is 1 minus the result, which is then multiplied by 100. For the median, it equals (1-.721) = 0.28*100 = 28%. Note that this is different from the 38% I calculated with the reverse scoring of the dummy variables. Interestingly, if I calculate the reciprocal of .721, I obtain 1/.721 = 1.38, the value I obtained before reverse scoring. Again, neither value, 38% or 28%, is correct. They are just different vantage points for examining the wage disparity.

Second, it only makes sense to work with logs if the metric of the outcome is measured on a ratio level. The ratios that the log transform works with can be misleading if the measure is interval or ordinal level. This is why the log transform approach is not appropriate for the behavioral commitment to a "green" lifestyle measure in the primary numerical example for this document.

Third, log transforms can only be made for values of 0 or more because the log of a negative number is undefined. If one or more of your scores has a negative value, you can't perform log transforms.

Profile Analysis in Quantile Regression

It is possible to perform profile analyses in quantile regression. The program on my website called *Compare discrete change* conducts profile analyses for conditional quantile regression. Using the behavioral commitment to a "green" lifestyle example, I might want to know what the predicted median Y is for people who have predictor scores of M1 = 49, M2 = 32, M3 = 42, YB = 31 and COV = 3 (i.e., all relatively low scores on the mediators) as compared with people who have predictor scores of M1 = 59, M2 = 42, M3 = 42, YB = 31, and COV = 3. Note that I constructed the second profile in a way that raises each knowledge score of M1 and M2 by 10 points relative to the first profile. For the first profile, the predicted median Y is 27.5 ±2.28 and for the second profile, the predicted median Y is 39.9 ±1.96, an increase of just over 12 points on the behavioral commitment scale. For a profile that raises M1 and M2 knowledge each by 15 points yielded a predicted median of 46.0 ±2.02 and for one that raises M1 and M2 knowledge each by 20 points yields a predicted median of 52.2 ±2.24, an increase of almost 25 points on the behavioral commitment scale relative to the first profile. How would the program designers react to these changes in M1 and M2 (ignoring M3 because of its documented unexpected irrelevance) in terms of their feasibility in a program revision

and in terms of how they translate into changes in the median of Y?

Baseline Outcomes in Quantile Regression

Many randomized trials that use quantile regression do not collect a measure of the outcome at baseline. Such trials map onto a posttest only control group design as discussed in Chapter 4. There are advantages to having a measure of the baseline outcome when conducting quantile regression analyses. We often want to know if an intervention has been more effective for people who, just prior to the initiation of the intervention, score low on the outcome as opposed to those who initially score moderately on the outcome as opposed to those who initially score high on the outcome. A strategy for doing so is to measure the outcome prior to the delivery of the intervention and again at the posttest. You then incorporate the baseline measure as a moderator of the intervention effect on the outcome using some form of moderated regression.

Some applications of quantile regression do not use randomized designs and focus on questions that render randomization impossible. For example, research in economics and sociology has studied the "motherhood penalty" to examine how having an additional child affects wage earnings. They use quantile regression to explore if the penalty is stronger at the lower end of the wage distribution as opposed to at the middle of the distribution versus at the upper end of the distribution. Sometimes designs are such that baseline outcome assessments are difficult logistically or impossible to obtain.

As noted, when we conduct conditional logistic regression, the quantiles often are such that conditionalization shifts people to various positions in the conditional quantiles. It is a mistake to characterize people in, say, the bottom 10% of the unconditional outcome distribution to be in the bottom 10% of the conditional distribution. Incorporating a baseline predictor into the analysis through moderation dynamics allows us to get a better handle on how those who are initially low scoring on the outcome respond to treatment as compared with those who are higher scoring. The third part of this book is devoted to moderator analyses and provide the tools for doing so.

Even in cases where one only has a posttest measure, some argue that quantile regression analysis can be used to infer moderation as a function of baseline assessments *in the presence of auxiliary assumptions*. One such assumption is that of **rank invariance**. Rank invariance refers to the extent to which individuals maintain their same rank in the two potential outcome distributions used to evaluate a treatment effect, i.e., the person's rank would be the same if she or he was in the control group as when he or she is in the intervention group. A variant of rank invariance is **rank similarity**, which is basically the same property as rank invariance but after controlling for one or more strategically chosen covariates, usually those that are thought to impact selection into the

treatment condition. There are tests of rank invariance and rank similarity but these are not clean (Dong & Shen, 2018; Frandsen & Lars, 2018). The tests surround evaluation of rank distribution equivalence of the intervention and control conditions after controlling for covariates that might invalidate rank invariance. In the absence of rank invariance or rank similarity, traditional interpretation of quantile regression coefficients as reflecting changes in distributions rather than specific individuals is necessitated. Given the presence of rank invariance or reasonable approximations to it, the interpretation of quantile regression and quantile treatment effects is enriched to provide perspectives on typical amounts of individual change at a given quantile - for example, the estimated motherhood penalty at the 80th quantile reflects the motherhood effect for women who would have occupied the 80th quantile in the absence of motherhood.

UNCONDITIONAL QUANTILE REGRESSION

Another form of quantile regression is **unconditional quantile regression**. In this approach, instead of estimating how a unit increase in a predictor affects the *conditional* quantile of the outcome as defined by covariates, the focus instead is on estimating how the *unconditional* outcome distribution shifts at a given quantile for a unit change in a predictor. In conditional quantile regression, the definition of quantiles is affected by the covariates in the model. This is not the case in unconditional quantile regression because the quantiles are defined only by the outcome distribution as a whole. A popular algorithm for implementing unconditional quantile regression is one suggested by Firpo et al. (2009). You can do so in the *Quantile regression* program on my website.

Coefficients associated with predictors using the Firpo et al. method are called **unconditional quantile partial effects** (UQPE). The magnitude of the coefficient is impacted by the unconditional quantile treatment effect (QTE) associated with the predictor, the composition of the predictor (also called its **share**), and the distribution of the outcome (e.g., its skewness; Borgen, Haupt & Wiborg, 2022).

For binary predictors scored 0 and 1, such as a treatment condition dummy variable, the quantile treatment effect (QTE) refers to how much a quantile value for a specified tau in the outcome distribution changes if everyone in the population has a score of 1 on the predictor versus if everyone has a score of zero on the predictor. It is analogous to comparing two predictor profiles in a profile analysis, a profile where the predictor equals 1 and a profile where the predictor equals zero. For continuous predictors, the QTE refers to how much a quantile value for a specified tau in the outcome distribution changes if everyone in the population has a score of $m+1$ on the predictor versus if everyone has a score of m.

The composition or share of a predictor is most easily understood with reference to a binary predictor. It refers to the proportion of individuals who are in one group as opposed to the other for the target predictor. In a traditional RCT or RET, the proportion of individuals in the intervention condition is often 0.50. However, some trials have a decidedly uneven split between the intervention and control conditions because they assign larger numbers of individuals to the former for strategic reasons. In economics, a study might focus on a non-randomized treatment variable, such as when a researcher seeks to evaluate the effects of being in a union on wage earnings using a large cross-sectional data base. The goal might be to use quantile regression to estimate the effects of being in a union versus not on wages at the lower end of the unconditional wage distribution, in the middle of the unconditional wage distribution, and at the upper end of the unconditional wage distribution (perhaps using taus of 0.15, 0.50, and 0.85, respectively). The proportion of workers in the population who are in a union might be 0.25, with 0.75 of workers not being union members. The share of union workers is 0.25. For continuous predictors, the composition of it is essentially the shape of its distribution, i.e., it distribution function.

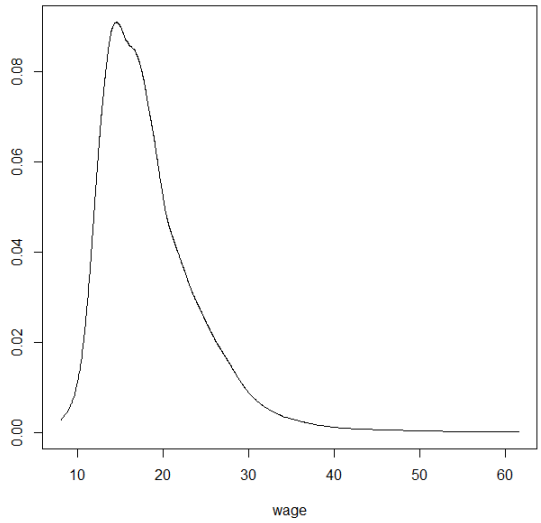
It turns out that the extent to which a change in a binary predictor produces changes in a given quantile at a specified tau in the unconditional outcome distribution is influenced by both the magnitude of the QTE for the predictor and the share of the population that has been exposed to the intervention. If the intervention is only given to half the population, thereby changing their treatment condition value from 0 to 1 while the remaining half (i.e., the control group) remains at a score of 0, then the unconditional quantile for the tau in question *for the outcome distribution* cannot change by the value of QTE. The amount of change instead will be a function of the strength of the QTE qualified by the share of the population that has been exposed to the intervention. If union membership increases by, say 10%, then the shift in the quantile for the tau in question for the wage distribution cannot equal the value of the QTE associated with union membership. It instead will be a function of the (revised) share of workers in unions as well as the QTE.

Borgen, Haupt and Wiborg (2022) show that similar dynamics apply to continuous predictors but they are more complicated than for binary predictors. And, the situation is further complicated by the shape of the outcome distribution. I do not want to get sidetracked into details of these matters here. Interested readers are referred to Borgen, Haupt & Wiborg, 2022 and Firpo et al. 2009). The important point is that the UQPE coefficient associated with a predictor reflects something different from the QTE because it takes into account the QTE, the component structure of the predictor, *and* the outcome distribution. For a continuous predictor, the UQPE coefficient estimates how much a one

unit increase in the predictor alters the quantile in question in the unconditional outcome distribution, taking these three factors into account. For a binary predictor, one can somewhat unrealistically interpret the UQPE coefficient as how much the quantile value at a given tau would change if all the individuals scored 0 on the predictor suddenly changed to a 1, but in practice, such radical changes are unlikely. For this and other technical reasons, researchers typically transform the UQPE coefficient for a binary predictor to reflect how much the quantile at a given tau changes given an increase in the "incidence rate" or share of the predictor (e.g., and increase of 0.10 or 10%). I illustrate this strategy below.

In both naturally occurring "experiments" like the aforementioned union example and in randomized trials like the behavioral commitment to a "green" lifestyle example, control variables typically need to be introduced to deal with confounds. Even randomized trials sometimes encounter the need for controls. For naturally occurring data, the unconditional outcome distribution usually is meaningful and reflects the current state of the population. In randomized trials, however, the unconditional outcome distribution can be artificial and representative of nothing. For example, half the individuals in the distribution is purposely and artificially given an intervention designed to change mediators and the outcome itself whereas the other half of individuals is explicitly denied the intervention. This produces an outcome distribution that is atypical. Often experimental trials are conducted in artificial settings that also can make the outcome distribution take on unrealistic or artificial forms. In cases where the unconditional distribution is distorted, modeling UQPEs may not make sense.

To illustrate the above concepts, I analyzed cross sectional data from a study of the effects of being a member of a union on wages. Here is the unconditional density plot of the wage distribution using annual income in units of thousands of dollars:



The distribution is positively skewed, which is not untypical of wage data. The 20th quantile equals 13.681 (or $13,681), the 50th quantile is 17.240, and the 80th quantile is 22.592.

One of my goals might be to examine the effect of being in a union on wages for "low" wage earners (the 20th quantile), for "middle" wage earners (the 50th quantile), and for "high" wage earners (the 80th quantile) while controlling for age, education and ethnicity. A second goal might be to estimate the effect of educational background (measured as highest grade completed in school) on wages at different points in the unconditional wage distribution while controlling for age, union status, and ethnicity. I do so using the Firpo et al. method in the Quantile regression program on my website. The mechanics of the method are briefly described in the Appendix.

Here are the estimated UQPE coefficients for the two target predictors from the analysis for the 20th, 50th, and 80th quantiles that included age, and ethnicity as additional predictors in addition to the two target predictors (remember, the outcome is measured in units of thousands of dollars) :

Predictor	20th Quantile	50th Quantile	80th Quantile
Union status	1.520* ±0.306	2.448* ±0.451	3.367* ±0.994
Highest grade completed	0.515* ±0.064	0.867* ±0.076	1.389* ±0.163

Union status is binary. To translate its coefficients to reflect increases in incidence rates, I multiply the coefficient in question by the increased rate I am interested in. For example, if a program increases union membership by 10% or 0.10, then the 20th quantile for the unconditional wage distribution is estimated to increase by .1520*.10 = .152 or $152, the median is estimated to increase by 2.448*.10 = .2448 or $245, and the 80th quantile is estimated to increase by 3.367*.10 = .337 = $337. If the program increases union membership by 25% or 0.25, then the 20th quantile for the unconditional wage distribution is estimated to increase by 1.520*.25 = .3800 or $380, the median is estimated to increase by 2.448*.25 = .6120 or $612, and the 80th quantile is estimated to increase by 3.367*.25 = .842 or $842.

For grade, if we increased everyone's highest grade completed by 1 grade, then the 20th quantile for the unconditional wage distribution is estimated to increase by 0.515 or $515, the median is estimated to increase by 0.867 or $867, and the 80th quantile is estimated to increase by 1.389 or $1,389.

Keep in mind that the various UQPE estimates control for the covariates (i.e., the other predictors in the equation) but they do so through marginalization rather than the

use of conditional quantiles per traditional quantile regression.

Some researchers like unconditional quantile regression because they say it is easier to characterize the policy implications of a predictor.

If you want to document unconditional quantile treatment effects to complement the above analyses, you can do so using the *Quantile treatment effects* program on my website. For union status, the estimated unconditional QTE at the 20^{th} quantile was 1.605 ±.462, at the 50^{th} quantile it was 2.348 ±.706, and at the 80^{th} quantile it was 3.009 ±.861. For grade, the estimated unconditional QTE at the 20^{th} quantile was 0.521 ±.071, at the 50^{th} quantile it was 0.884 ±.079, and at the 80^{th} quantile it was 1.225 ±.107.

CONCLUDING COMMENTS

In sum, traditional conditional quantile regression is used to study the effects of a predictor on an outcome across the conditional distribution of the outcome variable. When a covariate is introduced to control for a confound, the covariate can change an observation's place in the modeled distribution. This is analogous to implicitly creating subgroups defined by the covariates. Given this, some methodologists argue that quantile regression coefficients are local in the sense they are defined by the particular covariates that happen to be included in the model. Such effects, the argument goes, cannot be generalized to the effect of the predictor on the unconditional Y.

Unconditional quantile regression uses a different framework for analyzing predictor effects on Y across the outcome distribution. In unconditional quantile regression, the quantiles are not affected by the model covariates; the quantiles are described only as a function of the distribution of Y as a whole. As such, unconditional quantile regression reflects the effects of predictors on the unconditional distribution of the outcome, not on a conditional distribution defined by covariates. To be sure, the coefficients in unconditional quantile regression are used for statistical adjustments to control for them, but the adjustments use marginalization methods. Importantly, researchers using unconditional quantile regression must decide if the unconditional distribution studied is meaningful for their research questions. Unconditional distributions in artificial randomized trials may not lend themselves to meaningful unconditional quantile regression analysis.

APPENDIX: UNCONDITIONAL QUANTILE REGRESSION ALGORITHM

A popular algorithm for unconditional quantile regression is that proposed by Firpo et al. (2009). The algorithm transforms the outcome variable using a recentered influence function (RIF). The transformation for a given score on the outcome, y, and a quantile τ (e.g., τ = 0.50 for the median) is

$$\text{RIF}(y) = q_\tau + \frac{\tau - 1\{y \leq q_\tau\}}{f_Y(q_\tau)}$$

where $1\{y \leq q_\tau\}$ is a function that equals 1 when an observation's value of y is less than or equal to the value of the outcome at quantile τ, 0 otherwise, and $f_Y(q_\tau)$ is the density of y at quantile τ. This transformation results in only two values for the outcome depending on whether an observation is above or below the specified quantile. Once the transformation has been made, it is used as an outcome in an OLS regression with the predictors consisting of those of interest to the researcher. For a completed worked example, see Porter (2015) and for statistical details, see Firpo et al., (2009).

REFERENCES

Borgen, N., Haupt, A. & Wiborg, 0. (2022). A new framework for estimation of unconditional quantile treatment effects: The residualized quantile regression (RQR) model. SocArXiv, doi: 10.31235/osf.io/42gcb.

Dong, Y. & Shen, S. (2018). Testing for rank invariance or similarity in program evaluation, *Review of Economics and Statistics*, 100, 78–85.

Firpo, S. (2007). Efficient semiparametric estimation of quantile treatment effects. *Econometrica*, 75, 259–276.

Firpo, S., Fortin, N. & Lemieux, T. (2009). Unconditional quantile regressions. *Econometrica*, 77, 953-973.

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

Firpo, S. & Pinto, C. (2016). Identification and estimation of distributional impacts of interventions using changes in inequality measures. *Journal of Applied Econometrics*, 31, 457-486.

Frandsen, B. & Lars, J. (2018). Testing rank similarity. *Review of Economics and Statistics*, 100, 86-91.

Hao, L. & Naiman, D. (2007). *Quantile regression*. SAGE.

Wentz, S. (2018). What quantile regression does and doesn't do: A commentary on Petscher and Logan (2014). *Child Development*, 90, 1442-1452.