A Note on the Robustness of Quantile Treatment Effect Estimands

Karim Chalak‡
University of Virginia
August 9, 2019

Abstract

This note examines the robustness of two quantile treatment effect estimands to a perturbation away from the common effect assumption. The first estimand $Q_{Y_1 - Y_0}(\tau)$ is the $\tau$-quantile of the difference between the potential outcomes and the second estimand $Q_{Y_1}(\tau) - Q_{Y_0}(\tau)$ is the difference between the $\tau$-quantiles of the potential outcomes. To this end, this note provides a simple “trembling hand” example whereby the treatment effect deviates from the common effect $\beta$ for only one individual in a large population. As a result, each estimand deviates from $\beta$, and may have the opposite sign than $\beta$, in a distinct range of $\tau$. In general, this perturbation leads $Q_{Y_1 - Y_0}(\cdot)$ to differ from $\beta$ more severely but only in a small and extreme range of $\tau$ whereas it leads $Q_{Y_1}(\cdot) - Q_{Y_0}(\cdot)$ to differ from $\beta$ less severely but over a substantial and central range of $\tau$. This example suggests that researchers should carefully evaluate estimates or bounds for $Q_{Y_1 - Y_0}(\tau)$ and especially $Q_{Y_1}(\tau) - Q_{Y_0}(\tau)$ over a large range of $\tau$.

Keywords: heterogeneity, quantile treatment effect, robustness.

JEL Classification Codes: C18, C21.

1 Introduction

Quantile regression is a powerful tool to estimate aspects of the conditional distribution of a variable that may not be suitably captured by its conditional mean (see Koenker and Bassett, 1978). For instance, quantile regression analysis can help document the heterogenous impact that a randomized binary treatment $T$ may have on an outcome $Y$. Let $Y_1$ and $Y_0$ denote the potential outcomes under the treatment and control settings respectively. As discussed in

---

*Department of Economics, University of Virginia, Charlottesville, VA, USA. Email: chalak@virginia.edu.
†Acknowledgments: I thank John Pepper for helpful feedback. Any errors are the responsibility of the author.
e.g., Abadie and Cattaneo (2018), one useful measure of the impact of the treatment is the difference between the \( \tau \)-quantiles of the potential outcomes \( Q_{Y_1}(\tau) - Q_{Y_0}(\tau) \). This estimand is commonly referred to as the quantile treatment effect. An alternative useful measure is the \( \tau \)-quantile of the difference between the potential outcomes \( Q_{Y_1 - Y_0}(\tau) \). Whereas the expectation operator is linear, \( Q_{Y_1}(\tau) - Q_{Y_0}(\tau) \) and \( Q_{Y_1 - Y_0}(\tau) \) need not coincide. Roughly speaking, for continuous potential outcomes, these estimands coincide when each subject has the same rank in the distribution of \( Y_0 \) and \( Y_1 \) (see, e.g., Doksum (1974) and Lehmann (1974)).

The literature discusses various motivations and advantages for each of these estimands. First, as Angrist and Pischke (2009, p. 281) explain, \( Q_{Y_1}(\cdot) - Q_{Y_0}(\cdot) \) is informative about the treatment effect “on distributions, not on individuals.” In contrast, \( Q_{Y_1 - Y_0}(\cdot) \) is informative about the “quantiles of the distribution of the treatment effect” (Abadie and Cattaneo, 2018). For instance, \( Q_{Y_1}(\frac{1}{2}) - Q_{Y_0}(\frac{1}{2}) \) contrasts the medians of the potential outcome distributions \( F_{Y_1} \) and \( F_{Y_0} \) under the treatment and control settings whereas \( Q_{Y_1 - Y_0}(\frac{1}{2}) \) concerns the median of the distribution \( F_{Y_1 - Y_0} \) of the treatment effect \( Y_1 - Y_0 \). Second, these estimands may measure different aspects in welfare analysis. For example, Angrist and Pischke (2009, p. 285) discuss that “[\( Q_{Y_1}(\cdot) - Q_{Y_0}(\cdot) \)] are usually more important than \( Q_{Y_1 - Y_0}(\cdot) \) because comparison of social welfare typically requires only the distributions of \( Y_1 \) and \( Y_0 \).” Furthermore, Heckman, Smith, and Clements (1997, p. 491) study estimands, such as \( Q_{Y_1 - Y_0}(\tau) \), that concern the “distribution of impacts” and “are of interest to persons interested in ‘social justice’” in their evaluation of the benefits of a program. Finally, quantile regressions or instrumental variable quantile regressions can be used to point identify and estimate \( Q_{Y_1}(\tau) - Q_{Y_0}(\tau) \) when \( T \) is (conditionally) randomized or when an exogenous instrument is available (see e.g. Abadie, Angrist, and Imbens (2002) and Firpo (2007)). In contrast, \( Q_{Y_1 - Y_0}(\tau) \) is generally partially identified when \( T \) is randomized (see e.g. Heckman, Smith, and Clements (1997) and Fan and Park (2010)). Nevertheless, \( Q_{Y_1 - Y_0}(\tau) \), or quantities similar to it, may be point identified under further structure, such as rank invariance or similarity (e.g. Chernozhukov and Hansen, 2005) or in the linear random coefficient model (e.g. Hoderlein, Klemela, and Mammen, 2010). See also Hoderlein and Mammen (2007) and Sasaki (2015) who study the interpretation of the quantile regression estimand in nonparametric nonseparable structural systems.

In interpreting these estimands, it is useful to examine how robust they are to small deviations from benchmark assumptions. Indeed, quantile regression was put forward as a robust
alternative to the least squares estimator for the linear model (see Koenker and Bassett, 1978). In particular, as Heckman, Smith, and Clements (1997, p. 488) discuss, one important special case in which \( Q_{Y_1}(\cdot) - Q_{Y_0}(\cdot) \) and \( Q_{Y_1-Y_0}(\cdot) \) coincide is the "common effect' model where the programme is assumed to have the same impact on everyone" since in this simple yet central case, \( Y_1 - Y_0 \) is degenerate. This note studies the robustness of \( Q_{Y_1}(\cdot) - Q_{Y_0}(\cdot) \) and \( Q_{Y_1-Y_0}(\cdot) \) to a perturbation away from the common effect assumption. Here, "robustness signifies insensitivity to small deviations from the assumptions" (Huber, 1981, p. 1). To this end, the note provides a "trembling hand" example in arguably one of the simplest settings. In particular, it considers a large population of individuals who respond to the treatment identically. It then examines the consequences on \( Q_{Y_1}(\cdot) - Q_{Y_0}(\cdot) \) and \( Q_{Y_1-Y_0}(\cdot) \) of deviating from the common effect assumption for only one individual in the population.

2 A "Trembling Hand" Example

2.1 The Basic Setup

Consider a large finite population of individuals indexed by \( i = 1, 2, \ldots, n \). Suppose that the individuals in this population are exposed to a binary treatment \( T \). For example, \( T(i) \) denotes whether individual \( i \) receives a vaccine or a placebo. Let \( Y_1 \) and \( Y_0 \) denote the potential outcomes under the treatment and control settings respectively. For example, \( Y_t(i) \) may denote the health outcome for individual \( i \) under the treatment setting \((t = 1)\) or the control setting \((t = 0)\) (ignore any network or social interaction effects). Individual \( i \)'s outcome is observed in only one treatment setting and is given by \( Y(i) = [Y_1(i) - Y_0(i)]T(i) + Y_0(i) \). Let \( F_{Y_0} \) denote the discrete distribution of the potential outcome \( Y_0 \) under the control setting, with realizations \( Y_0(1), \ldots, Y_0(n) \). Further, denote the ordered realizations of \( Y_0 \) by \( Y_0[1], \ldots, Y_0[n] \) and adopt the shorthand notation

\[ Y_0[a] = d_a \quad \text{for } d_a \in \mathbb{R} \text{ and } a = 1, \ldots, n, \]

with \(-\infty \equiv d_0 < d_1 \leq d_2 \leq \ldots \leq d_n < d_{n+1} \equiv +\infty\). Note that, unlike \( Y_0(i) \), the ordered potential outcome value \( Y_0[i] \) need not correspond to individual \( i \). Similarly, let \( F_{Y_1} \) denote the distribution of the potential outcome \( Y_1 \) under the treatment setting and let \( Y_1[1], \ldots, Y_1[n] \) denote the ordered realizations of \( Y_1 \), arranged (weakly) from the smallest to the largest value. Last, suppose that the treatment effect is the constant \( \beta \in \mathbb{R} \) for every individual in the
Figure 1: Figure 1 illustrates the trembling hand example. The upper panel illustrates $Q_{Y_1 - Y_0} (\tau)$ by connecting the potential outcome realizations $Y_0(i)$ to $Y_1(i)$ for $i = 1, \ldots, n$ using a solid line. The lower panel illustrates $Q_{Y_1} (\tau) - Q_{Y_0} (\tau)$ by connecting the ordered potential outcome realizations $Y_0[a]$ to $Y_1[a]$ for $a = 1, \ldots, n$ using a dashed line.

population, except for individual $j$. In particular, for $\beta^* \in \mathbb{R}$, let

$$
Y_1(i) - Y_0(i) = \begin{cases} 
\beta & \text{for } i \in \{1, \ldots, n\}\setminus\{j\} \\
\beta^* & \text{for } i = j 
\end{cases}.
$$

(1)

Thus, for $\beta^* \neq \beta$, $\Pr(Y_1 - Y_0 = \beta^*) = \frac{1}{n} \equiv \epsilon$ and $\Pr(Y_1 - Y_0 = \beta) = \frac{n-1}{n} \equiv 1 - \epsilon$. For instance, everyone may benefit equally from taking the vaccine except for individual $j$ who experiences an adverse reaction.

Figure 1 illustrates the setup and the subsequent discussion. For simplicity, it considers a population of $n = 10$ individuals. The upper panel connects, using a solid line, the potential outcome realizations $Y_0(i)$ and $Y_1(i)$, $i = 1, \ldots, n$. The lower panel connects, using a dashed line, the ordered potential outcome realizations $Y_0[a]$ and $Y_1[a]$ for $a = 1, \ldots, n$. For instance, individual $h$ ranks second under both the treatment and control settings, $Y_0(h) = Y_0[2] = d_2$ and $Y_1(h) = Y_1[2] = d_2 + \beta$ whereas individual $j$ occupies position $b = 8$ under the control setting and position $c = 3$ under the treatment setting, $Y_0(j) = Y_0[b] = d_8$ and $Y_1(j) = Y_1[c] = d_8 + \beta^*$. 
2.2 The Quantile of the difference between the Potential Outcomes

First, examine $Q_{Y_1-Y_0}(\cdot)$ in this example. We have from (1) that, for $0 < \tau < 1$, 

$$Q_{Y_1-Y_0}(\tau) = \inf\{b : F_{Y_1-Y_0}(b) \geq \tau\} = \begin{cases} \min\{\beta, \beta^*\} & \text{if } \tau \leq \epsilon \\ \beta & \text{if } \epsilon < \tau \leq 1 - \epsilon \\ \max\{\beta, \beta^*\} & \text{if } 1 - \epsilon < \tau \end{cases}.$$ 

Thus, except when either $\tau \leq \epsilon$ or $1 - \epsilon < \tau$, we have that $Q_{Y_1-Y_0}(\tau)$ coincides with $\beta$. Otherwise, $Q_{Y_1-Y_0}(\tau) = \beta^*$ in one extreme $\epsilon$-range of $\tau$.

2.3 The difference between the Quantiles of the Potential Outcomes

Next, examine $Q_{Y_1}(\cdot) - Q_{Y_0}(\cdot)$. First, note that, for $t = 0, 1$ and $0 < \tau < 1$, 

$$Q_{Y_t}(\tau) = \inf\{y : F_{Y_t}(y) \geq \tau\} = Y_t[a(\tau)] \quad \text{where } a(\tau) = \min\{a \in \{1, \ldots, n\} : \tau \leq \frac{a}{n}\}. \quad (2)$$

Observe that if $\beta^* = \beta$ then

$$Q_{Y_1}[a] = d_a + \beta \quad \text{for } a = 1, \ldots, n.$$ 

Further, if $Y_0(j) = d_b$ and $\beta + d_{b-1} - d_b \leq \beta^* \leq \beta + d_{b+1} - d_b$ then the rank of an individual’s potential outcome under the control setting is preserved under the treatment setting, i.e., for any $i$ and $a$ if $Y_0(i) = Y_0[a]$ then $Y_1(i) = Y_1[a]$. More generally, let individual $j$’s potential outcomes under the control and treatment settings correspond to ranks $b$ and $c$ respectively, i.e., $Y_0(j) = Y_0[b]$ and $Y_1(j) = Y_1[c]$. As illustrated in Figure 1 for $c = 3$ and $b = 8$, this holds\(^1\) when\(^2\)

$$\beta^* \in [d_{c-1} - d_b + 1_{(b<c)}(d_c - d_{c-1}) + \beta, d_{c+1} - d_b - 1_{(c<b)}(d_{c+1} - d_c) + \beta]. \quad (3)$$

Then, using $Q_{Y_1}(\tau) = Y_t[a(\tau)]$ from (2), we obtain that, for $0 < \tau < 1$, (recall that $\epsilon \equiv \frac{1}{n}$)

$$Q_{Y_1}(\tau) - Q_{Y_0}(\tau) = \begin{cases} \beta & \text{if } \tau \leq \min\left\{\frac{b}{n}, \frac{c}{n}\right\} - \epsilon \\ \beta + d_{a(\tau)+1} - d_{a(\tau)} & \text{if } \frac{b}{n} - \epsilon < \tau \leq \frac{c}{n} - \epsilon \\ \beta^* + d_b - d_c & \text{if } \frac{c}{n} - \epsilon < \tau \leq \frac{\epsilon}{n} \\ \beta + d_{a(\tau)-1} - d_{a(\tau)} & \text{if } \frac{\epsilon}{n} < \tau \leq \frac{b}{n} \\ \beta & \text{if } \max\left\{\frac{b}{n}, \frac{c}{n}\right\} < \tau \end{cases}.$$ 

\(^1\)This follows since $Y_1[c] = Y_1(j) = Y_0(j) + \beta^* = d_b + \beta^*$ immediately succeeds $d_c + \beta$ (the value $Y_1[c]$ would take had $\beta^* = \beta$) when $b < c$ and immediately precedes $d_c + \beta$ when $c < b$.

\(^2\)\(1_{(A)}(B) = B\) if $A$ is true and is 0 otherwise.
Figure 2: Figure 2 plots the two quantiles estimands corresponding to the example in Figure 1 as $\tau$ ranges from 0 to 1. $Q_{Y_1 - Y_0}(\cdot)$ is depicted using a dark solid line and $Q_{Y_1}(\cdot) - Q_{Y_0}(\cdot)$ is depicted using a light dashed line.

In this example, $Q_{Y_1}(\tau) - Q_{Y_0}(\tau)$ generally differs from the common effect $\beta$ when $\min\{\frac{b}{n}, \frac{c}{n}\} - \epsilon < \tau \leq \max\{\frac{b}{n}, \frac{c}{n}\}$ and coincides with $\beta$ otherwise. In the limit, if $b = c$ then the rank of any individual’s potential outcome under the control setting is preserved under the treatment setting, and we obtain

$$Q_{Y_1}(\tau) - Q_{Y_0}(\tau) = \begin{cases} 
\beta & \text{if } \tau \leq \frac{b}{n} - \epsilon \\
\beta^* & \frac{b}{n} - \epsilon < \tau \leq \frac{b}{n} \\
\beta & \frac{b}{n} < \tau
\end{cases}.$$ 

Here, except for $\frac{b}{n} - \epsilon < \tau \leq \frac{b}{n}$, $Q_{Y_1}(\tau) - Q_{Y_0}(\tau)$ coincides with $\beta$.

2.4 Discussion

Figure 2 demonstrates how the perturbation in Figure 1 affects each of the quantile processes $Q_{Y_1 - Y_0}(\cdot)$ and $Q_{Y_1}(\cdot) - Q_{Y_0}(\cdot)$ in a distinct range of $\tau$. It plots $Q_{Y_1 - Y_0}(\cdot)$ using a dark solid line and $Q_{Y_1}(\cdot) - Q_{Y_0}(\cdot)$ using a light dashed line. As illustrated in Figure 2, the perturbation away from the common effect assumption leads the quantile process $Q_{Y_1 - Y_0}(\cdot)$ to differ from $\beta$ only for extreme values of $\tau$. The sign and magnitude of the effect perturbation $\beta^* - \beta$ determine the direction and severity of the discrepancy over this small $\epsilon$-range of $\tau$. In contrast, this
perturbation leads $Q_Y(\cdot) - Q_{\dot{Y}}(\cdot)$ to differ from $\beta$ over a possibly substantial and central range of $\tau$. The larger the deviation $|c - b|$ in the ranks of individual $j$’s potential outcomes under the treatment and the control settings is, the wider the range of $\tau$ where $Q_Y(\cdot) - Q_{\dot{Y}}(\cdot)$ may differ from $\beta$. The sign of $c - b$ determines the sign of the discrepancy over this range. Further, the magnitude of the discrepancy increases with either $d_{a(\tau)} - d_{a(\tau)-1}$ or $d_{a(\tau)+1} - d_{a(\tau)}$, the distance between the neighbor values of the potential outcome$^3 Y_{\dot{y}}$. In particular, the sign of $Q_Y(\cdot) - Q_{\dot{Y}}(\cdot)$ may differ from that of $\beta$ over a central range of $\tau$. For example, even when every $i \in \{1, ..., n\}\{j\}$ benefits from the vaccine by $\beta \in (0, \min_{\tau \in (\frac{c}{n}, \frac{b}{n})}(d_{a(\tau)} - d_{a(\tau)-1}))$, if $j$ experiences a strong adverse reaction $\beta^* < 0$ then $Q_Y(\tau) - Q_{\dot{Y}}(\tau) = \beta + d_{a(\tau)-1} - d_{a(\tau)} < 0$ for all $\tau \in (\frac{c}{n}, \frac{b}{n}]$. As illustrated in Figures 1 and 2, this range can be wide if $\beta^*$ corresponds to a substantial drop $c - b$ in the rank $c$ of individual $j$’s potential outcome under the treatment setting when compared to the rank $b$ of $j$’s potential outcome under the control setting. An analyst who bases his or her evaluation on $Q_Y(\tau) - Q_{\dot{Y}}(\tau)$ for $\tau \in (\frac{c}{n}, \frac{b}{n}]$ (e.g. on the “median treatment effect” when $\frac{c}{n} < \tau = \frac{1}{2} \leq \frac{b}{n}$) may conclude in this case that the vaccine is harmful.

In sum, this perturbation away from the comment effect $\beta$ generally leads $Q_{Y_1-Y_0}(\cdot)$ to differ from $\beta$ more severely but only in a small and extreme range of $\tau$ whereas it generally leads $Q_Y(\cdot) - Q_{\dot{Y}}(\cdot)$ to differ from $\beta$ less severely but over a substantial and central range of $\tau$.

Finally, it is worth emphasizing that robustness is only one of several favorable properties to consider when interpreting an estimand. Moreover, even according to this criterion, it may be possible that an estimand is robust to certain types of perturbations from a benchmark assumption but not to others. In particular, the results in this note pertain to the common effect benchmark assumption where the treatment effect $\beta$ is homogenous. In this case, $Q_{Y_1-Y_0}(\cdot)$ and $Q_Y(\cdot) - Q_{\dot{Y}}(\cdot)$ coincide with $\beta$. Moreover, this setting facilitates formalizing a useful notion of a perturbation away from $\beta$ and examining its consequences on these estimands. More generally, $Y_1(i) - Y_0(i) = \beta(i)$ may vary with $i$. In this case, $Q_{Y_1-Y_0}(\tau)$ may differ from $Q_Y(\tau) - Q_{\dot{Y}}(\tau)$ and the consequences of various perturbations of $\beta(\cdot)$ on $Q_{Y_1-Y_0}(\cdot)$ and $Q_Y(\cdot) - Q_{\dot{Y}}(\cdot)$ are more nuanced than in the common effect case. Future work may study extensions of the above analysis to such more general settings.

---

$^3$If $\frac{c}{n} - \epsilon < \tau \leq \frac{c}{n}$ then condition (3) gives that $[Q_Y(\tau) - Q_{\dot{Y}}(\tau)] - \beta \in [1_{(b < c)}(d_c - d_{c-1}) - (d_c - d_{c-1})).(d_{c+1} - d_c) - 1_{(c < b)}(d_{c+1} - d_c)]$. Thus, the magnitude of the discrepancy is bounded by either $d_c - d_{c-1}$ or $d_{c+1} - d_c$. 

---

7
This note examines the robustness of the quantile treatment estimands $Q_{Y_1-Y_0}(\tau)$ and $Q_{Y_1}(\tau) - Q_{Y_0}(\tau)$ to a perturbation away from the common effect assumption whereby the treatment effect is homogenous except for one individual in a large population. In general, this perturbation affects $Q_{Y_1-Y_0}(\cdot)$ more severally but in a small and extreme range of $\tau$ whereas it affects $Q_{Y_1}(\cdot) - Q_{Y_0}(\cdot)$ less severally but over a substantial and central range of $\tau$. This suggests that researchers should carefully evaluate estimates or bounds for $Q_{Y_1-Y_0}(\tau)$ and especially $Q_{Y_1}(\tau) - Q_{Y_0}(\tau)$ over a large range of $\tau$. This analysis calls for a more detailed study of the robustness of these two useful quantile treatment effect estimands in more general contexts.

References:


