Supplementary Material

Supplement to “Tolerating defiance? Local average treatment effects without monotonicity”  
(Quantitative Economics, Vol. 8, No. 2, July 2017, 367–396)

CLÉMENT DE CHAISEMARTIN  
Department of Economics, University of California at Santa Barbara

In this supplement, I show that one can estimate quantile treatment effects in the surviving-compliers subpopulation, that one can test the CD condition, and that my results extend to multivariate treatment and instrument, to fuzzy regression discontinuity designs, and to survey nonresponse.

S1. Testability and quantile treatment effects

In this section, I show that random instrument, exclusion restriction, and CD have a testable implication. Then I discuss how the testable implications of random instrument, exclusion restriction, and ND studied in Huber and Mellace (2012), Kitagawa (2015), and Mourifie and Wan (forthcoming) relate to the CD condition.

For any random variable $R$, let $S(R)$ denote the support of $R$. Let $U_{dz}$ be a random variable uniformly distributed on $[0, 1]$ denoting the rank of an observation in the distribution of $Y|D = d, Z = z$. If the distribution of $Y|D = d, Z = z$ is continuous, $U_{dz} = F_{Y|D = d, Z = z}(Y)$. If the distribution of $Y|D = d, Z = z$ is discrete, one can randomly allocate a rank to tied observations by setting $U_{dz} = F_{Y|D = d, Z = z}(Y) - F_{Y|D = d, Z = z}(Y^-) + V(F_{Y|D = d, Z = z}(Y) - F_{Y|D = d, Z = z}(Y^-))$, where $V$ is uniformly distributed on $[0, 1]$ and independent of $(Y, D, Z)$, and $Y^- = \sup\{y \in S(Y|D = d, Z = z) : y < Y\}$.

For every $d \in \{0, 1\}$, let $p_d = \frac{FS}{p(D = d|Z = d)}$. Notice that both $p_0$ and $p_1$ are included between 0 and 1. Finally, let

$$L = E(Y|D = 1, Z = 1, U_{11} \leq p_1) - E(Y|D = 0, Z = 0, U_{00} \geq 1 - p_0),$$

$$\bar{L} = E(Y|D = 1, Z = 1, U_{11} \geq 1 - p_1) - E(Y|D = 0, Z = 0, U_{00} \leq p_0).$$

**Theorem S1.** If Assumptions 1, 2, and 5 are satisfied,

$$L \leq W \leq \bar{L}. \tag{29}$$

The intuition of this result is as follows. Under random instrument, exclusion restriction, and CD, $E(Y_1 - Y_0|C_Y)$ is point identified: following Theorem 2.1 in the main paper,
it is equal to \( W \). It is also partially identified. The term \( C_V \) is included in \( \{D = 0, Z = 0\} \), and it accounts for \( p_0\% \) of this population. Therefore, \( E(\theta|C_V) \) cannot be larger than the mean of \( Y_0 \) of the \( p_0\% \) of this population with the largest \( Y_0 \). It also cannot be smaller than the mean of \( Y_0 \) of the \( p_0\% \) with the lowest \( Y_0 \) (see Horowitz and Manski (1995) and Lee (2009)). Combining this with a similar reasoning for \( Z \) and \( \varepsilon \) to construct a uniformly valid confidence interval for \( \varepsilon \).\footnote{See Huber and Mellace (2015), Kitagawa (2015), Machado, Shaikh, and Vytlacil (2013), and Mourifie and Wan (forthcoming) have developed statistical tests of these implications. I now discuss how these tests relate to the CD condition.}

As pointed out in Balke and Pearl (1997) and Heckman and Vytlacil (2005), random instrument, exclusion restriction, and ND also have testable implications. Huber and Mellace (2015), Kitagawa (2015), Machado, Shaikh, and Vytlacil (2013), and Mourifie and Wan (forthcoming) have developed statistical tests of these implications. I now discuss how these tests relate to the CD condition.

The test suggested in Huber and Mellace (2015) is not a test of CD. Huber and Mellace (2015) use the fact that \( E(\theta|NT) \) and \( E(\theta|AT) \) are both point and partially identified under random instrument, exclusion restriction, and ND. For instance, \( E(Y|D = 0, Z = 1) \) is equal to \( E(Y_0|NT) \) and \( E(Y|D = 0, Z = 0) \) is equal to \( (1 - p_0)E(Y_0|NT) + p_0E(Y_0|C) \). This implies

\[
E(Y|D = 0, Z = 0, U_{00} \leq 1 - p_0) \leq E(Y|D = 0, Z = 1) \\
\leq E(Y|D = 0, Z = 0, U_{00} \geq p_0).
\]

But under CD, \( E(Y|D = 0, Z = 1) \) is equal to \( \frac{p_{NT}}{p_{NT} + p_{F}} E(Y_0|NT) + \frac{p_{F}}{p_{NT} + p_{F}} E(Y_0|F) \), so the previous implication need no longer be true.
The tests suggested in Kitagawa (2015), Machado, Shaikh, and Vytacil (2013), and Mourifie and Wan (forthcoming) are not tests of CD, but they are tests of the CDM condition introduced hereafter.

**Assumption S1 (Compliers–Defiers for Marginals (CDM)).** There is a subpopulation of C denoted \( CF \) that satisfies

\[
\begin{align*}
P(C_F) &= P(F), \\
Y_0|C_F &\sim Y_0|F, \\
Y_1|C_F &\sim Y_1|F.
\end{align*}
\]

The CDM condition requires that a subgroup of compliers have the same size and the same marginal distributions of \( Y_0 \) and \( Y_1 \) as defiers. CDM is stronger than CD. On the other hand, it is invariant to the scaling of the outcome, while CD is not. With nonbinary outcomes, working under assumptions invariant to scaling might be desirable so one might then prefer to invoke CDM than CD, despite the fact it is a stronger assumption. Following the logic of Theorem 2.2 in the main paper, a sufficient condition for CDM to hold is that there be more compliers than defiers in each subgroup with the same value of \((Y_0/Y_1)\):

\[
P(F|Y_0, Y_1) \leq P(C|Y_0, Y_1).
\]

One can show that under CDM, the marginal distributions of \( Y_0 \) and \( Y_1 \) are identified for surviving-compliers. The estimands are the same as in Imbens and Rubin (1997):

\[
\begin{align*}
f_{Y_0|C_F}(y_0) &= \frac{P(D = 0|Z = 0)f_{Y|D=0,Z=0}(y_0) - P(D = 0|Z = 1)f_{Y|D=0,Z=1}(y_0)}{P(D = 0|Z = 0) - P(D = 0|Z = 1)}, \\
f_{Y_1|C_F}(y_1) &= \frac{P(D = 1|Z = 1)f_{Y|D=1,Z=1}(y_1) - P(D = 1|Z = 0)f_{Y|D=1,Z=0}(y_1)}{P(D = 1|Z = 1) - P(D = 1|Z = 0)}.
\end{align*}
\]

The testing procedures in Kitagawa (2015) and Mourifie and Wan (forthcoming) test whether the right-hand sides of the two previous equations are positive everywhere. Under CDM, these quantities are densities so they must be positive. These procedures are therefore tests of CDM. Machado, Shaikh, and Vytacil (2013) focus on binary outcomes. Their procedure tests inequalities equivalent to those considered in Kitagawa (2015). Therefore, their procedure is also a test of the CDM condition.

**S2. Multivariate treatment and instrument**

Results presented in the main paper extend to applications where treatment is multivariate. Assume that for every \( z \in \{0; 1\} \), \( D_z \in \{0, 1, 2, \ldots, J\} \) for some integer \( J \). Let \( (Y_j)_{j=0,1,2,\ldots,J} \) denote the corresponding potential outcomes. Let \( C^j = \{D_1 \geq j > D_0\} \) denote the subpopulation of compliers induced to go from a treatment below to above \( j \) by the instrument. Let \( F^j = \{D_0 \geq j > D_1\} \) denote the subpopulation of defiers induced
to go from above to below \( j \) by the instrument. As in Angrist and Imbens (1995), I assume that the c.d.f. of \( D|Z = 0 \) stochastically dominates that of \( D|Z = 1 \). This implies that \( P(C_j) \geq P(F_j) \). Consider the following CD condition.

**Assumption S2** (Compliers–Defiers for Multivariate Treatment (CDMU)). For every \( j \in \{1, 2, \ldots, J\} \), there is a subpopulation of \( C_j \), denoted \( C^j_F \), that satisfies

\[
P(C^j_F) = P(F^j),
\]

\[
E(Y_j - Y_{j-1}|C^j_F) = E(Y_j - Y_{j-1}|F^j).
\]

If CDMU is satisfied, one can show that \( W \) is equal to the average causal response,

\[
\sum_{j=1}^{J} w_j E(Y_j - Y_{j-1}|C^j_F),
\]

where \( C^j_F \) is a subset of \( C_j \) of size \( P(C_j) - P(F_j) \), and \( w_j \) are positive real numbers. This generalizes Theorem 1 in Angrist and Imbens (1995).

Their Theorem 2 considers the case where the instrument is multivariate: \( Z \in \{0, 1, 2, \ldots, K\} \). Let \( \beta \) denote the coefficient of \( D \) in a 2SLS regression of \( Y \) on \( D \) using the set of dummies \( \{1[Z = k]\}_{1 \leq k \leq K} \) as instruments. Angrist and Imbens (1995) show that if \( D_z' \geq D_z \) for every \( z' \geq z \), \( \beta \) is equal to a weighted average of average causal responses. Let \( C^{jz} = \{D_z \geq j > D_{z-1}\} \) (resp. \( F^{jz} = \{D_z \geq j > D_{z-1}\} \) ) denote the subpopulation of compliers induced to go from a treatment below to above \( j \) (resp. above to below \( j \) ) when the instrument moves from \( z - 1 \) to \( z \). Consider the following condition:

**Assumption S3** (Compliers–Defiers for Multivariate Treatment and Instrument (CDMU2)). For every \( j \in \{1, 2, \ldots, J\} \) and \( z \in \{0, 1, 2, \ldots, K\} \), there is a subpopulation of \( C^{jz} \), denoted \( C^{jz}_F \), that satisfies

\[
P(C^{jz}_F) = P(F^{jz}),
\]

\[
E(Y_j - Y_{j-1}|C^{jz}_F) = E(Y_j - Y_{j-1}|F^{jz}).
\]

Under CDMU2, one can show that \( \beta \) is equal to a weighted average of average causal responses for surviving-compliers.

### S3. Fuzzy regression discontinuity designs and survey nonresponse

My results also extend to many treatment effect models relying on ND type conditions. An important example is fuzzy regression discontinuity (RD) designs studied in Hahn, Todd, and Van der Klaauw (2001). Their Theorem 3 still holds if their Assumption A.3, an “ND at the threshold” assumption, is replaced by a “CD at the threshold” assumption. Let \( S \) denote the forcing variable and let \( s \) denote the threshold. Let \( C^s = \{C, S = s\} \) and \( F^s = \{F, S = s\} \), respectively, denote compliers and defiers at the threshold.
Assumption S4 (Compliers–Defiers at the Threshold (CDT)). There is a subpopulation of $C^x$, denoted $C_F^x$, that satisfies

\[
P(C_F^x | S = s) = P(F^x | S = s),
\]
\[
E(Y_1 - Y_0 | C_F^x, S = s) = E(Y_1 - Y_0 | F^x, S = s).
\]

Finally, ND type conditions have also been used in the sample selection and survey nonresponse literatures, for instance, in Lee (2009) or Behaghel, Crépon, Gurgand, and Le Barbanchon (2015). My results also extend to these applications: the results in Lee (2009) or Behaghel, Crépon, Gurgand, and Le Barbanchon (2015) remain valid under a weakening of the ND condition similar in spirit to the CD condition.

Appendix SA: Proofs

Proof of Theorem S1. Proving that $L$ and $\bar{L}$ are valid bounds for $E(Y_1 - Y_0 | C_F^x)$ relies on an argument similar to the proof of Proposition 1.a in Lee (2009). Due to a concern for brevity, I refer the reader to this paper for this part of the proof. The testable implication directly follows.

□

References


Co-editor Petra Todd handled this manuscript.

Manuscript received 10 August, 2015; final version accepted 24 September, 2016; available online 18 October, 2016.