Appendix

Table of Contents

A Figures **A.2**

B Relationship between the ACR and CCE **A.2**

C Identification of the CCE Without Homogeneity **A.3**

C.1 Proof that the CCE is Unconstrained Under Assumptions 1-3 . . . . . A.3

C.2 Bounded Unit Causal Effects A.5

D Identification of CCE with Multiple Instruments **A.5**

E Proof for Proposition 1 **A.6**

F Proof for Proposition 2 **A.6**

G Sharpness of Identification Results **A.7**

G.1 Sharpness of Proposition 1 . . . . . A.9

G.2 Sharpness of Proposition 2 . . . . . A.11

G.3 Illustration A.12

H Identification of CCE with Additional Boundedness Assumptions **A.13**

H.1 Bounded Unit Causal Effects with Homogeneity A.13

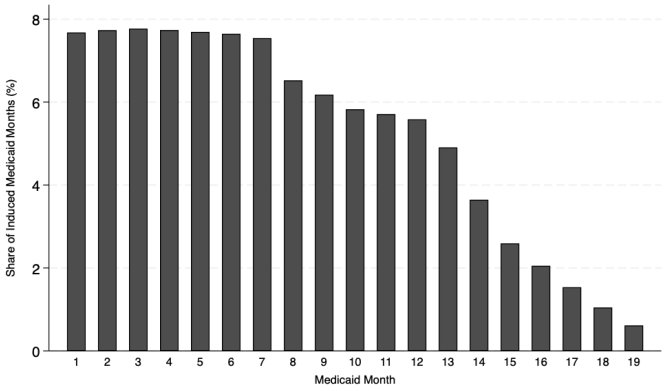
H.2 Bounded Outcomes with Homogeneity A.15

I Identification of Related Parameters **A.16**

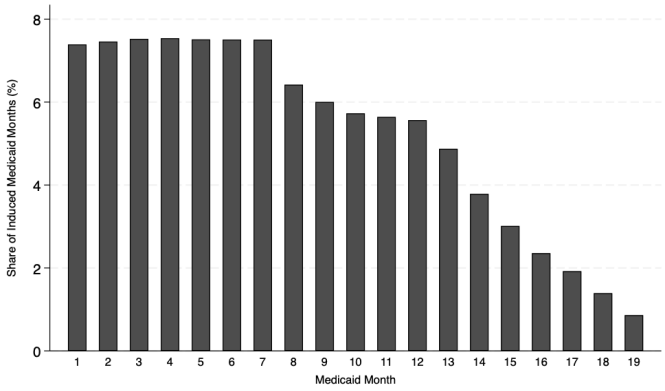
I.1 Proposition 1 . . . . . A.17

I.2 Proposition 2 . . . . . A.18

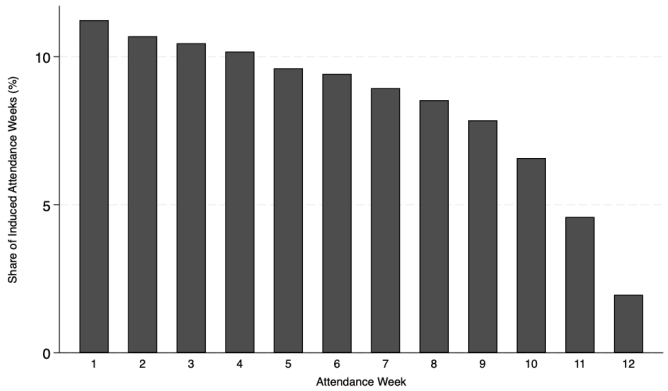
A Figures



(a) Taubman et al. (2014) No Prior ED Use Sample: Medicaid Months



(b) Taubman et al. (2014) Prior ED Use Sample: Medicaid Months



(c) Muralidharan, Singh and Ganimian (2019): Attendance Weeks

Figure A.1: Distribution of Estimation Unit Weights

B Relationship between the ACR and CCE

It will be convenient to introduce the following notation. Let β_{ij} denote the unit-causal effect for individual i at treatment intensity j: $\beta_{ij} = Y_i(j) - Y_i(j-1)$. Let δ_{ij} indicate whether the instrument induces unit i to cross treatment intensity j, $\delta_{ij} = \mathbb{1}\{D_i(1) \geq j > D_i(0)\}$.

Without loss of generality, suppose that i is indexed such that $D_i(1) > D_i(0) \iff i \leq N_c$. In this case, the ACR can be written as:

$$ACR = \frac{\sum_{i=1}^{N}\sum_{j=1}^{J}\delta_{ij}\beta_{ij}}{\sum_{i=1}^{N}\sum_{j=1}^{J}\delta_{ij}}$$

and the CCE can be written as

$$CCE = \frac{1}{N_c}\sum_{i=1}^{N}\sum_{j=1}^{J}\mathbb{1}(D_i(1) > D_i(0))\beta_{ij} = \frac{1}{N_c}\sum_{i=1}^{N_c}\sum_{j=1}^{J}\beta_{ij}$$

Define $\overline{\gamma}$ as the average unit causal effect among compliers: $\overline{\gamma} = CCE/J$. Contrasting $\overline{\gamma}$ with the ACR, the ACR is a weighted average of unit causal effects for compliers and weights corresponding to a specific treatment level. On the other hand, the $\overline{\gamma}$ considers the average effect of a unit increase in treatment for *all* types of compliers if they had all moved along the full length of the causal response. In contrast to the weighting scheme of the ACR, $\overline{\gamma}$ weighs all compliers equally.

Next, note that each unit causal effect, β_{ij}, can be decomposed as follows:

$$\beta_{ij} = \overline{\gamma} + \gamma_i + \gamma_j + \tilde{\gamma}_{ij}$$

where γ_i captures unit-level heterogeneity, $\gamma_i = \frac{1}{J}\sum_j \beta_{ij} - \overline{\gamma}$; γ_j captures heterogeneity across treatment intensities, $\gamma_j = \frac{1}{N_c}\sum_i^{N_c} \beta_{ij} - \overline{\gamma}$; and $\tilde{\gamma}_{ij}$ captures the remaining heterogeneity among unit causal effects, $\tilde{\gamma}_{ij} = \beta_{ij} - \overline{\gamma} - \gamma_i - \gamma_j$. Note that $E[\gamma_i] = E[\gamma_j] = E[\tilde{\gamma}_{ij}] = 0$.

Finally, let δ_i denote i's (mean) change in treatment intensity induced by the instrument, $\delta_i = \frac{1}{J}\sum\limits_{j=1}^{J} \delta_{ij}$; δ_j denote the share of compliers induced by the instrument to cross treatment intensity j, $\delta_j = \frac{1}{N_c}\sum\limits_{i=1}^{N_c} \delta_{ij}$; and $\overline{\delta}$ denote the mean increase in treatment units induced by the instrument, $\overline{\delta} = \frac{1}{N_c}\frac{1}{J}\sum_{i=1}^{N_c}\sum_{j=1}^{J} \delta_{ij}$. Substituting this decomposition and notation into the ACR formula, we have:

$$\begin{aligned}
N_c\, J\, \overline{\delta}\, ACR = \sum_{i=1}^{N_c}\sum_{j=1}^{J} \delta_{ij}\beta_{ij} &= \sum_{i=1}^{N_c}\sum_{j=1}^{J} \delta_{ij}\left(\overline{\gamma} + \gamma_i + \gamma_j + \tilde{\gamma}_{ij}\right) \\
&= \overline{\gamma}\, N_c\, J\, \overline{\delta} + \sum_i \gamma_i \sum_j \delta_{ij} + \sum_j \gamma_j \sum_i \delta_{ij} + \sum_i\sum_j \tilde{\gamma}_{ij}\delta_{ij} \\
&= \overline{\gamma}\, N_c\, J\, \overline{\delta} + \sum_i \gamma_i\, J\, \delta_i + \sum_j \gamma_j\, N_c\, \delta_j + \sum_i\sum_j \tilde{\gamma}_{ij}\delta_{ij} \\
&= \overline{\gamma}\, N_c\, J\, \overline{\delta} + \frac{N_c}{N_c}\, J \sum_i \gamma_i\, \delta_i + \frac{J}{J}\, N_c \sum_j \gamma_j\, \delta_j + \frac{N_c\, J}{N_c\, J}\sum_i\sum_j \tilde{\gamma}_{ij}\delta_{ij} \\
&= \overline{\gamma}\, N_c\, J\, \overline{\delta} + N_c\, J\, \left(\text{Cov}(\gamma_i, \delta_i) + \text{Cov}(\gamma_j, \delta_j) + \text{Cov}(\tilde{\gamma}_{ij}, \delta_{ij})\right)
\end{aligned}$$

Dividing through by $N_c\, J\, \overline{\delta}$, we have:

$$ACR - \overline{\gamma} = \frac{1}{\overline{\delta}}\left(\text{Cov}(\gamma_i, \delta_i) + \text{Cov}(\gamma_j, \delta_j) + \text{Cov}(\tilde{\gamma}_{ij}, \delta_{ij})\right) \tag{3}$$

To obtain Equation (3), we simply plug $\overline{\gamma} = CCE/J$ into the above.

C Identification of the CCE Without Homogeneity

This section explores the bounds on the CCE after relaxing Assumption 4, allowing for potential heterogeneity in the unit causal effects among compliers.

C.1 Proof that the CCE is Unconstrained Under Assumptions 1-3

Let $\delta_{ij} = \mathbb{1}\{D_i(1) \geq j > D_i(0)\}$. We begin by noting that the ACR can be written as

$$ACR = \frac{\sum\limits_{i=1}^{N}\sum\limits_{j=1}^{J} \delta_{ij}\,\beta_{ij}}{\sum\limits_{i=1}^{N}\sum\limits_{j=1}^{J} \delta_{ij}} = \frac{1}{\delta}\sum_{i=1}^{N}\sum_{j=1}^{J} \delta_{ij}\,\beta_{ij}$$

where $\delta = \sum_{i=1}^{N}\sum_{j=1}^{J}\delta_{ij}$. To see this, first note that $P[D_i(1) \geq j > D_i(0)] = \sum_{i=1}^{N}\delta_{ij}/N$. Therefore,

$$w_j = \sum_{i=1}^{N}\delta_{ij} \Big/ \sum_{j'=1}^{J}\sum_{i=1}^{N}\delta_{ij'}$$

Additionally, observe that

$$E[Y_i(j) - Y_i(j-1) \mid D_i(1) \geq j > D_i(0)] = \frac{\sum_{i=1}^{N}\beta_{ij}\delta_{ij}}{\sum_{i=1}^{N}\delta_{ij}}$$

Putting the above together,

$$\begin{aligned} ACR &= \sum_{j=1}^{J} w_j E[Y_i(j) - Y_i(j-1) \mid D_i(1) \geq j > D_i(0)] \\ &= \sum_{j=1}^{J} \frac{\sum_{i=1}^{N}\delta_{ij}}{\sum_{j=1}^{J}\sum_{i=1}^{N}\delta_{ij}} \frac{\sum_{i=1}^{N}\beta_{ij}\delta_{ij}}{\sum_{i=1}^{N}\delta_{ij}} = \frac{\sum_{j=1}^{J}\sum_{i=1}^{N}\beta_{ij}\delta_{ij}}{\sum_{j=1}^{J}\sum_{i=1}^{N}\delta_{ij}} = \frac{1}{\delta}\sum_{i=1}^{N}\sum_{j=1}^{J}\delta_{ij}\,\beta_{ij} \end{aligned}$$

Rearranging, we get:

$$\delta\, ACR = \sum_{i=1}^{N}\sum_{j=1}^{J}\delta_{ij}\,\beta_{ij} \tag{4}$$

Let N_c be the number of compliers, $N_c := \sum_{i=1}^{N}\mathbb{1}(D_i(1) > D_i(0))$.

$$\begin{aligned} CCE &= \frac{1}{N_c}\left[\sum_{i=1}^{N}\sum_{j=1}^{J}\mathbb{1}(D_i(1) > D_i(0))\beta_{ij}\right] \\ &= \frac{1}{N_c}\left[\sum_{i=1}^{N}\sum_{j=1}^{J}\mathbb{1}(D_i(1) > D_i(0))\Big(\beta_{ij} + \beta_{ij}\delta_{ij} - \beta_{ij}\delta_{ij}\Big)\right] \\ &= \frac{1}{N_c}\left[\sum_{i=1}^{N}\sum_{j=1}^{J}\mathbb{1}(D_i(1) > D_i(0))\beta_{ij}\delta_{ij} + \sum_{i=1}^{N}\sum_{j=1}^{J}\mathbb{1}(D_i(1) > D_i(0))(1-\delta_{ij})\beta_{ij}\right] \\ &= \frac{1}{N_c}\left[\sum_{i=1}^{N}\sum_{j=1}^{J}\beta_{ij}\delta_{ij} + \sum_{i=1}^{N}\sum_{j=1}^{J}\mathbb{1}(D_i(1) > D_i(0))(1-\delta_{ij})\beta_{ij}\right] \\ &= \frac{1}{N_c}\left[\delta\, ACR + \sum_{i=1}^{N}\sum_{j=1}^{J}\mathbb{1}(D_i(1) > D_i(0))(1-\delta_{ij})\beta_{ij}\right] \end{aligned}$$

where the last equality uses (4).

To see that the CCE is unbounded under assumptions 1-3, note that we can always increase (or decrease) β_{ij} arbitrarily for any i and j pair with $\delta_{ij} = 0$ and $D_i(1) > D_i(0)$, since these unit causal effects are unconstrained by the ACR or other

observable moments. Therefore, as long as every single complier doesn't move along the full length of the dose response function, i.e., there exists some i for which $D_i(1) > D_i(0)$ and either $D_i(1) < J$ or $D_i(0) > 0$, the CCE is unbounded absent additional assumptions.[11]

C.2 Bounded Unit Causal Effects

One approach to restricting the free unit-causal effects is to impose an additional assumption on the support of the unit-causal effects, $\beta_{ij} \in \left[\underline{\beta}, \overline{\beta}\right]$. Suppose that $\beta_{ij} = \beta$ for all (i, j) pairs such that $\delta_{ij} = 0$. We can then express the CCE (scaled by J) as a convex combination of the ACR and β:

$$CCE/J = \left(\frac{\delta}{J\,N_c}\right) ACR + \left(1 - \frac{\delta}{J\,N_c}\right) \beta$$

where the weight on the ACR, $\frac{\delta}{J\,N_c}$, corresponds to the fraction of the maximum potential effect of the instrument that is reflected in the ACR. We can therefore express the maximum CCE as

$$\begin{aligned}\overline{CCE} &= \left(\frac{\delta}{J\,N_c}\right) J\,ACR + \left(1 - \frac{\delta}{J\,N_c}\right) J\,\overline{\beta} \\ &= \frac{1}{N_c}\left[\delta\,ACR + (J\,N_c - \delta)\overline{\beta}\right]\end{aligned}$$

Analogously, the minimum CCE is

$$\underline{CCE} = \frac{1}{N_c}\left[\delta\,ACR + (J\,N_c - \delta)\underline{\beta}\right].$$

Note that the width of these bounds is increasing in N_c, which is not point-identified by the data. The maximum value of N_c, and hence the full range of CCE values consistent with the data, can be estimated based on the method described in Huang et al. (2016).

D Identification of CCE with Multiple Instruments

In this section, we consider identification of the CCE in the presence of additional identifying variation in the form of multiple instruments. Suppose there are $K \leq J$ mutually orthogonal binary instruments, $Z_k \in \{0, 1\}$. Assume that Assumption 1-3 are satisfied for each Z_k. Angrist and Imbens (1995) show that using K orthogonal indicators as instruments can equivalently be thought of as using a single $K + 1$ valued instrument, $Z_i \in \{0, 1, ..., K\}$. Let $\{w_j^k\}_{j=1}^J$ and ACR^k be the weighting scheme and ACR corresponding to instrument $d_k = \mathbb{1}(Z = k)$ respectively. Angrist and Imbens (1995) show that the K linearly independent dummy variables, $d_k = \mathbb{1}(Z_i = k)$, can be used as instruments to identify K linearly independent ACRs where

$$ACR^k = \frac{E[Y \mid Z = k] - E[Y \mid Z = 0]}{E[D \mid Z = k] - E[D \mid Z = 0]}$$

Recall that under Assumption 4 (homogeneity across complier types), we can re-write ACR^k as

$$ACR^k = \sum_{j=1}^{J} w_j^k \beta_j \tag{5}$$

for each $k = 1, ..., K$. When $K = J$, we have a system of J equations with J unknowns. Under an assumption that this system is full-rank, this allows for point-identification of each unit-causal effects, β_j. Therefore, when $K = J$, one can

[11] If all compliers do move along the full length of the dose response function, the setting can be analyzed as involving a binary treatment. In this case, the CCE is point-identified, and is equal to $ACR \times J$.

point-identify any function of the unit-causal effects including the CCE under Assumptions 1-4 and an additional full-rank assumption (requiring that the instruments induce compliers to move across distinct regions of the dose-response function).

Next, we consider identification in the case where $K < J$. In this case, we can use the additional identifying variation available through the additional instruments to tighten the bounds on the CCE. Specifically, we can incorporate (5) as a constraint in **LP.1** and **LP.2** for each $k = 1, .., K$:

$$\underset{\{\beta_j\}}{\text{Maximize/Minimize}} \quad \sum_{j=1}^{J} \beta_j \qquad \textbf{LP.3}$$

$$\text{subject to} \quad \sum_{j=1}^{J} \beta_j w_j^k = ACR^k \qquad \forall k = 1, ..., K \qquad \text{(C.1)}$$

$$\beta_j \geq 0 \qquad \forall\, j = 1, ..., J \qquad \text{(C.2)}$$

$$\beta_j \geq \beta_{j+1} \qquad \forall\, j = 1, ..., J-1 \qquad \text{(C.3)}$$

The extent to which incorporating additional instruments into the analysis will tighten the bounds on the CCE will depend on the degree to which the additional instruments induce compliers to move across different portions of the treatment dose response function. For example, if $J = 2$ and a first instrument primarily moves compliers from $j = 0$ to $j = 1$, incorporating a second instrument will tend to tighten the estimate of the CCE if it primarily moves compliers from $j = 1$ to $j = 2$. In contrast, it will tend to provide less new information if it also primarily moves compliers from $j = 0$ to $j = 1$.

E Proof for Proposition 1

To establish that $CCE \geq ACR/w_{\bar{j}}$, we will first show that CCE is minimized when $\beta_{\bar{j}}$ is the only non-zero unit causal effect, i.e., $\beta_j = 0$ for all $j \neq \bar{j}$. Proceeding by contradiction, suppose instead that CCE was minimized by a vector of unit treatment effects $(\beta_1, ..., \beta_J)$ with $\beta_k > 0$ for some $k \neq \bar{j}$. Consider a new vector of unit treatment effects, $(\beta_1^*, ..., \beta_J^*)$ identical to the first, except that the specified non-zero effect has been set to zero, $\beta_k^* = 0$, and the unit effect corresponding to the largest weight, $\beta_{\bar{j}}$, has been increased to maintain feasibility with respect to the ACR constraint: $\beta_{\bar{j}}^* = \beta_{\bar{j}} + \beta_k\, w_k/w_{\bar{j}}$. But, with these changes, the value of the objective function is lower than that of the original vector of unit effects:

$$\sum \beta_j^* - \sum \beta_j = \beta_k \frac{w_k}{w_{\bar{j}}} - \beta_k < 0,$$

where the last inequality follows from Assumption 5 and the definition of $\bar{j}$. Thus, the initial vector of unit treatment effects must not be the solution to the constrained minimization problem, proving the contradiction. The proof that $CCE \leq ACR/w_{\underline{j}}$ is analogous. ■

F Proof for Proposition 2

To show that $CCE \geq ACR/w_1$, we will first show that the CCE is minimized when $\beta_j = 0\ \forall\ j > 1$. Proceeding by contradiction, suppose that the CCE is minimized by a vector of unit treatment effects $(\beta_1, ..., \beta_J)$ with $\beta_j > 0$ for some $j > 1$. Let k denote the highest intensity non-zero unit effect, more precisely $k = \max_j\{2, ..., J \mid \beta_j > 0\}$. Along with Assumption 6, this implies $\beta_1 \geq \beta_2 \geq ... \geq \beta_k > 0$. Consider a new vector of unit treatment effects, $(\tilde{\beta}_1, ..., \tilde{\beta}_J)$, which is identical to the first vector, except that β_k has been set to zero, $\tilde{\beta}_k = 0$, and the first unit effect has been increased to maintain feasibility with respect to the ACR constraint: $\tilde{\beta}_1 = \beta_1 + \beta_k w_k/w_1$. But, with these changes, the value of the

objective function is lower than that of the original vector of unit effects:

$$\sum \tilde{\beta}_j - \sum \beta_j = \beta_k \frac{w_k}{w_1} - \beta_k < 0,$$

where the last inequality follows from Assumptions 5 and 7. Thus, the initial vector of unit treatment effects must not be the solution to the constrained minimization problem, proving the contradiction. Finally, to prove the result, note that when $\beta_j = 0 \; \forall \; j > 1$, it follows that $CCE = \sum_j \beta_j = \beta_1 = ACR/w_1$, where the last equality follows from the definition of the ACR.

To show that $CCE \leq ACR \times J$, we will first show that the CCE is maximized when the unit causal effects are equalized, i.e., when $\beta_1 = \beta_2 = ... = \beta_J$. Proceeding by contradiction, suppose that the CCE is maximized by a vector of unit treatment effects $(\beta_1, \beta_2, ..., \beta_J)$ with $\beta_k > \beta_{k+1}$ for some $k \in \{1, ..., J-1\}$ (concavity rules out the reverse ordering). Consider a new vector of unit treatment effects, $(\tilde{\beta}_1, ..., \tilde{\beta}_J)$, which is identical to the first vector, except that $\tilde{\beta}_k = \beta_k - \frac{(\beta_k - \beta_{k+1})w_{k+1}}{w_k + w_{k+1}}$ and $\tilde{\beta}_{k+1} = \beta_{k+1} + \frac{w_k(\beta_k - \beta_{k+1})}{w_k + w_{k+1}}$, so that $\tilde{\beta}_k = \tilde{\beta}_{k+1}$. Note that $(\tilde{\beta}_1, ..., \tilde{\beta}_J)$ is a feasible solution vector since it satisfies C.1-C.3. But, with these changes, the value of the CCE is higher than under that of the original vector of the unit effects:

$$\sum \tilde{\beta}_j - \sum \beta_j = \frac{(\beta_{k+1} - \beta_k)(w_{k+1} - w_k)}{w_{k+1} + w_k} > 0,$$

where the last inequality follows from the assumption that $\beta_k > \beta_{k+1}$ (Assumption 6) and that $w_k > w_{k+1}$ (Assumption 7). Thus, the initial vector of unit treatment effects must not be the solution to the constrained maximization problem, proving the contradiction. ■

G Sharpness of Identification Results

In this section, we prove that the bounds on the CCE in Proposition 1 and 2 are sharp, in the sense that all values of the CCE in the identified sets in Propositions 1 and 2 are compatible with the assumptions and the data (e.g., Kline and Tamer, 2023).

The data we observe is the joint distribution of (Y_i, D_i, Z_i). We will show that for identification of the CCE, the relevant moments of the joint distribution are $E[Y \mid D = j, Z = z]$ and $P[D = j \mid Z = z]$ for $j \in \{0, 1..., J\}$ and $z \in \{0, 1\}$. Under the assumptions of instrument validity (Assumptions 1-3), we establish the relationship between the unobserved joint distribution of $D_i(0)$ and $D_i(1)$ and the observed moments. Similarly, we show how the observed moments constrain the relationship between the outcome and treatment dosage under Assumptions 1-3. Next, we posit a dose-response relationship between the treatment and the outcome that attains the bounds in Propositions 1 and 2. We then show that this relationship is consistent with the assumptions and the observed moments, thereby establishing sharpness of our proposed bounds. We illustrate our argument in a simple example with three treatment levels ($J = 2$).

We begin by enumerating the relevant observed moments of the joint distribution of (Y_i, D_i, Z_i). Let $P[D_i(0) = m, D_i(1) = n] \equiv \pi_{mn}$ and $E[Y_i(j)|D_i(0) = m, D_i(1) = n] \equiv Y_{mn}(j)$ for $j, m, n \in \{0, 1, ..., J\}$. Using this notation, monotonicity (Assumption 3) implies $\pi_{mn} = 0$ for all m, n with $m > n$. Imposing independence (Assumption 2), we obtain the following constraints on the joint distribution of $D_i(0)$ and $D_i(1)$:

$$P[D = k \mid Z = 0] = P[D(0) = k \mid Z = 0] = P[D(0) = k] = \sum_{j=k}^{J} \pi_{kj} \tag{6}$$

$$P[D = k \mid Z = 1] = P[D(1) = k \mid Z = 1] = P[D(1) = k] = \sum_{j=0}^{k} \pi_{jk} \tag{7}$$

for each $k \in \{0, 1, ..., J\}$.

Relevance of the instrument (Assumption 1) implies that $E[D_i \mid Z_i = 1] - E[D_i \mid Z_i = 0] = \sum_{j=1}^{J} P[D_i(1) \geq j >$

$D_i(0)] = \sum_{j=1}^{J} \sum_{p=j}^{J} \sum_{q=0}^{j-1} \pi_{qp} > 0$.

Turning to the relationship between the treatment dosage and the outcome, $Y_{mn}(m)$ and $Y_{mn}(n)$ are similarly constrained by the observed moments under independence and monotonicity (Assumptions 2-3):

$$E[Y \mid Z = 0, D = k]P[D = k \mid Z = 0] = \sum_{j=k}^{J} \pi_{kj} Y_{kj}(k) \tag{8}$$

$$E[Y \mid Z = 1, D = k]P[D = k \mid Z = 1] = \sum_{j=0}^{k} \pi_{jk} Y_{jk}(k) \tag{9}$$

for each $k \in \{0, 1, ..., J\}$.

Next, we show that while we observe the entire joint distribution of (Y_i, D_i, Z_i), the only features of the distribution that are relevant for identification of the CCE are $E[Y|D = j, Z = z]$ and $P(D = j \mid Z = z)$ for $j \in \{0, 1, ..., J\}$ and $z \in \{0, 1\}$. We show this by expressing the CCE under homogeneity (Assumption 4) as a function of the features of the joint distribution of (Y_i, D_i, Z_i) that are constrained by (6)-(9). Specifically, we show that the CCE can be expressed as follows:

$$\begin{aligned} CCE = \frac{1}{P[D = J \mid Z = 1] - P[D = J \mid Z = 0]} \times \\ \Big[E[Y \mid Z = 1, D = J]P[D = J \mid Z = 1] - E[Y \mid Z = 0, D = J]P[D = J \mid Z = 0] \\ + E[Y \mid Z = 1, D = 0]P[D = 0 \mid Z = 1] - E[Y \mid Z = 0, D = 0]P[D = 0 \mid Z = 0] \\ + \sum_{j=1}^{J-1} \Big(\pi_{0j} Y_{0j}(0) - \pi_{jJ} Y_{jJ}(0) \Big) \Big] \end{aligned}$$

To see this, we start with the observed moment conditions (8) and (9) for $D = 0$ and $D = J$ respectively:

$$\begin{aligned} & E[Y \mid Z = 1, D = J]P[D = J \mid Z = 1] - E[Y \mid Z = 0, D = 0]P[D = 0 \mid Z = 0] \\ &= \sum_{j=0}^{J} \pi_{jJ} Y_{jJ}(J) - \sum_{j=0}^{J} \pi_{0j} Y_{0j}(0) \\ &= \pi_{0J} \Big(Y_{0J}(J) - Y_{0J}(0) \Big) + \sum_{j=1}^{J} \pi_{jJ} Y_{jJ}(J) - \sum_{j=0}^{J-1} \pi_{0j} Y_{0j}(0) \\ &= \pi_{0J} CCE + \pi_{JJ} Y_{JJ}(J) - \pi_{00} Y_{00}(0) + \sum_{j=1}^{J-1} \Big(\pi_{jJ} Y_{jJ}(J) - \pi_{0j} Y_{0j}(0) \Big) \\ &= \pi_{0J} CCE + \pi_{JJ} Y_{JJ}(J) - \pi_{00} Y_{00}(0) + \sum_{j=1}^{J-1} \Big(\pi_{jJ} CCE + \pi_{jJ} Y_{jJ}(0) - \pi_{0j} Y_{0j}(0) \Big) \\ &= \Big(\sum_{j=0}^{J} \pi_{jJ} - \pi_{JJ} \Big) CCE + \pi_{JJ} Y_{JJ}(J) - \pi_{00} Y_{00}(0) + \sum_{j=1}^{J-1} \Big(\pi_{jJ} Y_{jJ}(0) - \pi_{0j} Y_{0j}(0) \Big) \end{aligned}$$

Plugging in moment conditions (6) and (7) for $D = J$, and moment conditions (8) and (9) for $D = J$ and $D = 0$

respectively, allows us to rewrite the above as

$$
\begin{aligned}
&E[Y \mid Z=1, D=J]P[D=J \mid Z=1] - E[Y \mid Z=0, D=0]P[D=0 \mid Z=0] \\
&= \Big(P[D=J \mid Z=1] - P[D=J \mid Z=0]\Big)CCE \\
&\quad + E[Y \mid Z=0, D=J]P[D=J \mid Z=0] - E[Y \mid Z=1, D=0]P[D=0 \mid Z=1] \\
&\quad + \sum_{j=1}^{J-1} \Big(\pi_{jJ} Y_{jJ}(0) - \pi_{0j} Y_{0j}(0)\Big)
\end{aligned}
$$

Rearranging terms, we can re-write the CCE as follows,

$$
\begin{aligned}
CCE = \frac{1}{P[D=J \mid Z=1] - P[D=J \mid Z=0]} \times \\
\Bigg[E[Y \mid Z=1, D=J]P[D=J \mid Z=1] - E[Y \mid Z=0, D=J]P[D=J \mid Z=0] \\
+ E[Y \mid Z=1, D=0]P[D=0 \mid Z=1] - E[Y \mid Z=0, D=0]P[D=0 \mid Z=0] \\
+ \sum_{j=1}^{J-1} \Big(\pi_{0j} Y_{0j}(0) - \pi_{jJ} Y_{jJ}(0)\Big) \Bigg]
\end{aligned}
$$

This shows that the CCE is only a function of the features of the joint distribution of (Y_i, D_i, Z_i) that are constrained by (6)-(9).

G.1 Sharpness of Proposition 1

Recall that Proposition 1 shows that, under Assumptions 1-5, $CCE \in \left[\frac{ACR}{w_{\bar{j}}}, \frac{ACR}{w_{\underline{j}}}\right]$ where $\underline{j} = \arg\min_{j \in \{1,...J\}}\{w_j\}$ and $\bar{j} = \arg\max_{j \in \{1,...J\}}\{w_j\}$. To simplify exposition, we additionally assume that the weights are monotonically declining, so $\bar{j} = 1$ and $\underline{j} = J$; the same logic extends to the case where this condition does not hold .[12]

Consider the following candidate relationship between the outcome and treatment dosage for all complier types: $Y_{mn}(0) = Y_{mn}(1) = ... = Y_{mn}(J-1)$ and $Y_{mn}(J) - Y_{mn}(J-1) = ACR/w_J$ for all $n > m$ and $m, n \in \{0, 1, ..., J\}$.

We begin by showing that this candidate dose-response relationship is consistent with Assumptions 1-5. Since the choice of the dose-response relationship does not affect the validity of the instrument, it is consistent with Assumptions 1-3. Assumption 4 imposes that the unit causal effects, $Y_{mn}(j) - Y_{mn}(j-1)$ are constant across all $n > m$ and $m, n \in \{0, 1, ..., J\}$. Since we propose the same candidate dose-response relationship for all complier types, Assumption 4 is satisfied by construction. It is also easy to see that the unit causal effects for all compliers, $\beta_j = Y_{mn}(j) - Y_{mn}(j-1)$ are non-negative, satisfying Assumption 5.

Next, we show that this candidate dose-response relationship attains the maximum value of the CCE in Proposition 1. Under Assumption 4, the CCE can be written as $Y_{mn}(J) - Y_{mn}(0)$ for any $n > m$ and $m, n \in \{0, 1, ..., J\}$ (i.e., for any complier). Adding and subtracting $Y_{mn}(J-1)$, we have $CCE = Y_{mn}(J) - Y_{mn}(J-1) + Y_{mn}(J-1) - Y_{mn}(0) = ACR/w_J$, where the final equality follows from the fact that $Y_{mn}(J-1) = Y_{mn}(0)$ under the candidate dose response function. This shows that the candidate dose-response relationship attains the maximum value in Proposition 1.

To show that the upper bound in Proposition 1 is sharp, it remains to show that the candidate dose-response relationship is consistent with the observable moments. We prove this by showing that there exists a vector of unobserved potential outcomes that satisfies Equations (6)-(9) under the candidate dose-response relationship. We begin by plugging in the

[12] If set of weights $\{w_j\}$ are not monotonically declining, the dose-response relationship that would attain the bounds in Proposition 1 would maximize the unit effect at the range of the dose response function where the weight is smallest. The remainder of the proof would follow similarly.

dose-response relationship $Y_{mn}(0) = Y_{mn}(k)$ for $k \in \{1, ..., J-1\}$ in (8) and (9):

$$E[Y \mid Z=0, D=k]P[D=k \mid Z=0] = \pi_{kk}Y_{kk}(k) + \sum_{j=k+1}^{J} \pi_{kj}Y_{kj}(0)$$

$$E[Y \mid Z=1, D=k]P[D=k \mid Z=1] = \pi_{kk}Y_{kk}(k) + \sum_{j=0}^{k-1} \pi_{jk}Y_{jk}(0)$$

For $k = 0$ in (9),

$$E[Y \mid Z=1, D=0]P[D=0 \mid Z=1] = \pi_{00}Y_{00}(0)$$

Plugging the above into (8) for $k = 0$,

$$E[Y \mid Z=0, D=0]P[D=0 \mid Z=0] - E[Y \mid Z=1, D=0]P[D=0 \mid Z=1] = \sum_{j=1}^{J} \pi_{0j}Y_{0j}(0)$$

For $k = J$ in (8),

$$E[Y \mid Z=0, D=J]P[D=J \mid Z=0] = \pi_{JJ}Y_{JJ}(J)$$

Plugging the above into (9) for $k = J$,

$$E[Y \mid Z=1, D=J]P[D=J \mid Z=1] - E[Y \mid Z=0, D=J]P[D=J \mid Z=0] = \sum_{j=0}^{J-1} \pi_{jJ}Y_{jJ}(J)$$

Next, plugging in the dose-response relationship, $Y_{jJ}(J) = ACR/w_J + Y_{jJ}(0)$, into the above equation yields the following:

$$\begin{aligned} E[Y \mid Z=1, D=J]P[D=J \mid Z=1] - E[Y \mid Z=0, D=J]P[D=J \mid Z=0] \\ = \frac{ACR}{w_J}\left(\sum_{j=0}^{J-1} \pi_{jJ}\right) + \sum_{j=0}^{J-1} \pi_{jJ}Y_{jJ}(0) \end{aligned}$$

Together, this yields the following system of $2J$ equations:

$$E[Y \mid Z=0, D=k]P[D=k \mid Z=0] = \pi_{kk}Y_{kk}(k) + \sum_{j=k+1}^{J} \pi_{kj}Y_{kj}(0) \quad \text{for} \quad k = 1, ..., J-1$$

$$E[Y \mid Z=1, D=k]P[D=k \mid Z=1] = \pi_{kk}Y_{kk}(k) + \sum_{j=0}^{k-1} \pi_{jk}Y_{jk}(0) \quad \text{for} \quad k = 1, ..., J-1$$

$$E[Y \mid Z=0, D=0]P[D=0 \mid Z=0] - E[Y \mid Z=1, D=0]P[D=0 \mid Z=1] = \sum_{j=1}^{J} \pi_{0j}Y_{0j}(0)$$

$$\begin{aligned} E[Y \mid Z=1, D=J]P[D=J \mid Z=1] - E[Y \mid Z=0, D=J]P[D=J \mid Z=0] \\ = \frac{ACR}{w_J}\left(\sum_{j=0}^{J-1} \pi_{jJ}\right) + \sum_{j=0}^{J-1} \pi_{jJ}Y_{jJ}(0) \end{aligned}$$

In this system of equations, $Y_{mn}(0)$ is unknown for all $n > m$ and $m, n \in \{0, 1, ..., J\}$. Additionally, $Y_{kk}(k)$ is unknown for all $k \in \{1, ..., J-1\}$. Together, this implies that there are $\frac{J(J+1)}{2} + J - 1$ unknowns. For all $J \geq 2$, this implies that there are at least as many unknown potential outcomes as there are equations, and so there is at least 1 vector of unknown potential outcomes that satisfies this set of equations.

Together, this shows that the upper bound for the CCE in Proposition 1 can be attained under a dose-response relationship between the treatment and the outcome that is consistent with the assumptions and the data. This establishes that the bounds in Proposition 1 are sharp. The proof for the lower bound is symmetric.

G.2 Sharpness of Proposition 2

Recall that Proposition 2 shows that, under Assumptions 1-7, the CCE is bounded above by $ACR \times J$ and bounded below by ACR/w_1. Below, we provide a proof for the sharpness of the upper bound. The proof that the lower bound is sharp is symmetric to the proof in Proposition 1.

Consider the following relationship between the outcome and treatment dosage for all complier types: $Y_{mn}(j) - Y_{mn}(j-1) = ACR$ for all $j \in \{1, ..., J\}$ and all $n > m$.

We begin by showing that this candidate dose-response relationship is consistent with Assumptions 1-6. Since the choice of the dose-response relationship does not affect the validity of the instrument, it is consistent with Assumptions 1-3. Assumption 4 imposes that the unit causal effects, $Y_{mn}(j) - Y_{mn}(j-1)$ are constant across all $n > m$ and $m, n \in \{0, 1, ..., J\}$. Since we propose the same candidate dose-response relationship for all complier types, Assumption 4 is satisfied by construction. It is also easy to see that the unit causal effects for all compliers, $Y_{mn}(j) - Y_{mn}(j-1) = ACR$ are non-negative, satisfying Assumption 5. Assumption 6 imposes that the dose-response relationship between the treatment and outcome is (weakly) concave, i.e., $Y_{mn}(j) - Y_{mn}(j-1) \geq Y_{mn}(j+1) - Y_{mn}(j)$ for all $j \in \{1, 2, ..., J-1\}$. The candidate dose-response relationship we specify is such that $Y_{mn}(j) - Y_{mn}(j-1) = Y_{mn}(j+1) - Y_{mn}(j)$ for all $j \in \{1, 2, ..., J-1\}$, and is therefore compatible with Assumption 6.

Next, we show that the candidate dose-response relationship attains the maximum value of the CCE in Proposition 2. Under Assumption 4, the CCE can be written as $\sum_{j=1}^{J} Y_{mn}(j) - Y_{mn}(j-1)$ for any $n > m$ and $m, n \in \{0, 1, ..., J\}$. Plugging in $Y_{mn}(j) - Y_{mn}(j-1) = ACR$, we obtain that $CCE = ACR \times J$. This shows that our candidate dose-response relationship attains the maximum value in Proposition 2.

To show that the upper bound in Proposition 2 is sharp, it remains to show that the candidate dose-response relationship is consistent with the observable moments. We prove this by showing that there exists a vector of unobserved potential outcomes that satisfies Equations (6)-(9) under the specified dose-response relationship.

We begin by plugging in the dose-response relationship $Y_{mn}(k) = Y_{mn}(0) + ACR \times k$ for $k \in \{1, ..., J-1\}$ in (8) and (9):

$$E[Y \mid Z = 0, D = k]P[D = k \mid Z = 0] = \pi_{kk} Y_{kk}(k) + \sum_{j=k+1}^{J} \pi_{kj} \Big(Y_{kj}(0) + ACR \times k \Big)$$

$$E[Y \mid Z = 1, D = k]P[D = k \mid Z = 1] = \pi_{kk} Y_{kk}(k) + \sum_{j=0}^{k-1} \pi_{jk} \Big(Y_{jk}(0) + ACR \times k \Big)$$

For $k = 0$ in (9),

$$E[Y \mid Z = 1, D = 0]P[D = 0 \mid Z = 1] = \pi_{00} Y_{00}(0)$$

Plugging in the above into (8) for $k = 0$,

$$E[Y \mid Z = 0, D = 0]P[D = 0 \mid Z = 0] - E[Y \mid Z = 1, D = 0]P[D = 0 \mid Z = 1] = \sum_{j=1}^{J} \pi_{0j} Y_{0j}(0)$$

For $k = J$ in (8),

$$E[Y \mid Z = 0, D = J]P[D = J \mid Z = 0] = \pi_{JJ} Y_{JJ}(J)$$

Plugging in the above into (9) for $k = J$,

$$E[Y \mid Z = 1, D = J]P[D = J \mid Z = 1] - E[Y \mid Z = 0, D = J]P[D = J \mid Z = 0] = \sum_{j=0}^{J-1} \pi_{jJ} Y_{jJ}(J)$$

Next, plugging in the dose-response relationship, $Y_{jJ}(J) = ACR \times J + Y_{jJ}(0)$, into the above equation yields the following:

$$E[Y \mid Z = 1, D = J]P[D = J \mid Z = 1] - E[Y \mid Z = 0, D = J]P[D = J \mid Z = 0] = \sum_{j=0}^{J-1} \pi_{jJ} \Big(Y_{jJ}(0) + ACR \times J \Big)$$

Together, this yields the following system of $2J$ equations:

$$E[Y \mid Z = 0, D = k]P[D = k \mid Z = 0] = \pi_{kk} Y_{kk}(k) + \sum_{j=k+1}^{J} \pi_{kj} \Big(Y_{kj}(0) + ACR \times k \Big) \quad \text{for} \quad k = 1, ..., J-1$$

$$E[Y \mid Z = 1, D = k]P[D = k \mid Z = 1] = \pi_{kk} Y_{kk}(k) + \sum_{j=0}^{k-1} \pi_{jk} \Big(Y_{jk}(0) + ACR \times k \Big) \quad \text{for} \quad k = 1, ..., J-1$$

$$E[Y \mid Z = 0, D = 0]P[D = 0 \mid Z = 0] - E[Y \mid Z = 1, D = 0]P[D = 0 \mid Z = 1] = \sum_{j=1}^{J} \pi_{0j} Y_{0j}(0)$$

$$E[Y \mid Z = 1, D = J]P[D = J \mid Z = 1] - E[Y \mid Z = 0, D = J]P[D = J \mid Z = 0] = \sum_{j=0}^{J-1} \pi_{jJ} \Big(Y_{jJ}(0) + ACR \times J \Big)$$

In this system of equations, $Y_{mn}(0)$ is unknown for all $n > m$ and $m, n \in \{0, 1, ..., J\}$. Additionally, $Y_{kk}(k)$ is unknown for all $k \in \{1, ..., J-1\}$. Together, this implies that there are $\frac{J(J+1)}{2} + J - 1$ unknowns. For all $J \geq 2$, this implies that there are at least as many unknown potential outcomes as there are equations, and so there is at least 1 vector of unknown potential outcomes that satisfies this set of equations.

Together, this shows that the upper bound for the CCE in Proposition 2 can be attained under a dose-response relationship between the treatment and the outcome that is consistent with the assumptions and the data. The proof for the lower bound follows the same logic as Section G.1.[13]

G.3 Illustration

In this section, we illustrate the sharpness of the Proposition 1 bounds in the simple case where $J = 2$.

Assume the same candidate dose-response relationship as considered in Section G.1, $Y_{mn}(0) = Y_{mn}(1)$ and $Y_{mn}(2) = ACR/w_2$ for all $n > m$. We showed that this dose-response relationship attains the upper bound of the CCE provided in Proposition 1 and is consistent with Assumptions 1-5. Here, we provide a vector of the unknown potential outcomes that is consistent with (6)-(9) to illustrate the sharpness of the provided bounds. Re-writing Equations (8) and (9) using the

[13] The dose-response relationship under which the lower bound is attained is given by $Y_{mn}(1) - Y_{mn}(0) = ACR/w_1$ and $Y_{mn}(1) = ... = Y_{mn}(J)$ for all $n > m$. In addition to the proof in Section G.1, one only needs to show that this dose-response relationship is also consistent with Assumption 6. This is easy to see, since $\beta_1 = ACR/w_1 > 0 = \beta_j$ for all $j > 1$.

specified dose-response relationship, as in Section G.1:

$$E[Y \mid Z=0, D=1]P[D=1 \mid Z=0] = \pi_{11}Y_{11}(1) + \pi_{12}Y_{12}(0)$$
$$E[Y \mid Z=1, D=1]P[D=1 \mid Z=1] = \pi_{11}Y_{11}(1) + \pi_{01}Y_{01}(0)$$
$$E[Y \mid Z=1, D=2]P[D=2 \mid Z=1] - E[Y \mid Z=0, D=2]P[D=2 \mid Z=0] = \frac{ACR}{w_J}\Big(\pi_{02}+\pi_{12}\Big) + \pi_{02}Y_{02}(0) + \pi_{12}Y_{12}(0)$$
$$E[Y \mid Z=0, D=0]P[D=0 \mid Z=0] - E[Y \mid Z=1, D=0]P[D=0 \mid Z=1] = \pi_{01}Y_{01}(0) + \pi_{02}Y_{02}(0)$$

The vector of unknowns here is $\pi_{01}Y_{01}(0)$, $\pi_{02}Y_{02}(0)$, $\pi_{11}Y_{11}(1)$, and $\pi_{12}Y_{12}(0)$. Assume the following values of the unknown potential outcomes:

$$\pi_{01}Y_{01}(0) = E[Y \mid Z=1, D=1]P[D=1 \mid Z=1] - E[Y \mid Z=0, D=1]P[D=1 \mid Z=0]$$
$$\pi_{02}Y_{02}(0) = E[Y \mid Z=1, D=2]P[D=2 \mid Z=1] - E[Y \mid Z=0, D=2]P[D=2 \mid Z=0] - \frac{ACR}{w_J}\Big(\pi_{02}+\pi_{12}\Big)$$
$$\pi_{11}Y_{11}(1) = E[Y \mid Z=0, D=1]P[D=1 \mid Z=0]$$
$$\pi_{12}Y_{12}(0) = 0$$

It is easy to see simply by plugging in the assumed values of the unknown potential outcomes that this vector of unknowns satisfies the first three of the four moment conditions shown above. It remains to show that

$$E[Y \mid Z=0, D=0]P[D=0 \mid Z=0] - E[Y \mid Z=1, D=0]P[D=0 \mid Z=1] = \pi_{01}Y_{01}(0) + \pi_{02}Y_{02}(0)$$
$$\Leftrightarrow \frac{ACR}{w_J} = \frac{1}{\pi_{02}+\pi_{12}} \times \left(\sum_{j=0}^{2} E[Y \mid Z=1, D=j]P[D=j \mid Z=1] - E[Y \mid Z=0, D=j]P[D=j \mid Z=0] \right)$$

Using the law of total probability, the numerator on the RHS may be written as $E[Y \mid Z=1] - E[Y \mid Z=0]$. Next, it can be seen that $\pi_{02}+\pi_{12} = P[D_i(1) \geq 2 > D_i(0)]$. Dividing the numerator and denominator by the first stage, we obtain that the RHS is equal to ACR/w_J. This shows that the assumed values of the unknown potential outcomes are consistent with the observed moments and assumptions, and they attain the upper bound on the CCE provided in Proposition 1.

H Identification of CCE with Additional Boundedness Assumptions

This section considers bounds for the CCE under Assumptions 1-5, when either the unit causal effects or the support of the outcome are known to be bounded.

H.1 Bounded Unit Causal Effects with Homogeneity

We first consider the additional assumption that the unit causal effects are bounded from above, i.e., $\beta_j \in [0, \overline{\beta}]$. We are interested in solving the following linear program:

$$\underset{\{\beta_j\}}{\text{Maximize/Minimize}} \quad \sum_{j=1}^{J} \beta_j \qquad \textbf{LP.4}$$

$$\text{subject to} \quad \sum_{j=1}^{J} \beta_j w_j = ACR \qquad \text{(C.1)}$$
$$\beta_j \geq 0 \qquad \text{(C.2)}$$
$$\beta_j \leq \overline{\beta} \qquad \text{(C.3)}$$

For ease of notation, we assume that the weights are monotonically declining (Assumption 7).

Beginning with the upper bound, define $j_u = \max\{1, ..., J\}$ such that $\sum_{j=j_u}^{J} w_j \geq \frac{ACR}{\overline{\beta}}$. Note that j_u is well-defined under our assumptions; if not, it would imply $ACR > \overline{\beta} \sum_{j=1}^{J} w_j = \overline{\beta}$. Because the ACR is a weighted average of unit causal effects, all of which are bounded above by $\overline{\beta}$, this is impossible. Therefore, j_u must be well-defined. Next, consider the case in which $j_u = J$, which is obtained when $w_J \geq \frac{ACR}{\overline{\beta}}$, or equivalently, $\overline{\beta} \geq \frac{ACR}{w_J}$. In this case, the vector of unit-causal effects that attain the bounds in the absence of C.3 (as shown in Proposition 1) are still feasible, since $\overline{\beta} \geq \frac{ACR}{w_J} = \beta_J$ and $\overline{\beta} > \frac{ACR}{w_1} = \beta_1$. Therefore, in the case that $j_u = J$, C.3 does not bind and **LP.4** collapses into the case studied in Proposition 1.

The other possibility is that $j_u \in \{1, ..., J-1\}$. To find an upper bound for the CCE in this case, we will first show that the CCE is maximized when $\beta_j = \overline{\beta}$ for all $j > j_u$ and $\beta_j = 0$ for all $j < j_u$.

To prove $\beta_j = \overline{\beta}$ for all $j > j_u$, assume towards a contradiction that the CCE is maximized subject to C.1-C.3 by a vector of unit causal effects $\beta = (\beta_1, ..., \beta_J)$ where $\beta_{k_1} < \overline{\beta}$ for some $k_1 > j_u$. Observe that, by construction of j_u, we have:

$$\begin{aligned} \overline{\beta} &< \frac{ACR}{\sum_{j=j_u+1}^{J} w_j} \\ \implies \sum_{j=j_u+1}^{J} \overline{\beta} w_j &< ACR = \sum_{j=1}^{J} \beta_j w_j \end{aligned}$$

where the last equality follows from C.1. Because $\beta_j \leq \overline{\beta}$ for all j, this implies that there must exist a $k \leq j_u$ such that $w_k > 0$ and $\beta_k > 0$. We assume without loss of generality that there is only one such non-zero unit-causal effect for $k \leq j_u$.[14] Consider a new vector of unit causal effects $\beta^* = (\beta_1^*, ..., \beta_J^*)$ equal to β, except that $\beta_k^* = \beta_k - (\overline{\beta} - \beta_{k_1}) \frac{w_{k_1}}{w_k}$ and $\beta_{k_1}^* = \overline{\beta}$. Note that β^* satisfies C.1-C.3. Comparing the CCE under β^* and β:

$$\sum_{j=1}^{J} \beta_j^* - \sum_{j=1}^{J} \beta_j = \left(\beta_k^* - \beta_k\right) + \left(\beta_{k_1}^* - \beta_{k_1}\right) = -\left(\overline{\beta} - \beta_{k_1}\right) \frac{w_{k_1}}{w_k} + \left(\overline{\beta} - \beta_{k_1}\right) = \left(\overline{\beta} - \beta_{k_1}\right)\left(1 - \frac{w_{k_1}}{w_k}\right) > 0$$

where the final inequality follows since $\frac{w_{k_1}}{w_k} < 1$. Thus, the initial vector of unit treatment effects must not be the solution to the constrained maximization problem, proving that the CCE is maximized only when $\beta_j = \overline{\beta}$ for all $j > j_u$.

To prove $\beta_j = 0$ for all $j < j_u$, assume towards a contradiction that the CCE is maximized by a vector of unit causal effects $\beta = (\beta_1, ..., \beta_J)$ where $\beta_{k_2} > 0$ for some $k_2 < j_u$. Because β maximizes the CCE, we know from above that $\beta_j = \overline{\beta}$ for all $j > j_u$. Consider a new vector of unit treatment effects, $\beta^* = (\beta_1^*, ..., \beta_J^*)$, identical to β except that the specified non-zero effect has been set to zero, $\beta_{k_2}^* = 0$, and β_{j_u} has been increased to maintain feasibility with respect to the ACR constraint: $\beta_{j_u}^* = \beta_{j_u} + \beta_{k_2} w_{k_2} / w_{j'}$. For it to be feasible to increase β_{j_u} in this way, we need that $\beta_{j_u}^* \leq \overline{\beta}$. This follows from the construction of j_u:

$$\frac{ACR}{\sum_{j=j_u}^{J} w_j} \leq \overline{\beta} \iff \frac{ACR - \sum_{j=j_u+1}^{J} w_j \overline{\beta}}{w_{j_u}} \leq \overline{\beta}$$

[14] This is without loss of generality because we can always re-define the problem with β_k being a weighted average of all non-zero unit causal effects. For instance, if $\beta_p, \beta_q > 0$ for $p, q \leq j_u$, we could re-define $\beta_k = \frac{\beta_p w_p + \beta_q w_q}{w_p + w_q}$ and $w_k = w_p + w_q$.

To see this, note that

$$\frac{ACR - \sum_{j=j_u+1}^{J} w_j \overline{\beta}}{w_{j_u}} = \frac{\sum_{j=1}^{j_u} w_j \beta_j}{w_{j_u}} = \frac{w_{j_u}\beta_{j_u} + w_{k_2}\beta_{k_2}}{w_{j_u}} = \beta_{j_u} + \frac{w_{k_2}\beta_{k_2}}{w_{j_u}} = \beta^*_{j_u} \leq \overline{\beta}$$

where the first equality follows from the ACR constraint, the second equality follows from our assumption that β_k is the only non-zero unit causal effect for $k < j_u$, the fourth equality follows from our definition of $\beta^*_{j_u}$. But with these changes, the value of the objective function is higher than that of the original vector of unit effects:

$$\sum \beta^*_j - \sum \beta_j = \beta_k \frac{w_k}{w_{j_u}} - \beta_k > 0,$$

where the last inequality follows from the fact that $\frac{w_k}{w_{j_u}} > 1$. Thus, the initial vector of unit treatment effects must not be the solution to the constrained maximization problem. Together, this implies that the CCE is maximized only when $\beta_j = \overline{\beta}$ for all $j > j_u$ and $\beta_j = 0$ for all $j < j_u$. The constraint C.1 pins down the value for β_{j_u} and implies that the upper bound on the CCE is

$$\frac{ACR}{w_{j_u}} - \overline{\beta}\ \frac{\sum_{j=j_u+1}^{J} w_j}{w_{j_u}} + (J - j_u)\overline{\beta}$$

Turning to the lower bound, define $j_l = \min\{1, ..., J\}$ such that $\sum_{j=1}^{j_l} w_j \geq \frac{ACR}{\overline{\beta}}$. Note that j_l is well-defined under our assumptions; if not, it would imply $ACR > \overline{\beta} \sum_{j=1}^{J} w_j = \overline{\beta}$. Because the ACR is a weighted average of unit causal effects, all of which are bounded above by $\overline{\beta}$, this is impossible. Therefore, j_l must be well-defined. Next, consider the case in which $j_l = 1$, which is obtained when $w_1 \geq \frac{ACR}{\overline{\beta}}$, or equivalently, $\overline{\beta} \geq \frac{ACR}{w_1}$. In this case, the vector of unit-causal effects that attain the lower bounds in the absence of C.3 (as shown in Proposition 1) is still feasible, since $\overline{\beta} \geq \frac{ACR}{w_1} = \beta_1$. Therefore, in the case that $j_l = 1$, C.3 does not bind and **LP.4** collapses into the case studied in Proposition 1. The other possibility is that $j_l \in \{2, ..., J\}$. The CCE is minimized when $\beta_j = \overline{\beta}$ for all $j < j_l$ and $\beta_j = 0$ for all $j > j_l$. The constraint C.1 pins down the value for β_{j_l} and implies that the lower bound on the CCE is

$$\frac{ACR}{w_{j_l}} - \overline{\beta}\ \frac{\sum_{j=1}^{j_l-1} w_j}{w_{j_l}} + (j_l - 1)\overline{\beta}$$

The derivation for the lower bound is analogous to the proof for the upper bound.

H.2 Bounded Outcomes with Homogeneity

In this subsection, we consider bounds on the CCE under Assumptions 1-5 in the case that the outcome has bounded support between known values, $Y_i \in \left[\underline{Y}, \overline{Y}\right]$.

Let $P[D_i(0) = m, D_i(1) = n] \equiv \pi_{mn}$ and $E[Y_i(j)|D_i(0) = m, D_i(1) = n] \equiv Y_{mn}(j)$ for $j, m, n \in \{0, 1, ..., J\}$. In Appendix Section G, we establish how the observed moments constrain the relationship between the outcome and treatment dosage as well as the relationship between the unobserved joint distribution of $D_i(0)$ and $D_i(1)$ and the observed moments under Assumptions 1-3. Here, we incorporate the moment conditions as constraints in the following optimization problem.

For any $m, n \in \{0, 1, ..., J\}$ and $n > m$:

$$\underset{\{Y_{jk}(d), \pi_{jk}\}_{j,k,d=0}^{J}}{\text{Maximize/Minimize}} \quad Y_{mn}(J) - Y_{mn}(0) \qquad \textbf{LP.5}$$

$$\text{subject to} \quad Y_{mn}(j) - Y_{mn}(j-1) \geq 0 \qquad \text{(C.1)}$$

$$\sum_{j=k}^{J} \pi_{kj} = P[D = k \mid Z = 0] \text{ for all } k \in \{0, 1, ..., J\} \qquad \text{(C.2)}$$

$$\sum_{j=0}^{k} \pi_{jk} = P[D = k \mid Z = 1] \text{ for all } k \in \{0, 1, ..., J\} \qquad \text{(C.3)}$$

$$\sum_{j=k}^{J} \pi_{kj} Y_{kj}(k) = E[Y \mid Z = 0, D = k] P[D = k \mid Z = 0] \text{ for all } k \in \{0, 1, ..., J\} \qquad \text{(C.4)}$$

$$\sum_{j=0}^{k} \pi_{jk} Y_{jk}(k) = E[Y \mid Z = 1, D = k] P[D = k \mid Z = 1] \text{ for all } k \in \{0, 1, ..., J\} \qquad \text{(C.5)}$$

$$Y_{jk}(d) \in \left[\underline{Y}, \overline{Y}\right] \text{ for all } j, k, d \in \{0, 1, ..., J\} \qquad \text{(C.6)}$$

Absent C.6, **LP.5** simplifies to the case studied in Proposition 1. As shown in Appendix G, the bounds in Proposition 1 are sharp under C.1 through C.5 and the data available to the researcher. Adding C.6 allows us to tighten the bounds described in that Proposition. In particular, one implication of C.6 is that the magnitude of each unit causal effect does not exceed $\overline{Y} - \underline{Y}$. Hence, the bounds described in Appendix H.1 are valid for **LP.5**, with $\overline{\beta} = \overline{Y} - \underline{Y}$. However, this restriction does not necessarily exploit all of the information contained in C.6; C.6 also provides information concerning the level of the potential outcomes, which is not reflected in constraints relating to the magnitude of the treatment effect. As such, directly solving **LP.5** may yield tighter bounds compared to the bounds described in Appendix H.1.

I Identification of Related Parameters

Thus far, we have focused on identification of the effect of moving from the treatment's minimum intensity to the treatment's maximum intensity. In this section, we explore identification of a broader class of parameters. Specifically, we consider identification of the effect among compliers of moving from treatment intensity j_1 to j_2 where $j_1, j_2 \in \{0, 1, ..J\}$ and $j_1 < j_2$. We refer to this as the Complier Effect (CE) for j_1 to j_2: $CE(j_1, j_2) = E[Y_i(j_2) - Y_i(j_1) | D_i(1) > D_i(0)]$. Under Assumption 4, the unit-causal effects are homogeneous across different types of compliance subgroups. Specifically, Assumption 4 requires that $E[Y_i(j_2) - Y_i(j_1) | D_i(1) > D_i(0)] = E[Y_i(j_2) - Y_i(j_1) | D_i(1) = j_2, D_i(0) = j_1]$. Therefore, we may re-write $CE(j_1, j_2)$ as $CE(j_1, j_2) = E[Y_i(j_2) - Y_i(j_1) | D_i(1) = j_2, D_i(0) = j_1] = \sum_{j=j_1+1}^{j_2} \beta_j$. In the special case where $j_1 = 0$ and $j_2 = J$,

$$CE(0, J) = CCE = E[Y_i(J) - Y_i(0) | D_i(1) = J, D_i(0) = 0] = \sum_{j=1}^{J} \beta_j$$

I.1 Proposition 1

We begin by asking what may be identified for $CE(j_1, j_2)$ under Assumptions 1-5. As in Section 4, we cast the problem of identifying $CE(j_1, j_2)$ as the following linear optimization problem:

$$\underset{\{\beta_j\}}{\text{Maximize/Minimize}} \quad \sum_{j=j_1+1}^{j_2} \beta_j \qquad \textbf{LP.6}$$

$$\text{subject to} \quad \sum_{j=1}^{J} \beta_j w_j = ACR \qquad \text{(C.1)}$$

$$\beta_j \geq 0 \qquad \forall\, j = 1, ..., J \qquad \text{(C.2)}$$

In this section, we extend Proposition 1 and show the following:[15]

Proposition 1*: Let $\underline{j}(j_1, j_2) = \arg\min_{j \in \{j_1+1,...,j_2\}} \{w_j\}$ and $\overline{j}(j_1, j_2) = \arg\max_{j \in \{j_1+1,...,j_2\}} \{w_j\}$. Under Assumptions 1-5, the following sharp bounds hold:

$$CE(0, J) = CCE \in \left[\frac{ACR}{w_{\overline{j}(0,J)}}, \frac{ACR}{w_{\underline{j}(0,J)}} \right]$$

and for $(j_1, j_2) \neq (0, J)$

$$CE(j_1, j_2) \in \left[0, \frac{ACR}{w_{\underline{j}(j_1,j_2)}} \right]$$

.

Proof: We begin by showing that $CE(j_1, j_2)$ is minimized at 0, for any $(j_1, j_2) \neq (0, J)$. Recall that, under Assumption 5, $\beta_j \geq 0$. This implies that $CE(j_1, j_2)$ is always bounded below at 0 for all j_1, j_2: $CE(j_1, j_2) = \sum_{j=j_1+1}^{j_2} \beta_j \geq 0$. Next, we show that for any $(j_1, j_2) \neq (0, J)$, we cannot rule out that $CE(j_1, j_2)$ takes on a value 0. Consider the case where $j_1 = 0$. This implies that $j_2 \leq J - 1$. Define k such that $j_2 < k \leq J$. Then, constraint (C.1) may be satisfied with $\beta_k = \frac{ACR}{w_k}$ and $\beta_j = 0$ for all $j \neq k$. This implies that $\sum_{j=j_1+1}^{j_2} \beta_j = 0$ and shows that $CE(j_1, j_2)$ is minimized at 0. Similarly, consider the case where $j_1 \neq 0$. Define k such that $0 \leq k < j_1$. Then, constraint (C.1) may be satisfied with $\beta_k = \frac{ACR}{w_k}$ and $\beta_j = 0$ for all $j \neq k$. This shows that the $CE(j_1, j_2)$ is minimized at 0, for any $(j_1, j_2) \neq (0, J)$. This shows that $CE(j_1, j_2)$ is minimized at 0 for any $(j_1, j_2) \neq (0, J)$.

Next, we show that $CE(j_1, j_2)$ is maximized at $\frac{ACR}{w_{\underline{j}(j_1,j_2)}}$ where $\underline{j}(j_1, j_2) = \arg\min_{j \in \{j_1+1,...,j_2\}} \{w_j\}$. To establish that $CE(j_1, j_2)$ is bounded above by $\frac{ACR}{w_{\underline{j}(j_1,j_2)}}$, we will first show that the $CE(j_1, j_2)$ is maximized if and only if $\beta_j = 0$ for all $j < j_1$ and $j > j_2$ and then show that $CE(j_1, j_2)$ is maximized when $\beta_{\underline{j}(j_1,j_2)}$ is the only non-zero unit causal effect, i.e., $\beta_j = 0$ for all $j \neq \underline{j}(j_1, j_2)$.

Proceeding by contradiction to show the first step, suppose instead that $CE(j_1, j_2)$ was maximized by a vector of unit treatment effects $(\beta_1, ..., \beta_J)$ with $\beta_k > 0$ for some $k < j_1$ or $k > j_2$. Consider a new vector of unit treatment effects, $(\beta_1^*, ..., \beta_J^*)$ identical to $(\beta_1, ..., \beta_J)$, except that the specified non-zero effect has been set to zero, $\beta_k^* = 0$, and the unit effect corresponding to the smallest weight, $\beta_{\underline{j}(j_1,j_2)}$, has been increased to maintain feasibility with respect to the ACR constraint: $\beta^*_{\underline{j}(j_1,j_2)} = \beta_{\underline{j}(j_1,j_2)} + \beta_k\, w_k / w_{\underline{j}(j_1,j_2)}$. But, with these changes, the value of the objective function is larger

[15] In this proposition, we assume $w_k > 0$ for at least one $k \leq j_1$ or $k > j_2$. If $w_k = 0$ for all $k \leq j_1$ and $k > j_2$, then the lower bound on $CE(j_1, j_2)$ would be the same as the lower bound on the CCE.

than that of the original vector of unit effects:

$$\sum_{j=j_1+1}^{j_2} \beta_j^* - \sum_{j=j_1+1}^{j_2} \beta_j = \beta_k \frac{w_k}{w_{\underline{j}(j_1,j_2)}} > 0$$

where the final inequality follows by the assumption that $\beta_k > 0$.

Proceeding similarly by contradiction to show the second step, suppose instead that $CE(j_1, j_2)$ was maximized by a vector of unit treatment effects $(\beta_1, ..., \beta_J)$ with $\beta_k > 0$ for some $k \neq \underline{j}(j_1, j_2)$ and $j_1 \leq k \leq j_2$. Consider a new vector of unit treatment effects, $(\beta_1^*, ..., \beta_J^*)$ identical to $(\beta_1, ..., \beta_J)$, except that the specified non-zero effect has been set to zero, $\beta_k^* = 0$, and the unit effect corresponding to the smallest weight, $\beta_{\underline{j}(j_1,j_2)}$, has been increased to maintain feasibility with respect to the ACR constraint: $\beta^*_{\underline{j}(j_1,j_2)} = \beta_{\underline{j}(j_1,j_2)} + \beta_k \, w_k / w_{\underline{j}(j_1,j_2)}$. But, with these changes, the value of the objective function is larger than that of the original vector of unit effects:

$$\sum_{j=j_1+1}^{j_2} \beta_j^* - \sum_{j=j_1+1}^{j_2} \beta_j = \beta_k \frac{w_k}{w_{\underline{j}(j_1,j_2)}} - \beta_k > 0,$$

where the last inequality follows from the fact that $\beta_k > 0$ and the definition of $\underline{j}(j_1, j_2)$. Thus, the initial vector of unit treatment effects must not be the solution to the constrained maximization problem, proving the contradiction. This shows that $CE(j_1, j_2)$ is maximized at $\frac{ACR}{w_{\underline{j}(j_1,j_2)}}$ for any $(j_1, j_2) \neq (0, J)$. Finally, we note that one can use the same argument as in Appendix Section G to establish that these bounds are sharp.

I.2 Proposition 2

We consider identification for $CE(j_1, j_2)$ under Assumptions 1-7. As in Section 4, we may cast the problem of identifying $CE(j_1, j_2)$ as the following linear optimization problem:

$$\underset{\{\beta_j\}}{\text{Maximize/Minimize}} \quad \sum_{j=j_1+1}^{j_2} \beta_j \qquad \textbf{LP.7}$$

$$\text{subject to} \quad \sum_{j=1}^{J} \beta_j w_j = ACR \qquad \text{(C.1)}$$

$$\beta_j \geq 0 \qquad \forall \, j = 1, ..., J \qquad \text{(C.2)}$$

$$\beta_j \geq \beta_{j+1} \qquad \forall \, j = 1, ..., J-1 \qquad \text{(C.3)}$$

In this section, we extend Proposition 2 to show the following.

Proposition 2*: Under Assumptions 1-7, the following sharp bounds hold: $CCE \in \left[\frac{ACR}{w_1}, \, ACR \times J\right]$.

(a) When $j_1, j_2 \in \{0, 1, ..., J\}$ where $j_2 > j_1$, $CE(j_1, j_2) \leq \frac{ACR}{\sum_{j=1}^{j_2} w_j} \times (j_2 - j_1)$

(b) When $j_1, j_2 \in \{1, ..., J\}$ where $j_2 > j_1$, $CE(j_1, j_2) \geq 0$

(c) When $j_1 = 0$ and $j_2 \in \{1, ..., J\}$ where $j_2 \geq \frac{1}{w_1}$, $CE(0, j_2) \geq \frac{ACR}{w_1}$

(d) When $j_1 = 0$ and $j_2 \in \{1, ..., J\}$ where $j_2 < \frac{1}{w_1}$, $CE(0, j_2) \geq ACR \times j_2$

Proof: We begin with the proof of Proposition 2*(a) and show that the $CE(j_1, j_2)$ is maximized at $\frac{ACR}{\sum_{j=1}^{j_2} w_j} \times (j_2 - j_1)$.

We proceed in three steps. First, we show that $CE(j_1, j_2)$ is maximized when $\beta_j = 0$ for all $j > j_2$. In the second step, we show that the $CE(j_1, j_2)$ is maximized when $\beta_{j_1+1} = ... = \beta_{j_2}$. Finally, we show that it is only maximized when $\beta_1 = ... = \beta_{j_1+1}$.

Proceeding by contradiction for the first step, suppose that $CE(j_1, j_2)$ is maximized subject to constraints (C.1)-(C.3) by a vector of unit effects $\beta = (\beta_1, \beta_2, ..., \beta_J)$ with $\beta_k > 0$ for some $k > j_2$. Consider a new vector of unit treatment effects, $\tilde{\beta} = (\tilde{\beta}_1, \tilde{\beta}_2, ..., \tilde{\beta}_J)$, where $\tilde{\beta}_j = \beta_j + \frac{w_k \beta_k}{w_j j_2}$ for all $j \leq j_2$ and $\tilde{\beta}_k = 0$. To see that $\tilde{\beta}$ satisfies (C.1),

$$\sum_{j=1}^{J} \tilde{\beta}_j w_j = \sum_{j=1}^{j_2} \tilde{\beta}_j w_j = \sum_{j=1}^{j_2} \left(\beta_j + \frac{w_k \beta_k}{w_j j_2}\right) w_j = \sum_{j=1}^{j_2} \beta_j w_j + \sum_{j=1}^{j_2} \frac{w_k \beta_k}{j_2} = \sum_{j=1}^{j_2} \beta_j w_j + w_k \beta_k = ACR$$

where the first equality follows from the fact that $\tilde{\beta}_j = 0$ for all $j > j_2$ and the final equality follows by the assumption that the vector of unit effects β satisfies C.1. Additionally, $\tilde{\beta}$ satisfies C.2 since $\beta_j \geq 0$ for all j implies that $\tilde{\beta}_j = \beta_j + \frac{w_k \beta_k}{w_j j_2} \geq 0$. Finally, C.3 is satisfied since $\beta_j \geq \beta_{j+1}$ implies that $\beta_j + \sum_{j=1}^{j_2} \frac{w_k \beta_k}{w_j j_2} \geq \beta_{j+1} + \sum_{j=1}^{j_2} \frac{w_k \beta_k}{w_j j_2} \implies \tilde{\beta}_j \geq \tilde{\beta}_{j+1}$. Finally, comparing the objective functions under the two vectors of unit effects,

$$\sum_{j=j_1+1}^{j_2} \tilde{\beta}_j - \sum_{j=j_1+1}^{j_2} \beta_j = \sum_{j=j_1+1}^{j_2} \frac{w_k \beta_k}{w_j j_2} \geq 0$$

where the last inequality follows by the assumption that $\beta_k > 0$. This shows that $CE(j_1, j_2)$ is maximized when $\beta_j = 0$ for all $j > j_2$.

Second, we show that the $CE(j_1, j_2)$ is maximized when $\beta_{j_1+1} = ... = \beta_{j_2}$. Proceeding by contradiction, suppose that the $CE(j_1, j_2)$ is maximized by a vector of unit treatment effects $(\beta_1, \beta_2, ..., \beta_J)$ with $\beta_k > \beta_{k+1}$ for some $k \in \{j_1 + 1, ..., j_2 - 1\}$ (concavity rules out the reverse ordering). Consider a new vector of unit treatment effects, $(\tilde{\beta}_1, ..., \tilde{\beta}_J)$, which is identical to the first vector, except that $\tilde{\beta}_k = \beta_k - \frac{(\beta_k - \beta_{k+1}) w_{k+1}}{w_k + w_{k+1}}$ and $\tilde{\beta}_{k+1} = \beta_{k+1} + \frac{w_k (\beta_k - \beta_{k+1})}{w_k + w_{k+1}}$, so that $\tilde{\beta}_k = \tilde{\beta}_{k+1}$. Note that $(\tilde{\beta}_1, ..., \tilde{\beta}_J)$ is a feasible solution vector since it satisfies C.1-C.3. But, with these changes, the value of the objective function is higher than that of the original vector of the unit effects:

$$\sum_{j=j_1+1}^{j_2} \tilde{\beta}_j - \sum_{j=j_1+1}^{j_2} \beta_j = \frac{(\beta_{k+1} - \beta_k)(w_{k+1} - w_k)}{w_{k+1} + w_k} > 0,$$

where the last inequality follows from the assumption that $\beta_k > \beta_{k+1}$ (Assumption 6) and that $w_k > w_{k+1}$ (Assumption 7). Thus, the initial vector of unit treatment effects must not be the solution to the constrained maximization problem, proving the contradiction. This shows that the $CE(j_1, j_2)$ is maximized when $\beta_{j_1+1} = ... = \beta_{j_2}$.

Finally, we show that $CE(j_1, j_2)$ is maximized when $\beta_1 = ... = \beta_{j_1+1}$. Proceeding by contradiction, suppose that $CE(j_1, j_2)$ is maximized by a vector of unit effects $\beta = (\beta_1, \beta_2, ...\beta_J)$ where $\beta_j = 0$ for all $j > j_2$ and $\beta_j = \beta_{j+1}$ for all $j \leq j_2$ except $j = k$ for some $k \in \{1, 2, ..., j_1\}$. This implies that $\beta_k > \beta_{k+1}$ (concavity rules out the reverse ordering). We can re-write (C.1) under the vector β as:

$$\sum_{j=1}^{J} \beta_j w_j = \sum_{j=1}^{j_2} \beta_j w_j = \sum_{j=1}^{k} \beta_j w_j + \sum_{j=k+1}^{j_2} \beta_j w_j = \sum_{j=1}^{k} \beta_k w_j + \sum_{j=k+1}^{j_2} \beta_{j_1+1} w_j = ACR$$

where the first equality follows the assumption that $\beta_j = 0$ for all $j > j_2$ and the third equality follows from the assumption that $\beta_j = \beta_k$ for all $j \leq k$ and $\beta_j = \beta_{j_1+1}$ for all $j \in \{k+1, ..., j_2\}$.

Define $\tilde{w}_k = \sum_{j=1}^{k} w_j$ and $\tilde{w}_{j_1+1} = \sum_{j=k+1}^{j_2} w_j$. Then, $\beta_k \tilde{w}_k + \beta_{j_1+1} \tilde{w}_{j_1+1} = ACR$.

Consider a new vector of unit treatment effects, $(\tilde{\beta}_1, ..., \tilde{\beta}_J)$, with $\tilde{\beta}_k = \beta_k - \frac{(\beta_k - \beta_{j_1+1}) \tilde{w}_{j_1+1}}{\tilde{w}_k + \tilde{w}_{j_1+1}}$ and $\tilde{\beta}_{j_1+1} = \beta_{j_1+1} + \frac{\tilde{w}_k (\beta_k - \beta_{j_1+1})}{\tilde{w}_k + \tilde{w}_{j_1+1}}$, so that $\tilde{\beta}_k = \tilde{\beta}_{j_1+1}$. Note that $(\tilde{\beta}_1, ..., \tilde{\beta}_J)$ is a feasible unit vector since it satisfies C.1-C.3. But, with these

changes, the value of the objective function is higher than that of the original vector of the unit effects:

$$\sum_{j=j_1+1}^{j_2} \tilde{\beta}_j - \sum_{j=j_1+1}^{j_2} \beta_j = \frac{\tilde{w}_k(\beta_k - \beta_{j_1+1})}{\tilde{w}_k + \tilde{w}_{j_1+1}} > 0,$$

where the last inequality follows from the assumption that $\beta_k > \beta_{k+1} = \beta_{j_1+1}$. This shows that $CE(j_1, j_2)$ is maximized when $\beta_1 = ... = \beta_{j_2}$.

Finally, using the ACR constraint, this implies that

$$\sum_{j=1}^{J} \beta_j w_j = \sum_{j=1}^{j_2} \beta_j w_j = ACR \implies \beta_j = \frac{ACR}{\sum_{j=1}^{j_2} w_j}$$

Using the above, we can re-write $CE(j_1, j_2)$ as

$$CE(j_1, j_2) = \sum_{j=j_1+1}^{j_2} \beta_j = \sum_{j=j_1+1}^{j_2} \frac{ACR}{\sum_{j=1}^{j_2} w_j} = \frac{ACR(j_2 - j_1)}{\sum_{j=1}^{j_2} w_j}$$

Next, we prove Proposition 2*(b) and show that $CE(j_1, j_2)$ is minimized at 0 for any $j_1 > 0$. Recall that under Assumption 5, $\beta_j \geq 0$ for all j. This implies that $CE(j_1, j_2)$ is bounded below at 0 for any j_1, j_2, $CE(j_1, j_2) = \sum_{j=j_1+1}^{j_2} \beta_j \geq 0$. Next, we show that for any $j_1 > 0$, we cannot rule out that $CE(j_1, j_2)$ takes on a value of 0. The constraint (C.1) may be satisfied with $\beta_1 = \frac{ACR}{w_1}$ and $\beta_j = 0$ for all $j > 1$. Note that this vector of unit effects $(\beta_1, \beta_2, ..., \beta_J) = \left(\frac{ACR}{w_1}, 0, ..., 0\right)$ also satisfies Constraint (C.2) and (C.3) since $\beta_1 = \frac{ACR}{w_1} \geq \beta_j = 0$ for any $j > 1$. Under this vector of unit effects, $CE(j_1, j_2) = \sum_{j=j_1+1}^{j_2} \beta_j = 0$. This shows that $CE(j_1, j_2)$ is minimized at 0 for any $j_1 > 0$.

Next, we prove Proposition 2*(c) and show that the lower bound on $CE(0, j_2)$ is given by ACR/w_1 if $j_2 \geq \frac{1}{w_1}$. Suppose that $j_2 \geq \frac{1}{w_1}$. To show that $CE(0, j_2) \geq ACR/w_1$, we will first show that the $CE(0, j_2)$ is minimized when $\beta_j = 0 \; \forall \; j > 1$. Proceeding by contradiction, suppose that the $CE(0, j_2)$ is minimized by a vector of unit treatment effects $(\beta_1, ..., \beta_J)$ with $\beta_j > 0$ for some $j > 1$. Let k denote the highest intensity non-zero unit effect, more precisely $k = \max_j \{2, ..., J \mid \beta_j > 0\}$. Along with Assumption 6, this implies $\beta_1 \geq \beta_2 \geq ... \geq \beta_k > 0$. Consider a new vector of unit treatment effects, $(\tilde{\beta}_1, ..., \tilde{\beta}_J)$, which is identical to the first vector, except that β_k has been set to zero, $\tilde{\beta}_k = 0$, and the first unit effect has been increased to maintain feasibility with respect to the ACR constraint: $\tilde{\beta}_1 = \beta_1 + \beta_k w_k / w_1$. But, with these changes, the value of the objective function is lower than that of the original vector of unit effects:

$$\sum_{j=1}^{j_2} \tilde{\beta}_j - \sum_{j=1}^{j_2} \beta_j = \beta_k \frac{w_k}{w_1} - \beta_k < 0,$$

where the last inequality follows from Assumptions 5 and 7. Thus, the initial vector of unit treatment effects must not be the solution to the constrained minimization problem, proving the contradiction. Finally, to prove the result, note that when $\beta_j = 0 \; \forall \; j > 1$, it follows that $CE(0, j_2) = \sum_j \beta_j = \beta_1 = ACR/w_1$, where the last equality follows from the definition of the ACR. Next, we show that $CE(0, j_2) = ACR/w_1$ would not be a lower bound if $j_2 < \frac{1}{w_1}$. Proceeding by contradiction, suppose that the $CE(0, j_2)$ is minimized by a vector of unit treatment effects $\beta = (\beta_1, ..., \beta_J)$ with $\beta_1 = ACR/w_1$ and $\beta_j = 0$ for all $j > 1$. Consider a new vector of unit treatment effects, $\tilde{\beta} = (\tilde{\beta}_1, ..., \tilde{\beta}_J)$, where $\tilde{\beta}_j = \tilde{\beta}_{j+1}$ for all $j \in \{1, ..., J-1\}$. The ACR constraint implies that $\tilde{\beta}_j = ACR$ for all j. Comparing the objective

function under β and $\tilde{\beta}$,

$$\sum_{j=1}^{j_2} \tilde{\beta}_j - \sum_{j=1}^{j_2} \beta_j = ACR \times j_2 - \frac{ACR}{w_1} < 0$$

where the last inequality follows from the assumption that $j_2 < \frac{1}{w_1}$. This shows that $CE(0, j_2) = ACR/w_1$ would not be a lower bound if $j_2 < \frac{1}{w_1}$.

Finally, we prove Proposition 2*(d) and show that $CE(0, j_2)$ is minimized at $ACR \times j_2$ if $j_2 < \frac{1}{w_1}$. To show that $CE(0, j_2) \geq ACR \times j_2$, we will first show that $CE(0, j_2)$ is minimized when $\beta_1 = \beta_2 = ... = \beta_J$. Proceeding by contradiction, suppose that $CE(0, j_2)$ is minimized by a vector of unit effects $\beta = (\beta_1, \beta_2, ...\beta_J)$ where $\beta_j = \beta_{j+1}$ for all j except $j = k$ for some $k \in \{1, 2, ..., J-1\}$. This implies that $\beta_k > \beta_{k+1}$ (concavity rules out the reverse ordering). We can re-write (C.1) under the vector β as:

$$\sum_{j=1}^{J} \beta_j w_j = \sum_{j=1}^{k} \beta_j w_j + \sum_{j=k+1}^{J} \beta_j w_j = \sum_{j=1}^{k} \beta_k w_j + \sum_{j=k+1}^{J} \beta_{k+1} w_j = ACR$$

where the second equality follows from the assumption that $\beta_j = \beta_k$ for all $j \leq k$ and $\beta_j = \beta_{k+1}$ for all $j \in \{k+1, ..., J\}$. Define $\tilde{w}_k = \sum_{j=1}^{k} w_j$ and $\tilde{w}_{k+1} = \sum_{j=k+1}^{J} w_j$. Then, $\beta_k \tilde{w}_k + \beta_{k+1} \tilde{w}_{k+1} = ACR$.

Consider a new vector of unit treatment effects, $(\tilde{\beta}_1, ..., \tilde{\beta}_J)$, with $\tilde{\beta}_k = \beta_k - \frac{(\beta_k - \beta_{k+1})\tilde{w}_{k+1}}{\tilde{w}_k + \tilde{w}_{k+1}}$ and $\tilde{\beta}_{k+1} = \beta_{k+1} + \frac{\tilde{w}_k(\beta_k - \beta_{k+1})}{\tilde{w}_k + \tilde{w}_{k+1}}$, so that $\tilde{\beta}_k = \tilde{\beta}_{k+1}$. Note that $(\tilde{\beta}_1, ..., \tilde{\beta}_J)$ is a feasible unit vector since it satisfies C.1-C.3. Next, we show that with these changes, the value of the objective function is smaller than that of the original vector of the unit effects. If $k \geq j_2$,

$$\sum_{j=1}^{j_2} \tilde{\beta}_j - \sum_{j=1}^{j_2} \beta_j = -j_2 \times \frac{(\beta_k - \beta_{k+1})\tilde{w}_{k+1}}{\tilde{w}_k + \tilde{w}_{k+1}} < 0$$

where the last inequality follows from the assumption that $\beta_k > \beta_{k+1}$. If $k < j_2$,

$$\begin{aligned}
\sum_{j=1}^{j_2} \tilde{\beta}_j - \sum_{j=1}^{j_2} \beta_j &= \sum_{j=1}^{k} \left(\tilde{\beta}_j - \beta_j\right) + \sum_{j=k+1}^{j_2} \left(\tilde{\beta}_j - \beta_j\right) \\
&= -k \times \frac{(\beta_k - \beta_{k+1})\tilde{w}_{k+1}}{\tilde{w}_k + \tilde{w}_{k+1}} + (j_2 - k)\frac{\tilde{w}_k(\beta_k - \beta_{k+1})}{\tilde{w}_k + \tilde{w}_{k+1}} \\
&= \left(\frac{\beta_k - \beta_{k+1}}{\tilde{w}_k + \tilde{w}_{k+1}}\right)\left(j_2 \tilde{w}_k - k(\tilde{w}_k + \tilde{w}_{k+1})\right) \\
&\leq \left(\frac{\beta_k - \beta_{k+1}}{\tilde{w}_k + \tilde{w}_{k+1}}\right)\left(j_2 w_1 - k\right) < \left(\frac{\beta_k - \beta_{k+1}}{\tilde{w}_k + \tilde{w}_{k+1}}\right)\left(1 - k\right) \leq 0
\end{aligned}$$

where the fourth line follows from the fact that $w_1 > w_j$ for all $j > 1$ and $\tilde{w}_k + \tilde{w}_{k+1} = 1$, the fifth line follows from the fact that $j_2 < \frac{1}{w_1} \implies j_2 w_1 < 1$, and the final inequality follows from the fact that $\beta_k > \beta_{k+1}$ by assumption and $k \geq 1$.

This shows that when $j_2 < \frac{1}{w_1}$, $CE(0, j_2)$ is minimized when $\beta_1 = ... = \beta_J$. Using the ACR constraint, this implies that $\beta_j = ACR$ for all j. Using this to re-write $CE(0, j_2)$ as $CE(0, j_2) = \sum_{j=0}^{j_2} \beta_j = ACR \times j_2$. Finally, the same argument as in Appendix Section G can be used to establish that these bounds are sharp. ■

Proposition 2* also sheds light on the *width* of the bounds on $CE(j_1, j_2)$ under Assumptions 1-7. Here, we provide some discussion on the value of j_2 for which the bounds on the $CE(0, j_2)$ are the tightest. First, we note that the upper bound $CE(0, j_2)$ is increasing in $\frac{j_2}{\sum_{j=1}^{j_2} w_j}$. Therefore, increasing j_2 by one unit will increase the upper bound more if there

is a relatively greater drop-off in the share of compliers induced by the instrument to move beyond treatment intensity j_2 i.e., if w_j for $j > j_2$ is smaller. On the other hand, the lower bound on $CE(0, j_2)$ in Proposition 2* is maximized at treatment level j_2 where $j_2 > 1/w_1$. To summarize, under Assumptions 1-7, the bounds on the effect of moving from 0 to j_2 will be tightest when j_2 is such that w_j for $j > j_2$ is small *and* when $j_2 > 1/w_1$.

While Assumption 7 does not directly apply in our analysis of Taubman et al. (2014), this discussion clarifies how the width of the bounds might change depending on the distribution of compliers and the parameter of interest. In particular, we observe relatively tight bounds on the effect of one full-year effect of Medicaid coverage. To understand why this occurs, we can see in Figure A.1(a)-(b) that only about 17% of the compliers cross a treatment intensity of greater than 12 months, and $1/w_1 \approx 13$. This implies that, at $j_2 = 12$, the lower bound may be close to its maximum value and the upper bound may be close to its minimum value. Indeed, this is the case. Figure I.1 plots the lower bound and upper bound on the effect of moving from no Medicaid coverage to j_2 months of Medicaid coverage (shown on the x-axis). As can be seen from the figure, at $j_2 = 12$, the lower bound is only approximately 8% smaller than its maximum value and the upper bound is only approximately 9% larger than its minimum value; hence the bounds (the difference between the red and blue lines) is quite narrow around 12 months.

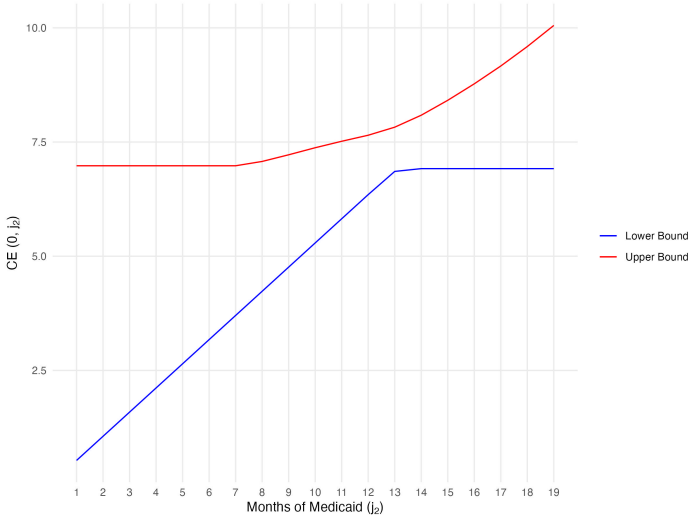


Figure I.1: Bounds on $CE(0, j_2)$ in Taubman et al. (2014)

Notes: The figure displays bounds on the effect of moving from 0 months of Medicaid coverage to j_2 months of Medicaid coverage on Emergency Department usage. Bounds on $CE(0, j_2)$ are displayed in percentage points on the y-axis. The bounds are computed by solving LP.7 for $j_2 = 1, ..., 19$.