Identification of the Average Treatment Effect when SUTVA is violated.

Lukas Laffers* Giovanni Mellace†

February 15, 2016

Abstract

In order to identify the Average Treatment Effect (ATE) of a binary treatment on an outcome of interest, we need to impose, often implicitly, the so called Stable Unit Treatment Value Assumption (SUTVA). In fact only under SUTVA we can observe at least one potential outcome for each individual. If SUTVA is violated, the ATE is not point identified even if the treatment has been randomly assigned. This paper derives sharp bounds on the ATE of an exogenous binary treatment on a binary outcome as a function of the share of the units α for which SUTVA is potentially violated. We also show how to derive the maximum value of α such that 0 (or any other value) is an extreme point of the bounds (i.e., the sign of the ATE is identified). Furthermore, after decomposing SUTVA in two separate assumptions, following the epidemiology literature, we provide weaker assumptions which might help sharpening our bounds. Furthermore, we show how some of our results can be extended to continuous outcomes. Finally we apply our bounds to two well known experiments, the US Job Corps training program and randomly assigned voucher for private schools in Colombia.

Keywords: SUTVA; Bounds; Average treatment effect; Sensitivity analysis.

*Matej Bel University, Dept. of Mathematics. E-mail: lukas.laffers@gmail.com, Web: http://www.lukaslaffers.com

†University of Southern Denmark, Dept. of Business and Economics, E-mail: giome@sam.sdu.dk, Web: http://sites.google.com/site/giovannimellace/

1 Introduction and Literature review

The name Stable Unit Treatment Value Assumption was originally proposed in Rubin (1980), but has been discuss way earlier. For example Cox (1958) assumes no interference between units. SUTVA plays a central role in the identification of causal effects. It has two implications i) it ensures that there exists as many potential outcomes as the number of value the treatment can take on (two for the binary case considered in this paper), ii) only under SUTVA we can observe at least one of the potential outcomes for each unit.

In most applications in economics SUTVA is often only implicitly assumed, although it is not always plausibly satisfied. For example SUTVA is violated in the presence of general equilibrium effects (see, Heckman et al. (1999)) which are likely to affect the evaluation of the effects of job training programs. SUTVA is likely violated in the presence of peer-effects in evaluating experiment designed to increase education (e.g., randomly assigned vouchers) or in the presence externalities and spillover effects.

Most of the literature has focused on either modeling General equilibrium effects (see, Heckman et al. (1999)) or dealt with other types of interaction effects (see Horowitz and Manski (1995), Sobel (1996)). However, SUTVA is also violated if some unit has access to different version of the treatment which may lead to a different value of the potential outcome. For this reason the recent literature in epidemiology decomposes SUTVA in two components which are somehow equivalent to the two main reasons why SUTVA can be violated.

This paper contribute to the literature in several ways. First we consider the simple binary outcome case. After fixing the share of units for which SUTVA is violated (i. e. the observed outcome differs from the potential outcome) we provide sharp bounds on the ATE which are function of this share. This allows to perform a sensitivity analysis of the point identified ATE (under SUTVA). In fact we show how to estimate the maximum share of units for which SUTVA can be violated without changing the conclusion about the sign of the ATE. We also show how the bounds can be sharpened and the sensitivity analysis can be improved by using observable covariates. Second we apply this sensitivity analysis to two well known experiment the US Job Corps already analyzed in Lee (2009) and the Colombia vouchers for private school already analyzed in Angrist et al. (2006). We find that the effect of the random assignment on both experiment is very sensitive to SUTVA violations and the shares of units for which SUTVA can be violated are very small but statistically different from 0. Third we decompose SUTVA in two separate assumptions, following the epidemiology literature, and we provide weaker alternative assumptions which can help narrowing the bounds. Finally we generalize some of our results for continuous outcomes. The paper is organized as follows, in Section 2 we introduce some necessary notation; in Section 3 we derive our bounds and provide the sensitivity analysis, in Section 4 we show the results of the empirical application, in Section 5 we look at the two components of SUTVA separately, Section 6 concludes. Proofs are provided in the appendix.

2 Setup and Notation

For each individual i in the population $\mathcal{I}$, we define:

- the observed binary outcome as $Y_i \in \mathcal{Y} = \{0, 1\}$,
- the observed binary treatment as $D_i \in \mathcal{D} = \{0, 1\}$, and
- the two potential outcomes as $(Y_i^0, Y_i^1) \in \mathcal{Y} \times \mathcal{Y}$.

We can observe the probability distribution of (Y, D) while the joint distribution of the potential outcomes outcomes $(Y(0), Y(1))$ is not observable, as we can only observe at most one potential outcome for each individual. We are interested in the average treatment effect, $ATE = E[Y(1) - Y(0)]$, which is a functional of the joint distribution of $(Y(0), Y(1), Y, D)$.

In Order to identify the ATE the first assumption that is (often implicitly) made is SUTVA:

Assumption 1: (SUTVA)

$$\forall d \in \mathcal{D}, \forall i \in \mathcal{I} : \quad \text{If } D_i = d \text{ then } Y_i(d) = Y_i.$$

Under SUTVA we can immediately relate observed and potential outcomes through the observational rule:

$$Y_i = D_i Y(1)_i + (1 - D_i) Y(0)_i.$$

As already discussed in the introduction, SUTVA requires that:

(i) There are no interaction effects.

(ii) The treatment is exhaustive, so that there are no hidden version of the treatment which can affect the potential outcomes.

We will denote the joint probability distribution of $(Y(0), Y(1), Y, D)$ by π, as in Figure 1, and by $S_i = I\{D_i = d \implies Y_i(d) = Y_i\}$ an indicator function equal to 1 if for individual i Assumption 1 holds.

SUTVA implies that

$$\pi_{13} = \pi_{14} = \pi_{22} = \pi_{24} = \pi_{31} = \pi_{32} = \pi_{41} = \pi_{43} = 0. \tag{1}$$

3 Results

3.1 Illustration: SUTVA is satisfied

Under SUTVA, the observed probabilities can be rewritten in terms of the unobserved joint probability distribution π in the following way:

$$\begin{aligned}
p_{00} &\equiv \Pr(Y=0, D=0) = \pi_{11} + \pi_{12}, \quad & E[Y(0)|D=0] &= \frac{\pi_{33} + \pi_{34}}{\Pr(D=0)}, \\
p_{01} &\equiv \Pr(Y=0, D=1) = \pi_{21} + \pi_{23}, \quad & E[Y(0)|D=1] &= \frac{\pi_{23} + \pi_{44}}{\Pr(D=1)} \\
p_{10} &\equiv \Pr(Y=1, D=0) = \pi_{33} + \pi_{34}, \quad & E[Y(1)|D=0] &= \frac{\pi_{12} + \pi_{34}}{\Pr(D=0)} \\
p_{11} &\equiv \Pr(Y=1, D=1) = \pi_{42} + \pi_{44}, \quad & E[Y(1)|D=1] &= \frac{\pi_{42} + \pi_{44}}{\Pr(D=1)}.
\end{aligned}$$

Also the observed mean outcome conditional on the treatment is equal to the mean potential outcome conditional on the treatment.

$$\begin{aligned}
E[Y|D=0] &= \frac{\pi_{33} + \pi_{34}}{\Pr(D=0)} = E[Y(0)|D=0], \\
E[Y|D=1] &= \frac{\pi_{42} + \pi_{44}}{\Pr(D=1)} = E[Y(1)|D=0].
\end{aligned}$$

The mean potential outcomes can be rewritten as

$$
\begin{aligned}
E[Y(0)] &= E[Y(0)|D=1]\cdot \Pr(D=1) + E[Y(0)|D=0]\cdot \Pr(D=0)\\
&= \pi_{23}+\pi_{44}+\pi_{33}+\pi_{34},\\
E[Y(1)] &= E[Y(1)|D=1]\cdot \Pr(D=1) + E[Y(1)|D=0]\cdot \Pr(D=0)\\
&= \pi_{42}+\pi_{44}+\pi_{12}+\pi_{34},
\end{aligned}
\tag{2}
$$

This implies that the ATE can be written as:

$$
E[Y(1)-Y(0)] = \pi_{42}+\pi_{12}-\pi_{23}-\pi_{33}. \tag{3}
$$

If we assume that the treatment is exogenous it is well known that the ATE is a function of only observable quantities and is therefore identified. We summarize this results in Lemma 1 after having formally defined exogeneity.

Assumption 2: (Exogenous Treatment Selection)

$$
\forall d \in \mathcal{D}: \quad E[Y(d)|D=1] = E[Y(d)|D=0].
$$

Lemma 1. *Under Assumptions 1 and 2, the ATE is identified.*

Proof of Lemma 1. Under Assumption 1, $E[Y(d)|D=d] = E[Y|D=d]$, and under Assumption 2, $E[Y(d)|D=1] = E[Y(d)|D=0]$ and hence $ATE = E[Y(1)-Y(0)] = E[Y|D=1] - E[Y|D=0]$ is identified from the data. □

3.2 SUTVA does not hold

When SUTVA does not hold the observed probabilities become

$$
\begin{aligned}
p_{00} &= \pi_{11}+\pi_{12}+\pi_{13}+\pi_{14}, \quad & E[Y(0)|D=0] &= \frac{\pi_{33}+\pi_{34}+\pi_{13}+\pi_{14}}{\Pr(D=0)},\\
p_{01} &= \pi_{21}+\pi_{23}+\pi_{22}+\pi_{24}, \quad & E[Y(0)|D=1] &= \frac{\pi_{23}+\pi_{44}+\pi_{24}+\pi_{43}}{\Pr(D=1)},\\
p_{10} &= \pi_{33}+\pi_{34}+\pi_{31}+\pi_{32}, \quad & E[Y(1)|D=0] &= \frac{\pi_{12}+\pi_{34}+\pi_{14}+\pi_{32}}{\Pr(D=0)},\\
p_{11} &= \pi_{42}+\pi_{44}+\pi_{41}+\pi_{43}, \quad & E[Y(1)|D=1] &= \frac{\pi_{42}+\pi_{44}+\pi_{22}+\pi_{24}}{\Pr(D=1)}.
\end{aligned}
\tag{4}
$$

The fundamental difference is that without SUTVA the potential outcomes for a given observed treatment value are not identified from the data, so the observed $E[Y|D=d]$ need not be equal to $E[Y(d)|D=d]$ anymore, i.e.:

$$E[Y|D=0] = \frac{\pi_{33}+\pi_{34}+\pi_{31}+\pi_{32}}{\Pr(D=0)} \neq \frac{\pi_{33}+\pi_{34}+\pi_{13}+\pi_{14}}{\Pr(D=0)} = E[Y(0)|D=0]$$
$$E[Y|D=1] = \frac{\pi_{42}+\pi_{44}+\pi_{41}+\pi_{43}}{\Pr(D=1)} \neq \frac{\pi_{42}+\pi_{44}+\pi_{22}+\pi_{24}}{\Pr(D=1)} = E[Y(1)|D=1].$$

The mean potential outcomes now become

$$\begin{aligned} E[Y(0)] &= E[Y(0)|D=1]\cdot\Pr(D=1) + E[Y(0)|D=0]\cdot\Pr(D=0) \\ &= \pi_{23}+\pi_{44}+\pi_{24}+\pi_{43}+\pi_{33}+\pi_{34}+\pi_{13}+\pi_{14}, \\ E[Y(1)] &= E[Y(1)|D=1]\cdot\Pr(D=1) + E[Y(1)|D=0]\cdot\Pr(D=0) \\ &= \pi_{42}+\pi_{44}+\pi_{22}+\pi_{24}+\pi_{12}+\pi_{34}+\pi_{14}+\pi_{32}, \end{aligned}$$

and therefore

$$E[Y(1)-Y(0)] = \pi_{42}+\pi_{12}+\pi_{22}+\pi_{32}-\pi_{23}-\pi_{33}-\pi_{13}-\pi_{43}.$$

The ATE can still be identified, but only at the price of strong additional assumptions. We propose an example of a sufficient condition that guarantee identification.
Assumption 3: (Balanced effect of the violation)

$$\begin{aligned} &\Pr(Y=1,S=0|D=1) - \Pr(Y=0,S=0|D=1) \\ &= \Pr(Y=1,S=0|D=0) - \Pr(Y=0,S=0|D=0) \end{aligned} \tag{5}$$

The Assumption 3 states that the difference between the probability of positive and negative outcome together with the violation of SUTVA is the same for treated and non-treated population. The following lemma shows that this assumption guarantees that the difference between the naive ATE estimated, $E[Y|D=1]-E[Y|D=0]$, and the ATE under Assumption 2, $E[Y(1)|D=1]-E[Y(0)|D=0]$ vanishes.

Lemma 2. *Under Assumptions 2 and 3, the ATE is identified.*

Proof of this lemma and all the other proofs are given in the Appendix.

3.3 Relaxing SUTVA

In this section we show propose a sensitivity analysis of the ATE to SUTVA violations.
Assumption 1α: (SUTVA violation share)

$$\Pr(\forall d \in \mathcal{D} : D_i = d \implies Y_i(d) = Y_i) \geq 1 - \alpha.$$

The sensitivity parameter α ranges between 0 and 1 and can be directly interpreted as a probability that SUTVA does not hold. This assumption implies that

$$\pi_{13} + \pi_{14} + \pi_{22} + \pi_{24} + \pi_{31} + \pi_{32} + \pi_{41} + \pi_{43} \leq \alpha.$$

Under the Assumption 1α, the ATE is no longer identified, but the following lemma provides its sharp bounds.

Lemma 3. *Under the Assumption 1α, the sharp bounds on the ATE are the following:*[1]

$$\begin{aligned} ATE &\in [ATE^{LB}, ATE^{UB}] \\ ATE^{LB} &= \max\{-p_{10} - p_{01} - \alpha, -1\}, \\ ATE^{UB} &= \min\{p_{00} + p_{11} + \alpha, 1\}. \end{aligned} \tag{6}$$

The width of these bounds is $1 + 2\alpha$ and they are therefore not useful in practice. We extend these results to the case with continuous outcome Y in Appendix C. In order to obtain meaningful bounds we also need to assume that the treatment is exogenous (Assumption 2).

[1] The dependence of ATE^{LB} and ATE^{UB} on α is suppressed for brevity.

Lemma 4. *Under the Assumptions 1α and 2, the sharp bounds on the ATE are the following:*

$$\begin{aligned}
ATE &\in [ATE^{LB}, ATE^{UB}] \\
&\text{if } p_{11} + p_{01} > p_{00} + p_{10}: \\
ATE^{LB} &= \frac{p_{11} - \min\{\max\{\alpha - p_{00}, 0\}, p_{11}\}}{p_{11} + p_{01}} - \frac{p_{10} + \min\{p_{00}, \alpha\}}{p_{00} + p_{10}}, \\
ATE^{UB} &= \frac{p_{11} + \min\{\max\{\alpha - p_{10}, 0\}, p_{01}\}}{p_{11} + p_{01}} - \frac{p_{10} - \min\{p_{10}, \alpha\}}{p_{00} + p_{10}}, \\
&\text{if } p_{11} + p_{01} < p_{00} + p_{10}: \\
ATE^{LB} &= \frac{p_{11} - \min\{p_{11}, \alpha\}}{p_{11} + p_{01}} - \frac{p_{10} + \min\{\max\{\alpha - p_{11}, 0\}, p_{00}\}}{p_{00} + p_{10}}, \\
ATE^{UB} &= \frac{p_{11} + \min\{p_{01}, \alpha\}}{p_{11} + p_{01}} - \frac{p_{10} - \min\{\max\{\alpha - p_{01}, 0\}, p_{01}\}}{p_{00} + p_{10}}.
\end{aligned} \tag{7}$$

The dependence of the bounds on the relaxation parameter α is visualized in Figure 3.

Lemma 4 allows us to detect the maximal possible violation of SUTVA, so that the sign of ATE is still identified.

Lemma 5. *Under Assumptions 1α and 2,* $ATE^{LB} \geq 0$ *if and only if*

$$0 \leq \alpha \leq \alpha^{+} \equiv \min\{\Pr(D=1), \Pr(D=0)\} \cdot [E(Y|D=1) - E(Y|D=0)]$$

and $ATE^{UB} \leq 0$ *if and only if*

$$0 \leq \alpha \leq \alpha^{-} \equiv -\min\{\Pr(D=1), \Pr(D=0)\} \cdot [E(Y|D=1) - E(Y|D=0)].$$

It is interesting test the hypothesis $H_0 : \alpha^{+} = 0$, so that the maximum possible violation of SUTVA assumption to guarantee positive ATE is zero. It is a question whether we can reject the hypothesis that positive ATE is robust to mild deviations of SUTVA assumption. Under the Assumptions 1α and 2 for $\alpha = 0$, the $ATE = E[Y|D = 1] - E[Y|D = 0] > \alpha^{+}$. This means that we may be in a situation where ATE is significantly different from 0, whereas α^{+} is not.

3.4 Narrowing the Bounds using Covariates

Suppose that a set covariates $X_i \in \mathbf{X}$ is also available and that all our assumptions hold also conditioned on X, such that $ATE = \int_{\mathcal{X}} ATE_x \Pr(X = x)dx$, where $ATE_x = E[Y(1) - Y(0)|X = x]$.

Assumption 2X: (Exogenous Treatment Selection with Covariates)

$$\forall d \in \mathcal{D} :, \forall x \in \mathcal{X} : \quad E[Y(d)|D = 1, X = x] = E[Y(d)|D = 0, X = x].$$

Lemma 6. *Under the Assumptions 1α and 2X, the sharp bounds on the ATE are the following:*

$$\begin{aligned}
ATE &\in \left[\overline{ATE}^{LB}, \overline{ATE}^{UB}\right] \\
\overline{ATE}^{LB} &= \int_{\mathbf{X}} ATE_x^{LB} \Pr(X = x)dx \\
\overline{ATE}^{UB} &= \int_{\mathbf{X}} ATE_x^{UB} \Pr(X = x)dx \\
&\text{if } p_{11|x} + p_{01|x} > p_{00|x} + p_{10|x} : \\
ATE_x^{LB} &= \frac{p_{11|x} - \min\{\max\{\alpha - p_{00|x}, 0\}, p_{11|x}\}}{p_{11|x} + p_{01|x}} - \frac{p_{10|x} + \min\{p_{00|x}, \alpha\}}{p_{00|x} + p_{10|x}}, \\
ATE_x^{UB} &= \frac{p_{11|x} + \min\{\max\{\alpha - p_{10|x}, 0\}, p_{01|x}\}}{p_{11|x} + p_{01|x}} - \frac{p_{10} - \min\{p_{10|x}, \alpha\}}{p_{00|x} + p_{10|x}}, \\
&\text{if } p_{11|x} + p_{01|x} < p_{00|x} + p_{10|x} : \\
ATE_x^{LB} &= \frac{p_{11|x} - \min\{p_{11|x}, \alpha\}}{p_{11|x} + p_{01|x}} - \frac{p_{10|x} + \min\{\max\{\alpha - p_{11|x}, 0\}, p_{00|x}\}}{p_{00} + p_{10}}, \\
ATE_x^{UB} &= \frac{p_{11|x} + \min\{p_{01|x}, \alpha\}}{p_{11|x} + p_{01|x}} - \frac{p_{10|x} - \min\{\max\{\alpha - p_{10|x}, 0\}, p_{01|x}\}}{p_{00|x} + p_{10|x}}.
\end{aligned} \tag{8}$$

Furthermore, $\overline{ATE}^{LB} \geq ATE^{LB}$ *and* $\overline{ATE}^{UB} \leq ATE^{UB}$.

In practice we may divide the sample into finite number of groups depending on how on the predicted value of the outcome variable.[2] The choice of the number of groups depends on the problem at hand. The larger the number the sharper are the bounds, but at the same time, the statistical uncertainty within the group increases.

When information about X is available, the maximum possible violation of SUTVA, $\alpha^+(\alpha^-)$ that guarantees positive (negative) ATE changes.

[2] Lee (2009) used all available covariates to construct a single variable that was discretized into five groups depending on the size of the value of the outcome it predicted.

Lemma 7. *Under the Assumptions 1α and 2X,* $\overline{ATE}^{LB} \geq 0$ *if and only if*

$$0 \leq \alpha \leq \overline{\alpha}^{+} \equiv \int_{\mathbf{X}} \min\{\alpha_x^{+}, 0\} \Pr(X = x) dx$$

and $\overline{ATE}^{UB} \leq 0$ *if and only if*

$$0 \leq \alpha \leq \overline{\alpha}^{-} \equiv \int_{\mathbf{X}} \min\{\alpha_x^{-}, 0\} \Pr(X = x) dx,$$

where

$$\alpha_x^{+} \equiv \min\{\Pr(D = 1, X = x), \Pr(D = 0, X = x)\} \cdot [E(Y|D = 1, X = x) - E(Y|D = 0, X = x)]$$
$$\alpha_x^{-} \equiv -\alpha_x^{+}.$$

We note that $\overline{\alpha}^{+} \leq \alpha^{+}$ (similarly $\overline{\alpha}^{-} \geq \alpha^{+}$), because for some x the quantity $E(Y|D = 1, X = x) - E(Y|D = 0, X = x)$ may be negative even though $E(Y|D = 1) - E(Y|D = 0) \geq 0$.

3.5 Estimation and Inference

The fact that the expressions for bounds, α^{+} and α^{-} involve minimum and maximum operators gives rise to a non-standard inferential procedure as no regular $\sqrt{n}$-consistent estimator exists (Hirano and Porter, 2012) and analog estimators may be severely biased in small samples. For this reason we use Intersection Bounds approach of Chernozhukov et al. (2013) that creates half-median-unbiased point estimates and confidence intervals.[3] This method corrects for the small sample bias *before* the max/min operator is applied.

4 Empirical Illustrations

This paper considers two empirical applications to illustrate the scope of usefulness of the presented results. The first one is the effect of job training assignment to U.S Job Corps program on the probability of employment four years after the assignment,

[3] Half-median-unbiased means that the estimate of the upper(lower) bound exceeds (lies below) its true value with probability at least one half asymptotically.

an intention-to-treat effect. Evaluations of this program aroused considerable interest among policy makers and researchers throughout the last decades, which is hardly surprising given the high cost of this program. We used data from National Job Corps Study also studied in Lee (2009). We refer the reader to Lee (2009) for extensive data description.

The second application looks at a school voucher experiment of Colombia's "Programa de Ampliacion de Cobertura de la Educacion Secundaria" (PACES) and analyze the impact of randomly assigned high school voucher to low income pupils that covered approximately a half of the cost of private secondary schooling on the probability that grades had to be repeated. This section uses the data previously studied in Angrist et al. (2006).

4.1 The Effect of Job Training Programme on Employment

Table 1 provides the summary statistics.

$Y \setminus D$	offered training ($D = 1$)	not offered training ($D = 0$)
working ($Y = 1$)	$p_{11} = 49.26\%$	$p_{10} = 31.63\%$
not working ($Y = 0$)	$p_{01} = 11.16\%$	$p_{00} = 7.94\%$
$n = 11146$	$\Pr(D = 1) = 60.43\%$	$\Pr(D = 0) = 39.57\%$

Table 1: Probability distribution of the working after 202 weeks indicator (Y) and of the randomized treatment (Job training programme JobCorps) offered status (D). Based on a dataset from Lee (2009). Missing values were removed.

Under SUTVA assumption and under the Exogenous Treatment Selection assumption the impact of the assignment on probability of employment is 1.6% with the lower 95% confidence bound positive at 0.1%. The minimal value of SUTVA relaxation that still yields positive ATE, α^+, is 0.954% and is statistically different from zero. The impact of different relaxations of SUTVA on ATE bounds are presented in Table 2 and visualised in Figure 3.

This analysis suggests that should we have some doubts about the mismeasurement or interaction effects within the tested subpopulation, the data cannot rule out a negative effect of the program assignment on the employment.

α	$[ATE^{LB}, ATE^{UB}]$ (CB^{LB}, CB^{UB})
0	[0.016, 0.016] (0.001, 0.031)
0.01	[-0.009, 0.041] (-0.023, 0.055)
0.05	[-0.111, 0.142] (-0.124, 0.155)
0.1	[-0.219, 0.269] (-0.230, 0.282)
0.2	[-0.384, 0.521] (-0.394, 0.537)
0.5	[-0.881, 1] (-0.893, 1)

α^+	0.954%
(CB^l, CB^u)	(0.076%, 1.213%)

Table 2: Bounds on ATE under different relaxations of SUTVA assumption. The left table presents estimates of bounds on ATE together with 95% confidence bounds. On the right had side, α^+ is the estimated maximum possible violation of SUTVA that still guarantees positive ATE. All estimates are half-median unbiased and based on Chernozhukov et al. (2013) using 9999 bootstrap samples and 200000 replications.

4.2 The Effect of School Vouchers on Never Repeating a Grade

See Angrist et al. (2006) for extensive data description.

$Y \setminus D$	offered voucher ($D = 1$)	not offered voucher ($D = 0$)
never repeated a grade ($Y = 1$)	$p_{11} = 43.71\%$	$p_{10} = 37.30\%$
repeated a grade ($Y = 0$)	$p_{01} = 8.41\%$	$p_{00} = 10.57\%$
$n = 1201$	$\Pr(D = 1) = 52.12\%$	$\Pr(D = 0) = 47.88\%$

Table 3: Probability distribution of never repeating a grade (Y) and of the randomized treatment (school vouchers offered) . Based on a dataset from Angrist et al. (2006). Missing values were removed.

Without SUTVA relaxation, the point identified ATE of voucher offered on the probability of never repeating a grade is 6% and it is statistically significant on the 95% confidence level. In order to maintain a positive effect, we may have no more 3.03% of the population that do not satisfy SUTVA, so the positive effect is more robust to the relaxation of SUTVA assumption that the previous example. Should the proportion of individuals that violate SUTVA assumption becomes larger than 10%, the data is uninformative about the direction of the effect with 95% confidence bound between -19.2% and 31.8%.

The results are summarized in Table 4 and depicted in Figure 4.

α	$[ATE^{LB}, ATE^{UB}]$ (CB^{LB}, CB^{UB})
	[0.060, 0.060] (0.009, 0.110)
0.01	[0.033, 0.092] (-0.014, 0.136)
0.05	[-0.050, 0.174] (-0.094, 0.215)
0.1	[-0.154, 0.278] (-0.192, 0.318)
0.2	[-0.348, 0.485] (-0.384, 0.528)
0.5	[-0.932, 1] (-0.969, 1)

α^{+}	3.03%
(CB^{l}, CB^{u})	(0.69%, 5.08%)

Table 4: Bounds on ATE under different relaxations of SUTVA assumption. The left table presents estimates of bounds on ATE together with 95% confidence bounds. On the right had side, α^{+} is the estimated maximum possible violation of SUTVA that still guarantees positive ATE. All estimates are half-median unbiased and based on Chernozhukov et al. (2013) using 9999 bootstrap samples and 200000 replications.

5 Extension: Decomposing SUTVA assumption

In the relaxed SUTVA assumption, we are completely agnostic about the mechanism that drives the violation. This may be an advantage or a disadvantage. In certain situations it may be desirable to distinguish between the different reasons why SUTVA may be violated.

In the epidemiology literature, Assumption 1 in this paper (which we call SUTVA assumption) was coined as a *Consistency assumption* (Cole and Frangakis, 2009). This assumption was further decomposed by VanderWeele (2009) into two components: *Treatment-variation irrelevance assumption* and *Consistency assumption*.

In order to make the distinction possible we introduce a new variable $H_i \in \mathcal{H}$, that denotes a hidden treatment of individual i. This may capture different conditions under which the treatment D is taken (e.g. different dose or length of exposure to treatment). Now, the potential outcome is a function of both observed and hidden treatment $Y(d,h)$. In this case the average treatment effect depends on the value of H and the quantity of interest may be the mean of average treatment effects for different values of the hidden treatment: $ATE = \int_{\mathcal{H}} ATE(h) \Pr(H = h) dh$, where $ATE(h) = E[Y(1,h) - Y(0,h)]$.

Assumption 1A: (Treatment-variation irrelevance assumption)

$$\forall d \in \mathcal{D}, \forall h, h' \in \mathcal{H}, \forall i \in \mathcal{I}: \quad D_i = d \implies Y_i(d,h) = Y_i(d,h'). \tag{9}$$

This means that there are no multiple versions of the treatment and the notation $Y_i(d)$ is justified and the quantity $ATE = E(Y(1) - Y(0))$ is well defined. It also means that there is no interference: $Y_i(d_i, \mathbf{d}_{-i}) = Y_i(d_i, \mathbf{d'}_{-i}), \forall \mathbf{d}_{-i}, \mathbf{d'}_{-i}$, where $\mathbf{d}_{-i}$ stands for the vector of treatments of individuals other than i.

Assumption 1B: (Consistency Assumption)

$$\forall d \in \mathcal{D}, \forall h \in \mathcal{H}, \forall i \in \mathcal{I}: \; D_i = d, \; H_i = h \implies Y_i(d,h) = Y_i. \tag{10}$$

This assumption states that the observed value of outcome Y_i is consistent with the potential outcome model formulation. A possible violation of this assumption is mismeasurement of the outcome or treatment.

We note that Assumptions 1A and 1B imply the following condition

$$\forall d \in \mathcal{D}, \forall h, h' \in \mathcal{H}, \forall i \in \mathcal{I}: \quad D_i = d,\ H_i = h \implies Y_i(d,h) = Y_i(d,h') = Y_i,$$

which reduces to Assumption 1 if there are no hidden treatments H.

Figure 5 depicts the Individual Average Treatment Effecton and the support of the joint probability distribution of $(Y^{00}, Y^{01}, Y^{10}, Y^{11}, Y, D, H)$ for binary hidden treatment H, where $Y^{dh} = Y(d,h)$.

Both Assumptions 1A and 1B are support restrictions and so we can relax these assumptions separately.

Assumption 1Aβ: (Relaxed Treatment-variation Irrelevance Assumption)

$$\Pr(\forall d \in \mathcal{D}, \forall h, h' \in \mathcal{H}: \ Y_i(d,h) = Y_i(d,h')) \geq 1 - \beta. \tag{11}$$

Assumption 1Bγ: (Relaxed Consistency Assumption)

$$\Pr(\forall d \in \mathcal{D}, \forall h \in \mathcal{H}: \ D_i = d,\ H_i = h \implies Y_i(d,h) = Y_i) \geq 1 - \gamma. \tag{12}$$

Assumption 2H: (Exogenous Treatment Selection with Hidden Treatment)

$$\forall d \in \mathcal{D},\ \forall h \in \mathcal{H}: \ E[Y(d,h)|D=1] = E[Y(d,h)|D=0].$$

We note that the Assumption 1Aβ has no identifying power once Assumption 1Bγ is assumed. This result does not change even when Assumption 2H is assumed. This observation is based on a simulation, where the bounds on ATE are calculated using a linear programming procedure described in Lafférs (2015). We note that there are recent advances in the literature of statistical inference of partially identified parameters that deal with random linear program of such form (Kaido et al., 2016). Further research is warranted.

6 Conclusion

This paper discussed the Stable Unit Treatment Value Assumption (SUTVA) assumptions and the implications of the violations of this assumption for the identification on the average treatment effect. SUTVA assumption is relaxed in a way that a certain fraction of the population may violate this assumption. We found analytic bounds on the ATE under the relaxed SUTVA and the Exogenous Treatment Selection assumption and demonstrated it on two empirical examples. Furthermore these results allow us to identify the maximum amount of SUTVA violation that would still result in a positive (negative) ATE. Following the epidemiology literature, this paper sketched the possible decomposition of SUTVA assumption that allows to distinguish between the different sources of SUTVA violation.

A Appendix

Proof of Lemma 2. The Assumption 2 together with (4) implies

$$\begin{aligned} ATE &= E[Y(1)] - E[Y(0)] = E[Y(1)|D=1] - E[Y(0)|D=0] \\ &= \frac{\pi_{42} + \pi_{44} + \pi_{22} + \pi_{24},}{\Pr(D=1)} - \frac{\pi_{33} + \pi_{34} + \pi_{13} + \pi_{14},}{\Pr(D=0)}. \end{aligned} \tag{A.1}$$

From (5) we can see that

$$E[Y|D=1] - E[Y|D=0] = \frac{\pi_{42} + \pi_{44} + \pi_{41} + \pi_{43},}{\Pr(D=1)} - \frac{\pi_{33} + \pi_{34} + \pi_{31} + \pi_{32}}{\Pr(D=0)}. \tag{A.2}$$

We note that under Assumption 3,

$$\frac{\pi_{41} + \pi_{43}}{\Pr(D=1)} - \frac{\pi_{22} + \pi_{24}}{\Pr(D=1)} = \frac{\pi_{31} + \pi_{32}}{\Pr(D=0)} - \frac{\pi_{13} + \pi_{14}}{\Pr(D=0)},$$

so that the equations (A.1) and (A.2) are equal. □

Proof of Lemma 3. We show the proof for the upper bound as the proof for the lower bound follows in an analogous way.

Let us further denote $ATE^{s}_{yd} = E[Y(1) - Y(0)|Y = y, D = d, S = s]$.

(i) Validity

$$\begin{aligned}
ATE =& \Big[ATE_{00}^{1} \cdot \Pr(S=1|Y=0,D=0) + ATE_{00}^{0} \cdot \Pr(S=0|Y=0,D=0)\Big] \cdot p_{00} \\
&+ \Big[ATE_{01}^{1} \cdot \Pr(S=1|Y=0,D=1) + ATE_{01}^{0} \cdot \Pr(S=0|Y=0,D=1)\Big] \cdot p_{01} \\
&+ \Big[ATE_{10}^{1} \cdot \Pr(S=1|Y=1,D=0) + ATE_{10}^{0} \cdot \Pr(S=0|Y=1,D=0)\Big] \cdot p_{10} \\
&+ \Big[ATE_{11}^{1} \cdot \Pr(S=1|Y=1,D=1) + ATE_{11}^{0} \cdot \Pr(S=0|Y=1,D=1)\Big] \cdot p_{11} \\
\leq& [1 \cdot \Pr(S=1|Y=0,D=0)] + 0 \cdot \Pr(S=0|Y=0,D=0)] \cdot p_{00} \\
&+ [0 \cdot \Pr(S=1|Y=0,D=1)] + 1 \cdot \Pr(S=0|Y=0,D=1)] \cdot p_{01} \\
&+ [0 \cdot \Pr(S=1|Y=1,D=0)] + 1 \cdot \Pr(S=0|Y=1,D=0)] \cdot p_{10} \\
&+ [1 \cdot \Pr(S=1|Y=1,D=1)] + 0 \cdot \Pr(S=0|Y=1,D=1)] \cdot p_{11} \\
=& \Pr(S=1|Y=0,D=0) \cdot p_{00} + \Pr(S=1|Y=1,D=1) \cdot p_{11} \\
&+ \Pr(S=0|Y=0,D=1) \cdot p_{01} + \Pr(S=0|Y=1,D=0) \cdot p_{10} \\
\leq& p_{00} + p_{11} + \min\{p_{01} + p_{10}, \alpha\} = \min\{p_{00} + p_{11} + \alpha, 1\},
\end{aligned}$$

Where the last inequality follows from the fact that $\Pr(S=0) \leq \alpha$.

(ii) Sharpness

Suppose that $\alpha < p_{01} + p_{10}$. Then there must exist constants $0 \leq \alpha_{01} \leq p_{01}$ and $0 \leq \alpha_{10} \leq p_{10}$, so that $\alpha = \alpha_{01} + \alpha_{10}$. The following specification for $\Pr(Y(0), Y(1), Y, D)$ is compatible with Assumption 1α and and with the distribution of (Y, D).

$$\begin{aligned}
&\pi_{12} = p_{00},\ \pi_{22} = \alpha_{01},\ \pi_{32} = \alpha_{10},\ \pi_{42} = p_{11},\ \pi_{21} = p_{01} - \alpha_{01},\ \pi_{34} = p_{10} - \alpha_{10}, \\
&\pi_{11} = \pi_{13} = \pi_{14} = \pi_{23} = \pi_{24} = \pi_{31} = \pi_{33} = \pi_{41} = \pi_{43} = \pi_{44} = 0.
\end{aligned}$$

Suppose now that $\alpha \geq p_{01} + p_{10}$.

$$\begin{aligned}
&\pi_{12} = p_{00},\ \pi_{22} = p_{01},\ \pi_{32} = p_{10},\ \pi_{42} = p_{11}, \\
&\pi_{11} = \pi_{13} = \pi_{14} = \pi_{21} = \pi_{23} = \pi_{24} = \pi_{31} = \pi_{33} = \pi_{34} = \pi_{41} = \pi_{43} = \pi_{44} = 0.
\end{aligned}$$

Figure 2 illustrates the sharpness part of the proof of Lemma 3.

□

Proof of Lemma 4. We show the proof for the upper bound and for $\pi_{11} + \pi_{01} > \pi_{00} + \pi_{10}$ as the proof for the lower bound and for $\pi_{11} + \pi_{01} < \pi_{00} + \pi_{10}$ follows in an analogous way.

(i) Validity

$$
\begin{aligned}
ATE &= E[Y(1)-Y(0)] = E[Y(1)|D=1] - E[Y(0)|D=0] \\
&= \frac{\pi_{42}+\pi_{44}+\pi_{22}+\pi_{24}}{\Pr(D=1)} - \frac{\pi_{33}+\pi_{34}+\pi_{13}+\pi_{14}}{\Pr(D=0)} \\
&= \frac{p_{11}-\pi_{41}-\pi_{43}+\pi_{22}+\pi_{24}}{p_{11}+p_{01}} - \frac{p_{10}-\pi_{31}-\pi_{32}+\pi_{13}+\pi_{14}}{p_{00}+p_{10}} \\
&\leq \frac{p_{11}+\pi_{22}+\pi_{24}}{p_{11}+p_{01}} - \frac{\pi_{10}-\pi_{31}-\pi_{32}}{p_{00}+p_{10}} \\
&\leq \frac{p_{11}+\min\{\max\{\alpha-p_{10},0\},p_{01}\}}{p_{11}+p_{01}} - \frac{p_{10}-\min\{p_{10},\alpha\}}{p_{00}+p_{10}} = ATE^{UB}.
\end{aligned}
$$

where the last inequality follows from inequalities $\pi_{31}+\pi_{32} \leq p_{10}$, $\pi_{22}+\pi_{24} \leq p_{01}$ and $\pi_{11}+\pi_{01} > \pi_{00}+\pi_{10}$.

(ii) Sharpness

Given that $\pi_{11}+\pi_{01} > \pi_{00}+\pi_{10}$, the following specification for $\Pr(Y(0),Y(1),Y,D)$ is compatible with Assumptions 1α, 2, with the distribution of (Y,D) and achieves the ATE^{UB}.

$$c_1 = \min\{p_{10},\alpha\},$$

$$c_2 = \min\{\max\{\alpha-p_{10},0\},p_{01}\},$$

$$
\begin{aligned}
\pi_{11} &= p_{00} - p_{00}\frac{p_{11}+c_2}{p_{11}+p_{01}}, & \pi_{21} &= p_{01} - c_2 - p_{01}\frac{p_{10}-c_1}{p_{00}+p_{10}}, \\
\pi_{12} &= p_{00}\frac{p_{11}+c_2}{p_{11}+p_{01}}, & \pi_{22} &= c_2, \\
\pi_{13} &= 0, & \pi_{23} &= p_{01}\frac{p_{10}-c_1}{p_{00}+p_{10}}, \\
\pi_{14} &= 0, & \pi_{24} &= 0, \\
\pi_{31} &= c_1 - c_1\frac{p_{11}+c_2}{p_{11}+p_{01}}, & \pi_{41} &= 0, \\
\pi_{32} &= c_1\frac{p_{11}+c_2}{p_{11}+p_{01}}, & \pi_{42} &= p_{11} - p_{11}\frac{p_{10}-c_1}{p_{00}+p_{10}}, \\
\pi_{33} &= p_{10} - c_1 - (p_{10}-c_1)\frac{p_{11}+c_2}{p_{11}+p_{01}}, & \pi_{43} &= 0, \\
\pi_{34} &= (p_{10}-c_1)\frac{p_{11}+c_2}{p_{11}+p_{01}}, & \pi_{44} &= p_{11}\frac{p_{10}-c_1}{p_{00}+p_{10}}.
\end{aligned}
$$

Straightforward manipulations show that the proposed specification is a proper probability distribution function.

□

Proof of Lemma 5. We only present the proof for $ATE^{LB} \geq 0$, as the the proof for $ATE^{UB} \leq 0$ is similar. Consider the case $\pi_{11}+\pi_{01} > \pi_{00}+\pi_{10}$. If $p_{00}+p_{11} \geq \alpha \geq p_{00}$, then

$$ATE^{LB} = \frac{p_{11}-(\alpha-p_{00})}{p_{11}+p_{01}} - 1,$$

so that $ATE^{LB} \geq 0$ would imply $p_{00} - p_{01} \geq \alpha$ which contradicts $\alpha \geq p_{00}$, so we have to have $\alpha \leq p_{00}$ and thus

$$ATE^{LB} = \frac{p_{11}}{p_{11} + p_{01}} - \frac{p_{10} + \alpha}{p_{00} + p_{10}} \geq 0 \iff \alpha \leq p_{11}\frac{p_{00} + p_{10}}{p_{11} + p_{01}} - p_{10} = \Pr(D = 0)\left[E(Y|D = 1) - E(Y|D = 0)\right].$$

Similarly, for $\pi_{11} + \pi_{01} > \pi_{00} + \pi_{10}$ we get that for $ATE^{UB} \leq 0$ we have to have $\alpha \leq p_{11}$ and therefore

$$ATE^{UB} = \frac{p_{11} - \alpha}{p_{11} + p_{01}} - \frac{p_{10}}{p_{00} + p_{10}} \leq 0 \iff \alpha \leq p_{11} - p_{10}\frac{p_{11} + p_{01}}{p_{00} + p_{10}} = \Pr(D = 1)\left[E(Y|D = 1) - E(Y|D = 0)\right],$$

which leads to the desired result.

□

Proof of Lemma 6. The proof is similar to the one or Proposition 1b in Lee (2009). The validity and sharpness of the bounds results from the application of Lemma 4 conditional on $X = x$. The second part follows from the fact that any ATE that is consistent with (Y, D, X) has to be consistent with (Y, D), that is ignoring the information about X cannot lead to a more informative result (sharpen the bounds). □

Proof of Lemma 7. Analoguous to the proof of Lemma 5 and hence ommited. □

B Figures

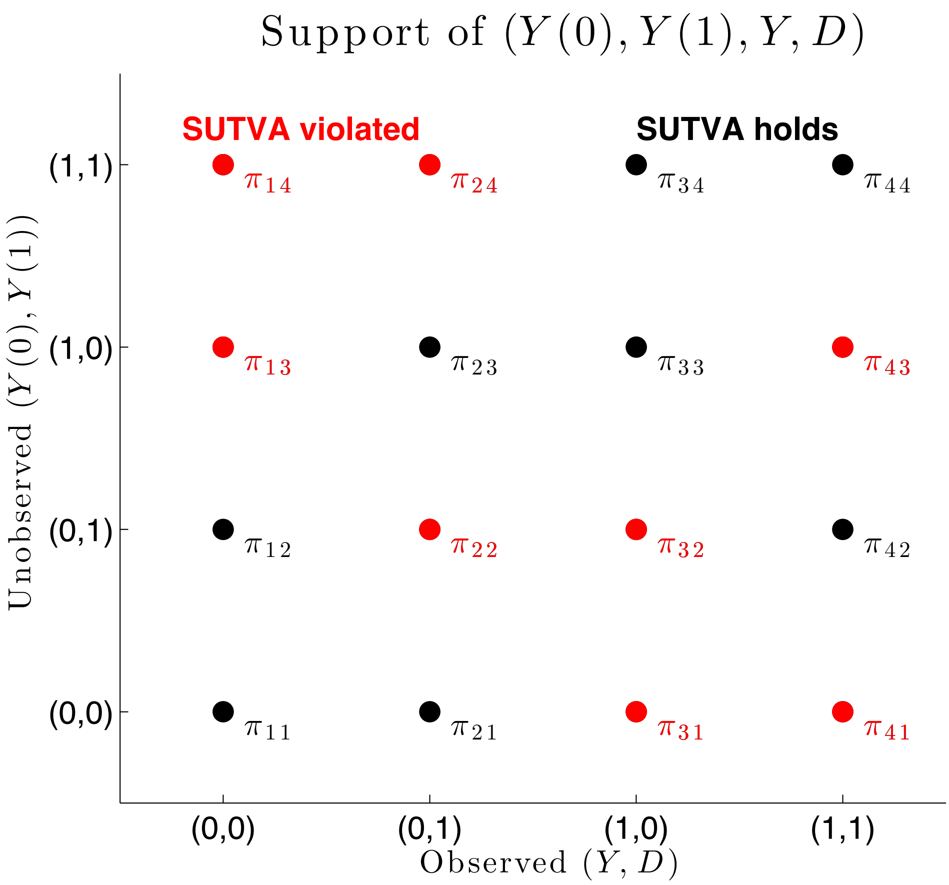

Figure 1: Joint probability distribution of $(Y(0), Y(1), Y, D)$. Under SUTVA, the red points must have zero probability mass.

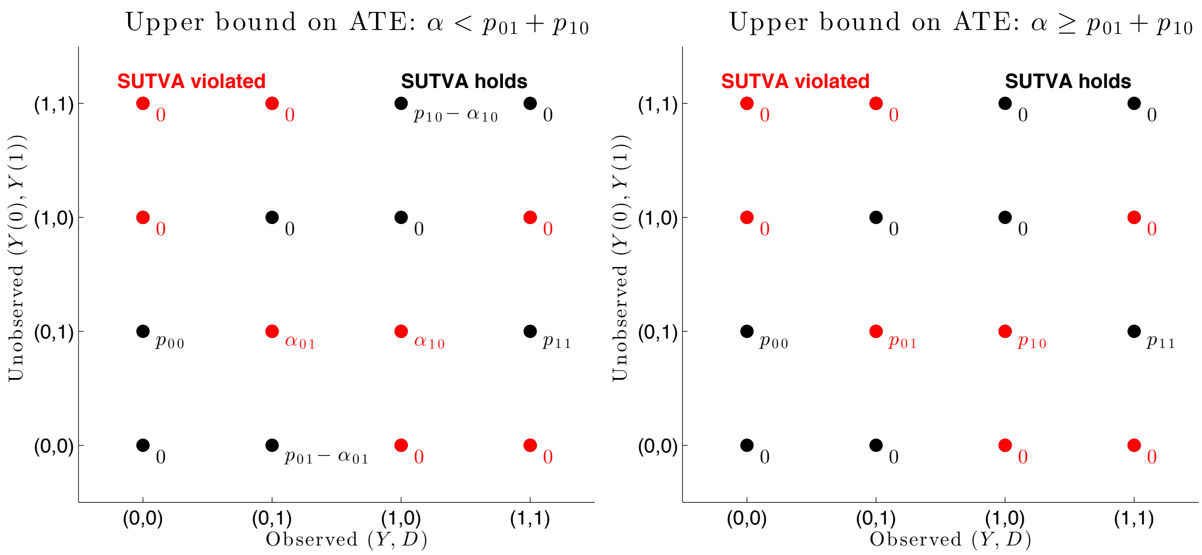

Figure 2: Visualisation of the sharpness part of the Lemma 3.

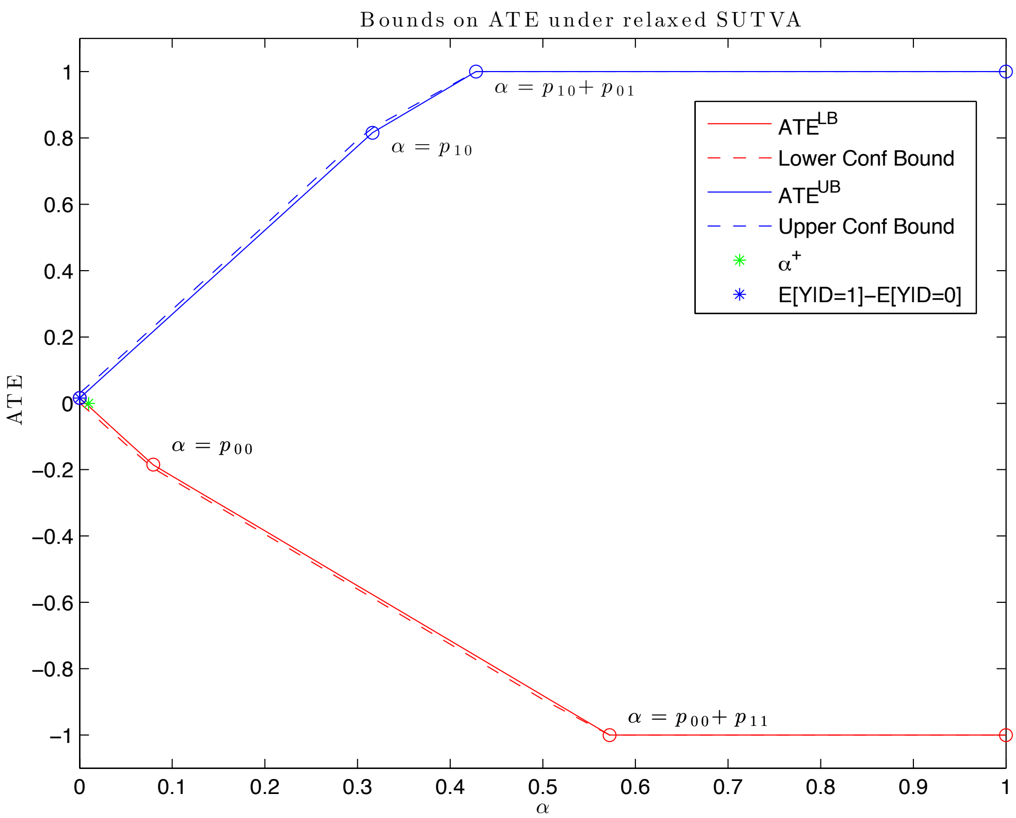

Figure 3: Sensitivity analysis to SUTVA assumption of the bounds on ATE of the assignment to job training on the probability of employment (Intention-to-Treat). All estimates are half-median unbiased and based on Chernozhukov et al. (2013).

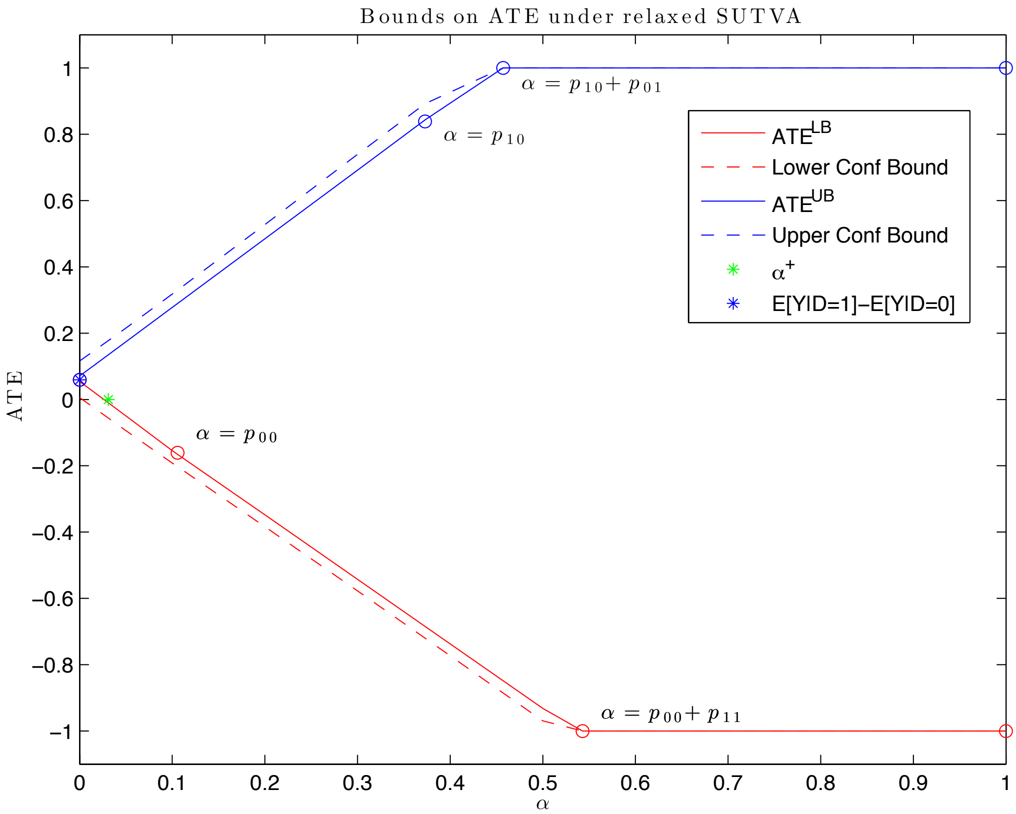

Figure 4: Sensitivity analysis to SUTVA assumption of the bounds on ATE of the school vouchers on the probability of never repeating a grade (Intention-to-Treat). All estimates are half-median unbiased and based on Chernozhukov et al. (2013).

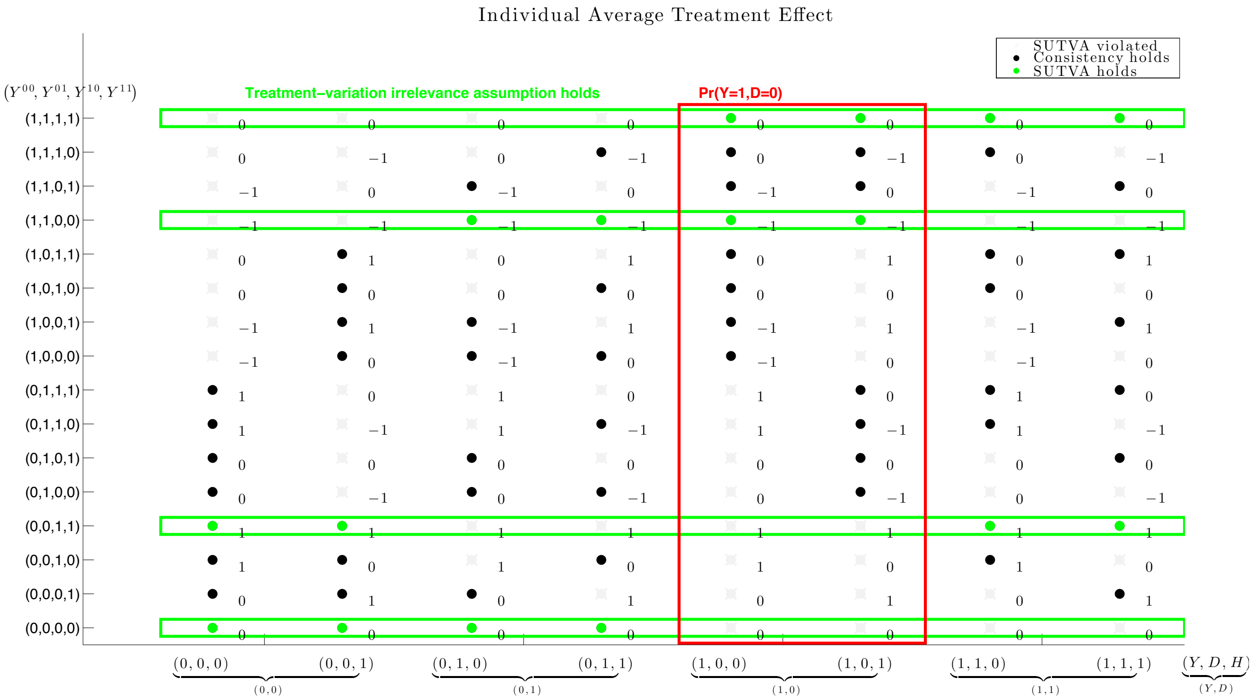

Figure 5: Individual Average Treatment Effect depicted on the support of the joint probability distribution of $(Y^{00}, Y^{01}, Y^{10}, Y^{11}, Y, D, H)$ for binary hidden treatment H. Note that only the proportions $\Pr(Y = y, D = d)$ are observed.

C Continuous Outcome

Notation: $\forall d \in \mathcal{D} : \pi^d = \pi^d(y_0, y_1, y) = f(y_0, y_1, y|d)$, $p^d = \Pr(D = d)$

$$\begin{aligned} \forall y \in \mathcal{Y} : \iint \pi^1(y_0, y_1, y)\, dy_0\, dy_1 &= f_Y(y|D=1) \\ \iint \pi^0(y_0, y_1, y)\, dy_0\, dy_1 &= f_Y(y|D=0) \end{aligned} \tag{C.1}$$

$$\begin{aligned} \forall y_0, y_1, y \in \mathcal{Y} : \pi^1(y_0, y_1, y) I\{y_1 \neq y\} = 0 \\ \forall y_0, y_1, y \in \mathcal{Y} : \pi^0(y_0, y_1, y) I\{y_0 \neq y\} = 0 \end{aligned} \tag{C.2}$$

$$\begin{aligned} \iiint y_1 \pi^1\, dy_0\, dy_1\, dy &= \iiint y_1 \pi^0\, dy_0\, dy_1\, dy \\ \iiint y_0 \pi^1\, dy_0\, dy_1\, dy &= \iiint y_0 \pi^0\, dy_0\, dy_1\, dy \end{aligned} \tag{C.3}$$

These restrictions state that π^d is compatible with the data (C.1), satisfy SUTVA assumption (C.2) and the Exogenous Treatment Selection assumption (C.3).

Given that $\pi^d \geq 0$ and t, conditions (C.2) can be rewritten as:

$$\iiint \pi^1(y_0, y_1, y) I\{y_1 \neq y\} + \pi^0(y_0, y_1, y) I\{y_0 \neq y\}\, dy_0\, dy_1\, dy = 0$$

and we can rewrite relaxed SUTVA (Assumption 1α) as

$$\iiint \pi^1(y_0, y_1, y) I\{y_1 \neq y\} + \pi^0(y_0, y_1, y) I\{y_0 \neq y\}\, dy_0\, dy_1\, dy \leq \alpha. \tag{C.4}$$

The $ATE = E[Y^1 - Y^0]$ can be rewritten in terms of π^d in the following way:

$$ATE = \iiint (y_1 - y_0)(\pi^1 p^1 + \pi^0 p^0)\, dy_0\, dy_1\, dy. \tag{C.5}$$

In order to to find meaningful bounds without the ETS assumption, we will need bounded support of the outcome, suppose now that $y \in \mathcal{Y} \subset [y_{\min}, y_{\max}]$.

Lemma 8. *Under the Assumption 1α, the sharp bounds on the ATE are the following:*

$$\begin{aligned} &ATE \in [ATE^{LB}, ATE^{UB}] \\ ATE^{LB} &= \max\left\{ p^1\left(E[Y|D=1] - y_{\max}\right) + p^0\left(y_{\min} - E[Y|D=0]\right) - \alpha(y_{\max} - y_{\min}),\ -(y_{\max} - y_{\min}) \right\} \\ ATE^{UB} &= \min\left\{ p^1\left(E[Y|D=1] - y_{\min}\right) + p^0\left(y_{\max} - E[Y|D=0]\right) + \alpha(y_{\max} - y_{\min}),\ y_{\max} - y_{\min} \right\}. \end{aligned} \tag{C.6}$$

Proof of Lemma 8. We show the proof for the upper bound as the proof for the lower bound follows in an analogous way.

(i) Validity

$$
\begin{aligned}
ATE &= \iiint (y_1 - y_0)(\pi^1 p^1 + \pi^0 p^0)\, dy_0\, dy_1\, dy \\
&= p^1 \iiint (y_1 - y_0)\pi^1\, dy_0\, dy_1\, dy \\
&+ p^0 \iiint (y_1 - y_0)\pi^0\, dy_0\, dy_1\, dy \\
&= p^1 \iiint y_1 \pi^1\, dy_0\, dy_1\, dy - p^1 \iiint y_0 \pi^1\, dy_0\, dy_1\, dy \\
&+ p^0 \iiint y_1 \pi^0\, dy_0\, dy_1\, dy - p^0 \iiint y_0 \pi^0\, dy_0\, dy_1\, dy \\
&= p^1 \iiint (y_1 - y_0) \left[\pi^1 I\{y_1 = y\} + \pi^1 I\{y_1 \neq y\}\right] dy_0\, dy_1\, dy \\
&+ p^0 \iiint (y_1 - y_0) \left[\pi^0 I\{y_0 = y\} + \pi^0 I\{y_0 \neq y\}\right] dy_0\, dy_1\, dy \\
&\leq p^1 \left(E[Y|D=1] - y_{\min}\right) \\
&+ p^1 (y_{\max} - y_{\min}) \iiint \pi^1(y_0, y_1, y) I\{y_1 \neq y\}\, dy_0\, dy_1\, dy \\
&+ p^0 \left(y_{\max} - E[Y|D=0]\right) \\
&+ p^0 (y_{\max} - y_{\min}) \iiint \pi^0(y_0, y_1, y) I\{y_0 \neq y\}\, dy_0\, dy_1\, dy \\
&= p^1 \left(E[Y|D=1] - y_{\min}\right) + p^0 \left(y_{\max} - E[Y|D=0]\right) + \alpha(y_{\max} - y_{\min})
\end{aligned}
\tag{C.7}
$$

(ii) Sharpness

The following specification for π^d is compatible with Assumptions 1α, with the distribution of (Y, D) and achieves the ATE^{UB}. Note that for $\alpha \leq p^1 E[Y|D=1] + p^0 E[Y|D=0]$ there exists α_0, α_1 such that $\alpha_0 \leq p^0 E[Y|D=0]$, $\alpha_1 \leq p^1 E[Y|D=1]$ and $\alpha = \alpha_0 + \alpha_1$. For $\alpha \leq p^1 E[Y|D=1] + p^0 E[Y|D=0]$:

$$
\begin{aligned}
\pi^0(y_0, y_1, y) &= ((1-\alpha_0) I\{y_0 = y\} + \alpha_0 I\{y_0 = y_{\min}\}) \cdot I\{y_1 = y_{\max}\} \cdot f_Y(y|D=0), \\
\pi^1(y_0, y_1, y) &= I\{y_0 = y_{\min}\} \cdot ((1-\alpha_1) I\{y_1 = y\} + \alpha_1 I\{y_1 = y_{\max}\}) \cdot f_Y(y|D=1),
\end{aligned}
\tag{C.8}
$$

and for $\alpha > p^1 E[Y|D=1] + p^0 E[Y|D=0]$ we set $\alpha_0 = p^0 E[Y|D=0]$ and $\alpha_1 = p^1 E[Y|D=1]$.

□

References

Angrist, J., E. Bettinger, and M. Kremer (2006): "Long-term educational consequences of secondary school vouchers: Evidence from administrative records in Colombia," *The American Economic Review*, 847–862.

Chernozhukov, V., S. Lee, and A. M. Rosen (2013): "Intersection Bounds: Estimation and Inference," *Econometrica*, 81, 667–737.

Cole, S. R. and C. E. Frangakis (2009): "The consistency statement in causal inference: a definition or an assumption?" *Epidemiology*, 20, 3–5.

Cox, D. R. (1958): "Planning of experiments." .

Heckman, J. J., R. J. LaLonde, and J. A. Smith (1999): "The economics and econometrics of active labor market programs," *Handbook of labor economics*, 3, 1865–2097.

Hirano, K. and J. R. Porter (2012): "Impossibility results for nondifferentiable functionals," *Econometrica*, 1769–1790.

Horowitz, J. L. and C. F. Manski (1995): "Identification and robustness with contaminated and corrupted data," *Econometrica: Journal of the Econometric Society*, 281–302.

Kaido, H., F. Molinari, and J. Stoye (2016): "Confidence Intervals for Projections of Partially Identified Parameters," *arXiv preprint arXiv:1601.00934*.

Laffférs, L. (2015): "Bounding Average Treatment Effects using Linear Programming," Tech. rep., Working paper.

Lee, D. S. (2009): "Training, wages, and sample selection: Estimating sharp bounds on treatment effects," *The Review of Economic Studies*, 76, 1071–1102.

Rubin, D. B. (1980): "Randomization Analysis of Experimental Data: The Fisher Randomization Test Comment," *Journal of the American Statistical Association*, 75, 591–593.

VanderWeele, T. J. (2009): "Concerning the consistency assumption in causal inference," *Epidemiology*, 20, 880–883.