Point Identification of LATE with Two Imperfect Instruments[*]

Rui Wang[†]

June 5, 2024

Abstract

This paper provides point identification results for the local average treatment effect (LATE) using two imperfect instruments. The classical approach (Imbens and Angrist (1994)) establishes the identification of LATE via an instrument that satisfies exclusion, monotonicity, and independence. My paper uses two instruments but imposes weaker assumptions on both instruments. The first instrument is allowed to have homogeneous direct effects on the outcome, and the second instrument does not need to satisfy monotonicity. The direct effects of the first instrument can be identified through exogenous variation in the second instrument, and thus the local average treatment effect is point identified.

Keywords: local average treatment effect, instrumental variables, exclusion restriction, monotonicity, point identification

JEL: C31, C36

1 Introduction

Instrumental variables are widely used to estimate causal effects with endogenous treatment. When treatment effects are heterogeneous, Imbens and Angrist (1994) and Angrist, Imbens, and Rubin (1996) show that the local average treatment effect (LATE) is identified as the instrumental variable (IV) estimand with a valid instrument. The instrumental variable is required to exert no direct effects on the outcome (exclusion), weakly increase treatment status (monotonicity), and be independent of the potential outcome and potential treatment (independence).

In practice, it is a challenging task to find a valid instrument that satisfies all the assumptions simultaneously. In estimating returns to schooling, several papers such as Uusitalo (1999) use family background variables as an instrument; however, these variables may have direct effects on people's earnings via family education and thus violate the exclusion restriction. The exclusion assumption may also fail in encouragement experiments. For example, Hirano et al. (2000) study the effects of the flu vaccine on the prevalence of influenza and use random encouragement to take the vaccine as an instrument. Their paper shows that the encouragement to take the vaccine has direct effects on the outcome since it may remind people to take other actions to prevent the flu.

Studies such as Kitagawa (2015), Huber and Mellace (2015), Kédagni and Mourifie (2016), and Mourifié and Wan (2017) develop different methods to test the assumptions of the instrumental variable and reject the validity of some instruments. For example, college proximity is used as an instrument in Card (1993), Kling (2001), and Carneiro, Heckman, and Vytlacil (2011), but its validity is rejected by Huber and Mellace (2015) and Mourifié and Wan (2017).

Motivated by these findings, this paper proposes a new approach to identify LATE. This approach uses two instruments while imposing weaker assumptions on both instruments compared to the standard IV assumptions. The first instrument is allowed to violate the exclusion restriction, and the second instrument does not need to satisfy monotonicity. Therefore, the

2

first instrument can affect the outcome through both direct effects and treatment effects. The second instrument is introduced to separate the two effects. By exploiting exogenous variation in the second instrument, the direct effects of the first instrument are identified, and therefore LATE defined by the first instrument is point identified.

The paper develops estimators for LATE and the direct effects, and establishes their asymptotic properties. I compare this approach by using two imperfect instruments with the IV estimand using a single instrument via Monte Carlo simulations. The results show that the IV estimand can have a large bias with nonzero direct effects and that the bias increases when the direct effects increase. The method with two instruments presented in this paper performs uniformly regardless of the direct effects; thus, it has a more robust performance concerning violations of the exclusion restriction.

Here I discuss some potential choices of the two instruments in various applications. One example is the effect of participation in the food stamp program on health outcomes. [1] The first instrument can be an increase in the benefits of the program. Such an increase is likely to affect participation monotonically, but the increased benefit may directly affect health outcomes. The second instrument can be whether the benefit of the program is issued electronically or in hard copy. The delivery method is unlikely to have a direct effect on health outcomes, but it may not satisfy monotonicity, as some people prefer electronic delivery but others prefer hard-copy delivery.

The second example is the study of the effect of having a third child on mothers' labor supply, as studied in Angrist and Evans (1998). The first instrument can be whether there is a financial subsidy for having a third child. Financial support is likely to encourage a third child, but may directly affect people's incentive for labor participation. The second instrument could be whether the first two children have the same sex. This instrument seems unlikely to directly affect labor participation, but it may violate monotonicity since different people may prefer different sibling-sex composition.

---

[1]See, e.g., DeBono, Ross, and Berrang-Ford (2012), Kreider et al. (2012), and Gundersen, Kreider, and Pepper (2017).

Another example is estimating the effect of early achievement for children (e.g., kindergarten performance) on subsequent outcomes (e.g., earnings), which has been studied by Chetty et al. (2011). One potential instrument is the random assignment of teachers/classrooms. But as argued by Kolesár et al. (2015), the exclusion restriction might be violated because an experienced teacher may not only improve kindergarten performance, but also have direct effects on children's subsequent outcomes. Nevertheless, we can use a second instrument to separate the direct effects. One potential choice is the class size, which is shown to have no significant effects on earnings in Chetty et al. (2011), but may not satisfy monotonicity as class size effects could be heterogeneous.[2]

1.1 Related Literature

This paper contributes to the literature studying heterogeneous treatment effects with endogeneity using instrumental variables. Imbens and Angrist (1994), Angrist, Imbens, and Rubin (1996), and Heckman and Vytlacil (2005) identify treatment effects using one instrument that satisfies the conditions of exclusion, monotonicity, and independence simultaneously. My paper complements the literature by using two instruments while relaxing one assumption on each instrument.

Several papers relax the exclusion restriction under the heterogeneous treatment effects framework. Hirano et al. (2000) study different violations of exclusion restrictions for subgroups and apply parametric models and the Bayesian approach for inference. Flores and Flores-Lagunes (2013) derive partial identification for LATE by employing a weak monotonicity assumption of mean potential outcomes within or across subgroups. Mealli and Pacini (2013) study partial identification of intention-to-treat effects through a secondary outcome.

De Chaisemartin (2017) relaxes the monotonicity assumption and shows that the IV estimand estimates the local average treatment for a subgroup of compliers under a "compliers-

---

[2]For example, Maasoumi, Millimet, and Rangaprasad (2005) show opposite signs of class size effects for students below and above the median.

defiers" condition. He further provides sufficient conditions for the "compliers-defiers" condition to hold. Kedagni (2021) relaxes the independence assumption of the instrument with potential treatment. His paper establishes partial identification results by using an additional instrument that serves as a proxy of the first instrument.

The paper is also related to the literature that relaxes the exclusion restriction of the instrument in a linear regression model. Hahn and Hausman (2005) derive the bias of different estimators with direct effects. Nevo and Rosen (2012) provide partial identification for the model parameter under the assumptions of a correlation between the instrument and the error term. Conley, Hansen, and Rossi (2012) employ different assumptions on the effect of the instrument on the outcome to conduct inference for the model parameter. Kolesár et al. (2015) allow for direct effects and develop an estimator under an orthogonality condition of the direct effects.

The remainder of the paper is organized as follows. Section 2 presents the heterogeneous treatment effect framework. Section 3 derives the identification result. Section 4 develops an estimator for LATE and examines its finite sample performance via simulations. Section 5 studies an extension and Section 6 concludes.

2   Heterogeneous Treatment Effects Model

The analysis focuses on the heterogeneous treatment effects framework introduced in Imbens and Angrist (1994) and Angrist, Imbens, and Rubin (1996). Let $Y \in \mathcal{Y}$ denote an outcome, $D \in \{0, 1\}$ denote a binary treatment, and $Z \in \{0, 1\}$ denote a binary instrument. The objective is to learn the effects of treatment D on outcome Y, but the treatment can be endogenous. Instrument Z is used to address the endogeneity issue. Observed variables are (Y, D, Z).

Following Imbens and Angrist (1994) and Angrist, Imbens, and Rubin (1996), I use counterfactual variables to describe the data generating process. Let D_z denote the potential

treatment given the instrument $Z = z$ and $Y_{d,z}$ denote the potential outcome given the instrument and treatment $Z = z, D_z = d$. Let Y_d denote the potential outcome given $D = d$, which is given as $Y_d = Y_{d,1} Z + Y_{d,0}(1 - Z)$. Observed variables (Y, D) are generated by

$$D = D_1 Z + D_0(1 - Z),$$

$$Y = Y_1 D + Y_0(1 - D).$$

The population can be divided into four subgroups based on the potential treatments (D_1, D_0): always takers (AT): $D_1 = D_0 = 1$; compliers (CP): $D_z = z$ for $z \in \{0, 1\}$; never takers (NT): $D_1 = D_0 = 0$; and defiers (DF): $D_z = 1 - z$ for $z \in \{0, 1\}$.

The following summarizes the assumptions in Imbens and Angrist (1994) and Angrist, Imbens, and Rubin (1996).

Assumption 1 (IV Validity)**.**

(i) *Exclusion:* $Y_{d,1} = Y_{d,0} \equiv Y_d$ *for any* $d \in \{0, 1\}$;

(ii) *Monotonicity:* $D_1 \geq D_0$;

(iii) *Independence:* $Z \perp\!\!\!\perp (Y_{1,1}, Y_{1,0}, Y_{0,1}, Y_{0,0}, D_1, D_0)$;

(iv) *Relevance:* $\Pr(D_1 > D_0) > 0$;

(v) $0 < \Pr(Z = 1) < 1$.

Assumption 1 (i) requires instrument Z to have no direct effects on the potential outcome, (ii) indicates that the instrument weakly increases the potential treatment (no defier), (iii) refers to the independence of the instrument with all potential variables, (iv) guarantees the existence of compliers, and (v) needs nonzero variation in instrument Z.

Under Assumption 1, Imbens and Angrist (1994) show that the average treatment effect for compliers can be identified as the IV estimand:

$$\text{LATE} \equiv E[Y_1 - Y_0 \mid \text{CP}] = \frac{E[Y \mid Z = 1] - E[Y \mid Z = 0]}{E[D \mid Z = 1] - E[D \mid Z = 0]} \equiv \text{IV}. \tag{1}$$

This identification result for LATE relies on the validity of instrument Z in Assumption 1. However, it may be difficult to find a single instrument satisfying all conditions in Assumption 1. The identification result may fail if one of the assumptions does not hold.

Angrist, Imbens, and Rubin (1996) (Section 5) discuss the sensitivity of the IV estimand to deviations from the IV validity assumptions. The IV estimand involves both direct effects and treatment effects when the exclusion restriction fails, and it is the combination of treatment effects for compliers and defiers when the monotonicity condition is violated. This paper proposes a new approach to identify LATE by using two instruments jointly but requiring weaker assumptions on both instruments. The first instrument is allowed to violate the exclusion restriction, and the second instrument does not need to satisfy the monotonicity condition.

3   Identifying LATE with Two Invalid Instruments

This section focuses on identifying the local average treatment effect defined by instrument Z, while allowing instrument Z to have nonzero direct effects on the outcome. I introduce an additional binary instrument $W \in \{0, 1\}$ to identify the direct effects and LATE defined by instrument Z. My approach uses two instruments (Z, W) but relaxing one main assumption on both instruments. Instrument Z is allowed to have direct effects on the outcome, and instrument W does not need to satisfy monotonicity. Instrument W is assumed to have no direct effects on the outcome such that the potential outcome is not indexed by instrument W.

When the exclusion restriction is relaxed, instrument Z will affect the outcome through both direct effects and treatment effects. The direct effect is $Y_{d,1} - Y_{d,0}$ for each individual with $D = d$. Let $\rho_{G,d} = E[Y_{d,1} - Y_{d,0} \mid G]$ denote the average direct effect for group $G \in \{\mathrm{AT}, \mathrm{NT}, \mathrm{CP}\}$ given treatment $D = d$.

Next, I present assumptions on the two instruments (Z, W).

Assumption 2 (Instruments $Z\&W$)**.**

    (i) *Direct effects:* $\rho_{G,d} = \rho_d$ *for any* $G \in \{AT, NT, CP\}$, *where* ρ_d *is a unknown constant for* $d \in \{1, 0\}$;

   (ii) *Monotonicity:* $D_1 \geq D_0$;

  (iii) *Independence:* $(Z, W) \perp\!\!\!\perp Y_{d,z} \mid (D_1, D_0)$ *for any* $d, z \in \{0, 1\}$, *and* $Z \perp\!\!\!\perp (D_1, D_0) \mid W$;

  (iv) *Relevance:* $\Pr(D_1 > D_0) > 0$;

   (v) $0 < \Pr(Z = z, W = w) < 1$ *for any* $z, w \in \{0, 1\}$.

Assumption 2 (i) relaxes the exclusion restriction in Assumption 1 and allows instrument Z to have nonzero direct effects on the outcome, $\rho_d \neq 0$. The average of the direct effects is assumed to be homogeneous across different subgroups but may vary given different treatments, $\rho_1 \neq \rho_0$. One example for potential outcome $Y_{d,z}$ is given as $Y_{d,z} = \rho_d z + u_d$, where ρ_d is a unknown constant. Then the direct effect is ρ_d for each individual given treatment $D = d$. Section 5 studies an extension allowing for different direct effects across subgroups and provides partial identification of LATE. The rest of the conditions in Assumption 2 are similar to Assumption 1 except for introducing the independence condition of instrument W with the potential outcome.

Under the monotonicity condition of instrument Z, we can divide the population into three subgroups $\{AT, NT, CP\}$ as described in Section 2. The objective is to identify the average treatment effect for the compliers defined by instrument Z:

$$\text{LATE} = E[Y_1 - Y_0 \mid CP].$$

Under the assumption of no direct effects ($\rho_1 = \rho_0 = 0$), LATE is identified as

$$\text{LATE} = \text{IV}_1 \equiv \frac{E[Y \mid Z = 1, W = 1] - E[Y \mid Z = 0, W = 1]}{E[D \mid Z = 1, W = 1] - E[D \mid Z = 0, W = 1]}.$$

8

The IV_1 estimand is similar to the IV estimand except it is conditional on the additional variable W. However, the above identification result does not apply if Z has nonzero direct effects on the outcome. My paper uses the additional instrument $W \in \{0, 1\}$ to help identify the direct effects (ρ_1, ρ_0) and LATE.

The following describes a relevance condition of instrument W.

Assumption 3 (Instrument W). *Relevance:* $\Pr(G \mid W = 1) \Pr(CP \mid W = 0) \neq \Pr(G \mid W = 0) \Pr(CP \mid W = 1)$ *for any* $G \in \{AT, NT\}$.

Assumption 3 is a relevance condition of instrument W. It requires that instrument W is correlated with the potential treatment such that the size of always takers (or never takers) relative to compliers varies when instrument W changes. Assumption 3 does not impose monotonicity on instrument W, so instrument W can affect treatment in any direction. This assumption is testable since the conditional probability of all types $G \in \{AT, NT, CP\}$ is identified, as shown in Appendix A.1.

Section 1 provides some examples of the two instruments in various applications. When estimating the effects of the food stamp program on health outcomes, the first instrument can be an increase in the benefits of the program. The increased benefits may directly affect health outcomes so the exclusion restriction fails. The second instrument can be whether the benefit of the program is issued electronically or in hard copy. This instrument may not satisfy the monotonicity condition since some people prefer electronic delivery but others prefer hard-copy delivery.

Theorem 1. *Under Assumptions 2-3, direct effects* (ρ_1, ρ_0) *and local average treatment effect* LATE *are point identified.*

When the exclusion restriction is relaxed, instrument Z can induce both treatment effects by switching the treatment status and direct effects on the outcome. We are unable to distinguish treatment effects and direct effects without further information, so the standard identification result for treatment effects no longer applies. This paper uses additional in-

strument W, which can serve as an instrument for the imperfect instrument Z, to address instrument Z's violation of the exclusion restriction. By exploiting variation in instrument W, this approach can identify the direct effects of instrument Z on outcome Y. Then the local average treatment effect is identified by subtracting the direct effects.

Theorem 1 provides an alternative approach to identify treatment effects. This approach uses two instruments while relaxing one of the assumptions on the two instruments. This method can be applied to scenarios where there are multiple options of instruments but the instruments are imperfect in different dimensions. The availability of multiple instruments is discussed in the literature including Card (2001), Hausman et al. (2012), and Kolesár et al. (2015). Moreover, instrument W can help identify the direct effects of instrument Z on the treated and untreated outcome, meaning that this approach can be used to test whether instrument Z has direct effects on the outcome.

4   Estimation

This section provides an estimation method for direct effects (ρ_1, ρ_0) and local average treatment effect LATE. Suppose that we have an $i.i.d$ sample $(Y_i, D_i, Z_i, W_i)_{i=1}^N$. As shown in Appendix A.1, LATE is identified as the IV_1 estimand minus a weighted average of the two direct effects (ρ_1, ρ_0):

$$\text{LATE} = \text{IV}_1 - \rho_1 w_1^\rho - \rho_0 w_0^\rho,$$

where $w_1^\rho = \frac{\Pr(\text{AT}|W=1)}{\Pr(\text{CP}|W=1)} + \Pr(Z=0)$, $w_0^\rho = \frac{\Pr(\text{NT}|W=1)}{\Pr(\text{CP}|W=1)} + \Pr(Z=1)$, and the formula for the two direct effects (ρ_1, ρ_0) is described below.

To estimate LATE, we need to estimate all terms in the above expression of LATE. The first term IV_1 is given as

$$\begin{aligned}
\text{IV}_1 &= \frac{E[Y \mid Z=1, W=1] - E[Y \mid Z=0, W=1]}{E[D \mid Z=1, W=1] - E[D \mid Z=0, W=1]} \\
&= \frac{E[YZW]E[W] - E[YW]E[ZW]}{E[DZW]E[W] - E[DW]E[ZW]}.
\end{aligned}$$

The estimator for the IV_1 estimand can be developed by replacing the population expectation with the sample mean:

$$\widehat{\text{IV}}_1 = \frac{\sum_i (Y_i Z_i W_i) \sum W_i - \sum_i (Y_i W_i) \sum_i (Z_i W_i)}{\sum_i (D_i Z_i W_i) \sum W_i - \sum_i (D_i W_i) \sum_i (Z_i W_i)}.$$

Now I construct estimators for the two weights (w_1^ρ, w_0^ρ). Let $\widehat{\Pr}(G \mid w)$ denote the estimator for the conditional probability $\Pr(G \mid w)$ of the three subgroups $G \in \{\text{AT}, \text{NT}, \text{CP}\}$ given $W = w$, constructed as follows:

$$\widehat{\Pr}(\text{AT} \mid w) = \frac{\sum_i D_i (1 - Z_i) \mathbb{1}\{W_i = w\}}{\sum_i (1 - Z_i) \mathbb{1}\{W_i = w\}},$$

$$\widehat{\Pr}(\text{NT} \mid w) = \frac{\sum_i (1 - D_i) Z_i \mathbb{1}\{W_i = w\}}{\sum_i Z_i \mathbb{1}\{W_i = w\}},$$

$$\widehat{\Pr}(\text{CP} \mid w) = \frac{\sum_i D_i Z_i \mathbb{1}\{W_i = w\}}{\sum_i Z_i \mathbb{1}\{W_i = w\}} - \frac{\sum_i D_i (1 - Z_i) \mathbb{1}\{W_i = w\}}{\sum_i (1 - Z_i) \mathbb{1}\{W_i = w\}}.$$

Then the two weights (w_1^ρ, w_0^ρ) can be estimated as

$$\hat{w}_1^\rho = \frac{\widehat{\Pr}(\text{AT} \mid 1)}{\widehat{\Pr}(\text{CP} \mid 1)} + 1 - \bar{Z}, \qquad \hat{w}_0^\rho = \frac{\widehat{\Pr}(\text{NT} \mid 1)}{\widehat{\Pr}(\text{CP} \mid 1)} + \bar{Z},$$

where $\bar{Z} = \frac{1}{N} \sum_i Z_i$.

We only need to develop estimators for the two direct effects (ρ_1, ρ_0). I focus on the estimator for direct effect ρ_1, and the idea also applies to ρ_0. As shown in Appendix A.1, direct effect ρ_1 is identified as

$$\rho_1 = \frac{r_1(1) \Pr(\text{CP} \mid 0) - r_1(0) \Pr(\text{CP} \mid 1)}{\Pr(\text{AT} \mid 1) \Pr(\text{CP} \mid 0) - \Pr(\text{AT} \mid 0) \Pr(\text{CP} \mid 1)},$$

where $r_1(w)$ is defined as

$$r_1(w) \equiv E[YD \mid Z = 1, w] - E[YD \mid Z = 0, w].$$

The estimator for $r_1(w)$ is developed as follows:

$$\hat{r}_1(w) = \frac{\sum_i Y_i D_i Z_i \mathbb{1}\{W_i = w\}}{\sum_i Z_i \mathbb{1}\{W_i = w\}} - \frac{\sum_i Y_i D_i (1 - Z_i) \mathbb{1}\{W_i = w\}}{\sum_i (1 - Z_i) \mathbb{1}\{W_i = w\}}.$$

Then estimator $\hat{\rho}_1$ for direct effect ρ_1 can be established by replacing all terms with their estimators:

$$\hat{\rho}_1 = \frac{\hat{r}_1(1)\widehat{\Pr}(\mathrm{CP} \mid 0) - \hat{r}_1(0)\widehat{\Pr}(\mathrm{CP} \mid 1)}{\widehat{\Pr}(\mathrm{AT} \mid 1)\widehat{\Pr}(\mathrm{CP} \mid 0) - \widehat{\Pr}(\mathrm{AT} \mid 0)\widehat{\Pr}(\mathrm{CP} \mid 1)}.$$

Estimator $\hat{\rho}_0$ for ρ_0 can be established similarly, so it is omitted here. Local average treatment effect LATE is estimated as

$$\widehat{\mathrm{LATE}} = \widehat{\mathrm{IV}}_1 - \hat{\rho}_1 \hat{w}_1^\rho - \hat{\rho}_0 \hat{w}_0^\rho.$$

The asymptotic properties of $\widehat{\mathrm{LATE}}$ and the two direct effects $(\hat{\rho}_1, \hat{\rho}_0)$ are derived in Appendix A.2. The estimators $\widehat{\mathrm{LATE}}$ and $(\hat{\rho}_1, \hat{\rho}_0)$ are $\sqrt{N}$ consistent assuming all variances and covariances are finite, and the expressions of asymptotic variances are shown in Appendix A.2.

4.1 Simulation Study

This section examines the finite sample performance of estimator $\widehat{\mathrm{LATE}}$ by using two instruments (Z, W) via Monte Carlo simulations. I compare the estimator in this paper with the IV estimator in Imbens and Angrist (1994) and show that the method in this paper works more robustly when the direct effects are nonzero. The IV_1 estimator (conditional on W) performs slightly worse than the IV estimator, so I display the results of the IV estimator for comparison.

The simulation setup is as follows. Instrument Z follows the Bernoulli distribution with probability $p = 0.5$. Potential treatment D_z is $D_z = \mathbb{1}\{z \geq \epsilon\}$, and observed treatment D is $D = D_1 Z + D_0(1 - Z)$. Potential outcome Y_d is given $Y_d = a_d + \rho_d Z + u_d$ for $d \in \{0, 1\}$,

where $a_1 = 1$, $a_0 = 0$. Observed outcome Y is given $Y = Y_1 D + Y_0(1 - D)$. I consider four different cases of direct effects: $\rho_1 = \rho_0 = \rho \in \{0, 0.5, 1, -1\}$. Instrument W is given as $W = \mathbb{1}\{v \leq D_1 + D_0\}$.

Latent variables (ϵ, u_1) follow a standard multivariate distribution with correlation $c = 0.5$, and this correlation captures the endogeneity of treatment D. Error terms u_0 and v follow a standard normal distribution and are independent of all other variables. I consider sample size $N = \{1000, 4000, 16000\}$, and the repetition number is $B = 5000$.

The assumptions of the two instruments (Z, W) are satisfied under this setup. Instrument Z is allowed to have nonzero direct effects on outcome $\rho_d \neq 0$, and the direct effects are the same across subgroups. The independence condition of Z is satisfied since it is independent of all variables. The monotonicity and relevance conditions are satisfied by the definition of potential treatment D_z. Instrument W is independent of the potential outcome given the potential treatment since error term v is independent of all variables. The relevance of instrument W is satisfied because W depends on the potential outcome.

Let $\hat{\theta}_{zw}$ denote the estimator by using the two instruments (Z, W) in this paper, and let $\hat{\theta}_z$ denote the IV estimator by using only instrument Z in Imbens and Angrist (1994). Let $(\hat{\rho}_1, \hat{\rho}_0)$ denote the estimators for the direct effects (ρ_1, ρ_0) in this paper. To compare the two estimators $(\hat{\theta}_{zw}, \hat{\theta}_z)$, I report four evaluations of the two estimators: bias, standard deviation (SD), root mean-squared error (rMSE), and median of absolute deviation (MAD). Under the simulation setup, the true local average treatment effect θ_0 is given as

$$\theta_0 = E[Y_1 - Y_0 \mid \mathrm{CP}] = E[a_1 - a_0 + (\rho_1 - \rho_0)Z + u_1 - u_0 \mid \epsilon \in (0, 1)]$$
$$= 1 + \frac{(\phi(0) - \phi(1))}{2(\Phi(1) - \Phi(0))},$$

where ϕ and Φ denote the PDF and CDF of the standard normal distribution, respectively.

Table 1 presents the performance of the two estimators $\hat{\theta}_{zw}$ and $\hat{\theta}_z$ under different direct effects $\rho \in \{0, 0.5, 1, -1\}$ and sample size $N \in \{1000, 4000, 16000\}$. Estimator $\hat{\theta}_z$ performs better when the direct effect is zero but can have a large bias with nonzero direct effects.

13

The bias increases when the direct effect is larger, and it does not decrease even when the sample size increases. When the direct effect is negative $\rho < 0$, the bias becomes negative and estimator $\hat{\theta}_z$ may have the wrong signs of the true treatment effect. Estimator $\hat{\theta}_{zw}$ by using two instruments performs uniformly under different direct effects, and it performs better than estimator $\hat{\theta}_z$ with nonzero direct effects. This pattern becomes more significant when the sample size increases. The comparison in Table 1 shows that estimator $\hat{\theta}_{zw}$ by using two instruments has a more robust performance for nonzero direct effects.

Table 1: Performance Comparisons of $\hat{\theta}_{zw}$ and $\hat{\theta}_z$

N	ρ	$\hat{\theta}_{zw}$				$\hat{\theta}_z$			
		Bias	SD	rMSE	MAD	Bias	SD	rMSE	MAD
1000	0	0.025	0.504	0.505	0.396	0.006	0.183	0.184	0.146
	0.5	0.025	0.504	0.505	0.396	1.481	0.245	1.501	1.481
	1	0.025	0.504	0.505	0.396	2.956	0.342	2.976	2.956
	-1	0.025	0.504	0.505	0.396	-2.944	0.269	2.956	2.944
4000	0	0.009	0.230	0.230	0.183	-0.001	0.092	0.092	0.073
	0.5	0.009	0.230	0.230	0.183	1.466	0.119	1.470	1.466
	1	0.009	0.230	0.230	0.183	2.932	0.165	2.936	2.932
	-1	0.009	0.230	0.230	0.183	-2.933	0.133	2.936	2.933
16000	0	0.001	0.113	0.113	0.090	0.001	0.046	0.046	0.036
	0.5	0.001	0.113	0.113	0.090	1.465	0.060	1.466	1.465
	1	0.001	0.113	0.113	0.090	2.930	0.083	2.931	2.930
	-1	0.001	0.113	0.113	0.090	-2.929	0.066	2.930	2.929

Table 2 presents the performance of the estimators $(\hat{\rho}_1, \hat{\rho}_0)$ for the direct effects. The performance of $(\hat{\rho}_1, \hat{\rho}_0)$ does not depend on the true direct effects, so I only report the results under different sample sizes. When the sample size increases, the bias and deviation of $(\hat{\rho}_1, \hat{\rho}_0)$ shrink dramatically.

Appendix A.4 presents more simulation results about the two estimators $(\hat{\theta}_{zw}, \hat{\theta}_z)$ under different probabilities of the three subgroups.

Table 2: Performance of Direct Effects $\hat{\rho}_1$ and $\hat{\rho}_0$

N	$\hat{\rho}_1$				$\hat{\rho}_0$			
	Bias	SD	rMSE	MAD	Bias	SD	rMSE	MAD
1000	-0.015	0.210	0.210	0.160	0.006	0.351	0.351	0.275
4000	-0.004	0.095	0.095	0.075	-0.003	0.168	0.168	0.133
16000	0.000	0.047	0.047	0.038	-0.001	0.083	0.083	0.066

5   Extension

Section 3 establishes point identification results of LATE when the average of direct effects is assumed to be homogeneous across subgroups. This assumption can allow for heterogeneous direct effects within a subgroup but assumes same direct effects across subgroups. This section further relaxes this assumption and provides partial identification results under heterogeneous direct effects across subgroups.

Recall that $\rho_{G,d} = E[Y_{d,1} - Y_{d,0} \mid G]$ denote the average direct effect for group $G \in \{\mathrm{AT}, \mathrm{NT}, \mathrm{CP}\}$ given treatment $D = d$. I consider that the heterogeneity in direct effects between subgroups can be bounded by a known number.

Assumption 4. *Direct effects:* $|\rho_{\mathrm{CP},1} - \rho_{\mathrm{AT},1}| \leq k_1$ *and* $|\rho_{\mathrm{CP},0} - \rho_{\mathrm{NT},0}| \leq k_0$, *where* $k_1, k_0 \geq 0$ *are known constants.*

Assumption 4 relaxes Assumption 2 (i) and allows heterogeneous direct effects across different subgroups. This assumption requires that the difference in direct effects between subgroups is not too large and can be bounded by a known number k_d. The information about k_d depends on specific applications. For example, when the instrument is whether there is an increase (or decrease) in the benefits of a social program, the direct effect on people who did not participate in the program is zero $\rho_{\text{NT},0} = \rho_{\text{CP},0} = 0$ so that $k_0 = 0$. The value of k_1 can be developed when the support of the outcome is bounded such as binary outcomes.

As shown in Section 4, when the average of direct effects is the same across different groups $k_1 = k_0 = 0$, LATE is identified as

$$\text{LATE} = \text{IV}_1 - \rho_1 w_1^\rho - \rho_0 w_0^\rho \equiv \widetilde{\text{IV}}.$$

The next proposition derives sharp bounds on LATE.

Proposition 1. *Under Assumptions 2 (ii)-(iv) and Assumptions 3-4, the sharp bounds for* LATE *are given as follows:*

$$\text{LATE} \in \left[\widetilde{\text{IV}} - k_1 \Pr(Z = 0) - k_0 \Pr(Z = 1), \widetilde{\text{IV}} + k_1 \Pr(Z = 0) + k_0 \Pr(Z = 1) \right].$$

Proposition 1 shows that LATE can be still bounded by using two instruments under heterogeneous direct effects. The bounds are tighter when the heterogeneity (k_1, k_0) in direct effects between different subgroups is smaller, and point identification is achieved when the heterogeneity is zero. The bounds on LATE can be established similarly and be further tightened if the direction of the difference in direct effects is known such as $0 \le \rho_{\text{CP},1} - \rho_{\text{AT},1} \le k_1$ and $0 \le \rho_{\text{CP},0} - \rho_{\text{NT},0} \le k_0$.

6 Conclusion

This paper proposes a new approach to point identify LATE by using two instruments while imposing weaker assumptions on both instruments compared to Imbens and Angrist (1994). The first instrument is allowed to violate the exclusion restriction, so it can have nonzero direct effects on the outcome. Then the IV estimand involves both treatment effects and direct effects, so the standard identification result does not apply. This paper uses an additional imperfect instrument, which does not need to satisfy monotonicity. By exploiting variation in the second instrument, we can identify the direct effects of the first instrument and then identify LATE. Based on the identification results, an estimator for LATE is developed and it is shown to perform more robustly than the IV estimator with nonzero direct effects.

This paper relaxes different assumptions for the two instruments to achieve point identification of treatment effects. It would be worthwhile to investigate the identifying power of multiple instruments when they violate the same assumption such as the exclusion restriction. Moreover, the paper considers two instruments and relaxes one main assumption for each of the two instruments. It would be interesting to explore whether point identification can be achieved under weaker conditions on instruments when there are more than two instruments available.

References

J. D. Angrist and W. N. Evans. Children and their parents' labor supply: Evidence from exogenous variation in family size. *American Economic Review*, 88(3):450–477, 1998.

J. D. Angrist, G. W. Imbens, and D. B. Rubin. Identification of causal effects using instrumental variables. *Journal of the American statistical Association*, 91(434):444–455, 1996.

D. Card. Using geographic variation in college proximity to estimate the return to schooling. National Bureau of Economic Research Cambridge, Mass., USA, 1993.

D. Card. Estimating the return to schooling: Progress on some persistent econometric problems. *Econometrica*, 69(5):1127–1160, 2001.

P. Carneiro, J. J. Heckman, and E. J. Vytlacil. Estimating marginal returns to education. *American Economic Review*, 101(6):2754–81, 2011.

R. Chetty, J. N. Friedman, N. Hilger, E. Saez, D. W. Schanzenbach, and D. Yagan. How does your kindergarten classroom affect your earnings? evidence from project star. *The Quarterly journal of economics*, 126(4):1593–1660, 2011.

T. G. Conley, C. B. Hansen, and P. E. Rossi. Plausibly exogenous. *Review of Economics and Statistics*, 94(1):260–272, 2012.

C. De Chaisemartin. Tolerating defiance? local average treatment effects without monotonicity. *Quantitative Economics*, 8(2):367–396, 2017.

N. L. DeBono, N. A. Ross, and L. Berrang-Ford. Does the food stamp program cause obesity? a realist review and a call for place-based research. *Health & Place*, 18(4):747–756, 2012.

C. A. Flores and A. Flores-Lagunes. Partial identification of local average treatment effects with an invalid instrument. *Journal of Business & Economic Statistics*, 31(4):534–545, 2013.

C. Gundersen, B. Kreider, and J. V. Pepper. Partial identification methods for evaluating food assistance programs: a case study of the causal impact of snap on food insecurity. *American Journal of Agricultural Economics*, 99(4):875–893, 2017.

J. Hahn and J. Hausman. Estimation with valid and invalid instruments. *Annales d'Economie et de Statistique*, pages 25–57, 2005.

J. A. Hausman, W. K. Newey, T. Woutersen, J. C. Chao, and N. R. Swanson. Instrumental variable estimation with heteroskedasticity and many instruments. *Quantitative Economics*, 3(2):211–255, 2012.

J. J. Heckman and E. Vytlacil. Structural equations, treatment effects, and econometric policy evaluation 1. *Econometrica*, 73(3):669–738, 2005.

K. Hirano, G. W. Imbens, D. B. Rubin, and X.-H. Zhou. Assessing the effect of an influenza vaccine in an encouragement design. *Biostatistics*, 1(1):69–88, 2000.

M. Huber and G. Mellace. Testing instrument validity for late identification based on inequality moment constraints. *Review of Economics and Statistics*, 97(2):398–411, 2015.

G. W. Imbens and J. D. Angrist. Identification and estimation of local average treatment effects. *Econometrica*, 62(2):467–475, 1994.

D. Kedagni. Identifying treatment effects in the presence of confounded types. *Journal of Econometrics*, 2021.

D. Kédagni and I. Mourifie. Testing the iv zero-covariance assumption and identification with an invalid instrument. *Available at SSRN 2861669*, 2016.

T. Kitagawa. A test for instrument validity. *Econometrica*, 83(5):2043–2063, 2015.

J. R. Kling. Interpreting instrumental variables estimates of the returns to schooling. *Journal of Business & Economic Statistics*, 19(3):358–364, 2001.

M. Kolesár, R. Chetty, J. Friedman, E. Glaeser, and G. W. Imbens. Identification and inference with many invalid instruments. *Journal of Business & Economic Statistics*, 33 (4):474–484, 2015.

B. Kreider, J. V. Pepper, C. Gundersen, and D. Jolliffe. Identifying the effects of snap (food stamps) on child health outcomes when participation is endogenous and misreported. *Journal of the American Statistical Association*, 107(499):958–975, 2012.

E. Maasoumi, D. L. Millimet, and V. Rangaprasad. Class size and educational policy: Who benefits from smaller classes? *Econometric Reviews*, 24(4):333–368, 2005.

F. Mealli and B. Pacini. Using secondary outcomes and covariates to sharpen inference in instrumental variable settings. *Journal of the American Statistical Association*, 108: 1120–1131, 2013.

I. Mourifié and Y. Wan. Testing local average treatment effect assumptions. *Review of Economics and Statistics*, 99(2):305–313, 2017.

A. Nevo and A. M. Rosen. Identification with imperfect instruments. *Review of Economics and Statistics*, 94(3):659–671, 2012.

R. Uusitalo. Return to education in finland. *Labour Economics*, 6(4):569–580, 1999.

A Appendix

A.1 Proof of Theorem 1

Proof. I first look at the expression of LATE, and divide it into two groups $Z = 1$ and $Z = 0$ as follows:

$$
\begin{aligned}
\text{LATE} =& E[Y_1 - Y_0 \mid \text{CP}] \\
=& E[Y_{1,1} - Y_{0,1} \mid \text{CP}, Z = 1] \Pr(Z = 1) + E[Y_{1,0} - Y_{0,0} \mid \text{CP}, Z = 0] \Pr(Z = 0) \\
=& E[Y_{1,1} - Y_{0,1} \mid \text{CP}] \Pr(Z = 1) + E[Y_{1,0} - Y_{0,0} \mid \text{CP}] \Pr(Z = 0).
\end{aligned}
$$

The above condition holds by the independence condition of instrument Z in Assumption 2. Using the condition $E[Y_{d,1} - Y_{d,0} \mid \text{CP}] = \rho_d$ to substitute $E[Y_{0,1} \mid \text{CP}]$ and $E[Y_{1,0} \mid \text{CP}]$ leads to the following implication:

$$
\text{LATE} = E[Y_{1,1} - Y_{0,0} \mid \text{CP}] - \rho_1 \Pr(Z = 0) - \rho_0 \Pr(Z = 1).
$$

To prove LATE is identified, we need to show that the two direct effects (ρ_1, ρ_0) and $(E[Y_{11} \mid \mathrm{CP}], E[Y_{0,0} \mid \mathrm{CP}])$ are identified. To prove it, I first show that the conditional probability of the three subgroups $\{\mathrm{AT}, \mathrm{NT}, \mathrm{CP}\}$ given $W = w$ is identified. When $Z = 0$, only always takers (AT) are treated so that the probability of always takers is identified as

$$\Pr(D = 1 \mid Z = 0, w) = \Pr(D = 1 \mid \mathrm{AT}, Z = 0, w) \Pr(\mathrm{AT} \mid Z = 0, w)$$

$$= \Pr(\mathrm{AT} \mid w).$$

The last equality holds since the probability of being treated conditional on always takers is one and instrument Z is independent of potential treatments given W in Assumption 2.

Similarly, the conditional probability of never takers and compliers given $W = w$ can be derived as

$$\Pr(\mathrm{NT} \mid w) = \Pr(D = 0 \mid Z = 1, w);$$

$$\Pr(\mathrm{CP} \mid w) = \Pr(D = 1 \mid Z = 1, w) - \Pr(D = 1 \mid Z = 0, w).$$

Now we are ready to show that the two direct effects (ρ_1, ρ_0) and $(E[Y_{11} \mid \mathrm{CP}], E[Y_{0,0} \mid \mathrm{CP}])$ are identified by using variation in instrument W. The expectation of YD conditional on $(Z = 1, W = w)$ can be the expressed as a mixture of always takers and compliers:

$$E[YD \mid Z = 1, w]$$
$$= E[Y_{1,1} \mid \mathrm{AT}, Z = 1, w] \Pr(\mathrm{AT} \mid Z = 1, w) + E[Y_{1,1} \mid \mathrm{CP}, Z = 1, w] \Pr(\mathrm{CP} \mid Z = 1, w) \quad (2)$$
$$= E[Y_{1,1} \mid \mathrm{AT}] \Pr(\mathrm{AT} \mid w) + E[Y_{1,1} \mid \mathrm{CP}] \Pr(\mathrm{CP} \mid w).$$

The above condition holds by the independence conditions of the two instruments (Z, W) in Assumptions 2.

Similarly, the expectation of YD given $(Z = 0, W = w)$ can be expressed as

$$E[YD \mid Z = 0, w] = E[Y_{1,0} \mid \mathrm{AT}] \Pr(\mathrm{AT} \mid w). \quad (3)$$

Taking the difference between (2) and (3) leads to the following condition: for any $w \in \{0, 1\}$,

$$r_1(w) \equiv E[YD \mid Z = 1, w] - E[YD \mid Z = 0, w]$$
$$= \rho_1 \Pr(\text{AT} \mid w) + E[Y_{1,1} \mid \text{CP}] \Pr(\text{CP} \mid w). \tag{4}$$

Equation (4) comes from the condition $E[Y_{d,1} - Y_{d,0} \mid \text{AT}] = \rho_d$ in Assumption 2 (i). Since condition (4) holds for any $w \in \{0, 1\}$, using variation in instrument W can identify direct effect ρ_1 and $E[Y_{1,1} \mid \text{CP}]$ as follows:

$$\rho_1 = \frac{r_1(1) \Pr(\text{CP} \mid 0) - r_1(0) \Pr(\text{CP} \mid 1)}{\Pr(\text{AT} \mid 1) \Pr(\text{CP} \mid 0) - \Pr(\text{AT} \mid 0) \Pr(\text{CP} \mid 1)},$$
$$E[Y_{1,1} \mid \text{CP}] = \frac{r_1(1) - \rho_1 \Pr(\text{AT} \mid 1)}{\Pr(\text{CP} \mid 1)}.$$

The relevance condition of instrument W in Assumption 3 guarantees that the denominator of direct effect ρ_1 is nonzero: $\Pr(\text{AT} \mid 1) \Pr(\text{CP} \mid 0) - \Pr(\text{AT} \mid 0) \Pr(\text{CP} \mid 1) \neq 0$. The relevance condition of instrument Z in Assumption 2 implies that there exists $w \in \{0, 1\}$ such that $\Pr(\text{CP} \mid w) > 0$. For simplicity, I assume that $\Pr(\text{CP} \mid 1) > 0$.

Similarly, I use the expectation of $Y(1 - D)$ under different values of (Z, W) to identify direct effect ρ_0 and $E[Y_{0,0} \mid \text{CP}]$. Let $r_0(w)$ be defined as

$$r_0(w) = E[Y(1 - D) \mid Z = 1, w] - E[Y(1 - D) \mid Z = 0, w].$$

By using variation in $r_0(w)$ with respect to w, direct effect ρ_0 and $E[Y_{0,0} \mid \text{CP}]$ can be identified as follows:

$$\rho_0 = \frac{r_0(1) \Pr(\text{CP} \mid 0) - r_0(0) \Pr(\text{CP} \mid 1)}{\Pr(\text{NT} \mid 1) \Pr(\text{CP} \mid 0) - \Pr(\text{NT} \mid 0) \Pr(\text{CP} \mid 1)},$$
$$E[Y_{0,0} \mid \text{CP}] = -\frac{r_0(1) - \rho_0 \Pr(\text{NT} \mid 1)}{\Pr(\text{CP} \mid 1)}.$$

Therefore, LATE is identified since the two direct effects (ρ_1, ρ_0) and $(E[Y_{1,1} \mid \text{CP}], E[Y_{0,0} \mid \text{CP}])$ are shown to be identified. Plugging into the expression of $(E[Y_{1,1} \mid \text{CP}], E[Y_{0,0} \mid \text{CP}])$

into the formula of LATE leads to following expression:

$$\text{LATE} = E[Y_{1,1} - Y_{0,0} \mid \text{CP}] - \rho_1 \Pr(Z=0) - \rho_0 \Pr(Z=1)$$

$$= \frac{r_1(1) + r_0(1)}{\Pr(\text{CP} \mid 1)} - \rho_1 \left(\frac{\Pr(\text{AT} \mid 1)}{\Pr(\text{CP} \mid 1)} + \Pr(Z=0) \right) - \rho_0 \left(\frac{\Pr(\text{NT} \mid 1)}{\Pr(\text{CP} \mid 1)} + \Pr(Z=1) \right)$$

$$\equiv \text{IV}_1 - \rho_1 w_1^\rho - \rho_0 w_0^\rho,$$

where $\text{IV}_1 = \frac{E[Y|Z=1,W=1] - E[Y|Z=0,W=1]}{E[D|Z=1,W=1] - E[D|Z=0,W=1]}$, $w_1^\rho = \frac{\Pr(\text{AT}|1)}{\Pr(\text{CP}|1)} + \Pr(Z=0)$, and $w_0^\rho = \frac{\Pr(\text{NT}|1)}{\Pr(\text{CP}|1)} + \Pr(Z=1)$. Therefore, LATE is identified as all terms in the above expression are shown to be identified.

$\square$

A.2 Asymptotic Properties of $\widehat{\text{LATE}}$ and $(\hat{\rho}_1, \hat{\rho}_0)$

This section derives the asymptotic properties of estimator $\widehat{\text{LATE}}$ for the local average treatment effect and the two estimators $(\hat{\rho}_1, \hat{\rho}_0)$ for the two direct effects. Suppose that the variances of all variables (Y, D, Z, W) are finite, and the variances of the product of any two, three, and four variables are finite.

As shown in Section 4 , the expression for estimator $\widehat{\text{LATE}}$ is given as follows:

$$\widehat{\text{LATE}} = \widehat{\text{IV}}_1 - \hat{\rho}_1 \hat{w}_1^\rho - \hat{\rho}_0 \hat{w}_0^\rho.$$

The analysis proceeds by deriving the asymptotic properties of the above terms $(\widehat{\text{IV}}_1, \hat{w}_1^\rho, \hat{w}_0^\rho, \hat{\rho}_1, \hat{\rho}_0)$ respectively, then the asymptotic property of $\widehat{\text{LATE}}$ can be derived accordingly.

Let μ_Y denote the expectation of the random variable Y. Let $\phi_i^Y = Y_i - \mu_Y$ for any random variable Y, let $\phi_i^{YZ} = Y_i Z_i - E[YZ]$ for any two random variables (Y, Z), and let $\phi_i^{YZW} = Y_i Z_i W_i - E[YZW]$ for any three random variables (Y, Z, W).

I first look at estimator $\widehat{\text{IV}}_1$. The numerator of estimator $\widehat{\text{IV}}_1$ is given as

$$\frac{1}{N}\sum_i (Y_i Z_i W_i)\bar{W} - \frac{1}{N}\sum_i (Y_i W_i)\frac{1}{N}\sum_i (Z_i W_i) - (E[YZW]E[W] - E[YW]E[ZW])$$
$$= \frac{1}{N}\sum_i (E[YZW]\phi_i^W + \mu_W \phi_i^{YZW} - E[YW]\phi_i^{ZW} - E[ZW]\phi_i^{YW}) + O_p\left(\frac{1}{N}\right).$$

The last equality holds by applying Taylor expansion and the fact that $\frac{1}{N}\sum_i (Y_i Z_i W_i) - E[YZW] = O_p(1/\sqrt{N})$, $\bar{W} - \mu_W = O_p(1/\sqrt{N})$, $\frac{1}{N}\sum_i (Y_i W_i) - E[YW] = O_p(1/\sqrt{N})$, and $\frac{1}{N}\sum_i (Z_i W_i) - E[ZW] = O_p(1/\sqrt{N})$.

Let $\phi_i^{\text{IV}_N}$ denote the influence function of the numerator term $\frac{1}{N}\sum_i (Y_i Z_i W_i)\bar{W} - \frac{1}{N}\sum_i (Y_i W_i)\frac{1}{N}\sum_i (Z_i W_i)$

defined as

$$\phi_i^{\text{IV}_N} = E[YZW]\phi_i^W + \mu_W \phi_i^{YZW} - E[YW]\phi_i^{ZW} - E[ZW]\phi_i^{YW}.$$

Similarly, the influence function of the denominator term $\frac{1}{N}\sum_i (D_i Z_i W_i)\bar{W} - \frac{1}{N}\sum_i (D_i W_i)\frac{1}{N}\sum_i (Z_i W_i)$ is derived as

$$\phi_i^{\text{IV}_D} = E[DZW]\phi_i^W + \mu_W \phi_i^{DZW} - E[DW]\phi_i^{ZW} - E[ZW]\phi_i^{DW}.$$

Then by applying Taylor expansion, the asymptotic property of $\widehat{\text{IV}}_1$ is derived as follows:

$$\widehat{\text{IV}}_1 - \text{IV}_1 = \frac{1}{N}\sum_i \left\{ \frac{\phi_i^{\text{IV}_N} - \text{IV}_1 \phi_i^{\text{IV}_D}}{E[YZW]E[W] - E[YW]E[ZW]} \right\} + O_p\left(\frac{1}{N}\right)$$
$$\equiv \frac{1}{N}\sum_i \phi_i^{\text{IV}_1} + O_p\left(\frac{1}{N}\right).$$

The approach for deriving influence functions of the two weights and direct effects is similar. I focus on the properties of $\hat{w}_1^\rho$ and $\hat{\rho}_1$, and the analysis applies to $(\hat{w}_0^\rho, \hat{\rho}_0)$. Estimator $\hat{w}_1^\rho$ is given as

$$\hat{w}_1^\rho = \frac{\widehat{\Pr}(\text{AT} \mid 1)}{\widehat{\Pr}(\text{CP} \mid 1)} + \widehat{\Pr}(Z = 0).$$

To derive the asymptotic property of $\hat{w}_1^\rho$, I need to show the asymptotic property of the

conditional probability of the two subgroups. Let ϕ_i^{AT1}, ϕ_i^{AT0}, ϕ_i^{CP1}, ϕ_i^{CP0} denote the influence function for the four conditional probabilities $\widehat{\Pr}(\text{AT} \mid 1)$, $\widehat{\Pr}(\text{AT} \mid 0)$, $\widehat{\Pr}(\text{CP} \mid 1)$, $\widehat{\Pr}(\text{CP} \mid 0)$ respectively, which are derived as follows:

$$\phi_i^{\text{AT1}} = \frac{\phi_i^{D\tilde{Z}W}}{E[\tilde{Z}W]} - \frac{E[D\tilde{Z}W]\phi_i^{\tilde{Z}W}}{(E[\tilde{Z}W])^2}, \qquad \phi_i^{\text{AT0}} = \frac{\phi_i^{D\tilde{Z}\tilde{W}}}{E[\tilde{Z}\tilde{W}]} - \frac{E[D\tilde{Z}\tilde{W}]\phi_i^{\tilde{Z}\tilde{W}}}{(E[\tilde{Z}\tilde{W}])^2},$$

$$\phi_i^{\text{CP1}} = \frac{\phi_i^{DZW}}{E[ZW]} - \frac{E[DZW]\phi_i^{ZW}}{(E[ZW])^2} - \left\{ \frac{\phi_i^{D\tilde{Z}W}}{E[\tilde{Z}W]} - \frac{E[D\tilde{Z}W]\phi_i^{\tilde{Z}W}}{(E[\tilde{Z}W])^2} \right\},$$

$$\phi_i^{\text{CP0}} = \frac{\phi_i^{DZ\tilde{W}}}{E[Z\tilde{W}]} - \frac{E[DZ\tilde{W}]\phi_i^{Z\tilde{W}}}{(E[Z\tilde{W}])^2} - \left\{ \frac{\phi_i^{D\tilde{Z}\tilde{W}}}{E[\tilde{Z}\tilde{W}]} - \frac{E[D\tilde{Z}\tilde{W}]\phi_i^{\tilde{Z}\tilde{W}}}{(E[\tilde{Z}\tilde{W}])^2} \right\},$$

where $\tilde{Z} = 1 - Z$ and $\tilde{W} = 1 - W$.

Then the influence function for $\hat{w}_1^{\rho}$ is derived as follows by applying Taylor expansion:

$$\hat{w}_1^{\rho} - w_1^{\rho} = \frac{1}{N} \sum_i \left\{ \frac{\phi_i^{\text{AT1}}}{\Pr(\text{CP} \mid 1)} - \frac{\Pr(\text{AT} \mid 1)\phi_i^{\text{CP1}}}{(\Pr(\text{CP} \mid 1))^2} - \phi_i^Z \right\} + O_p\left(\frac{1}{N}\right)$$

$$\equiv \frac{1}{N} \sum_i \phi_i^{w_1^{\rho}} + O_p\left(\frac{1}{N}\right).$$

Now we only need to derive the influence function of estimator $\hat{\rho}_1$, given as

$$\hat{\rho}_1 = \frac{\hat{r}_1(1)\widehat{\Pr}(\text{CP} \mid 0) - \hat{r}_1(0)\widehat{\Pr}(\text{CP} \mid 1)}{\widehat{\Pr}(\text{AT} \mid 1)\widehat{\Pr}(\text{CP} \mid 0) - \widehat{\Pr}(\text{AT} \mid 0)\widehat{\Pr}(\text{CP} \mid 1)}.$$

Let ϕ_i^{d1} denote the influence function for the denominator of $\hat{\rho}_1$, derived as

$$\phi_i^{d1} = \Pr(\text{AT} \mid 1)\phi_i^{\text{CP0}} + \Pr(\text{CP} \mid 0)\phi_i^{\text{AT1}} - \Pr(\text{AT} \mid 0)\phi_i^{\text{CP1}} - \Pr(\text{CP} \mid 1)\phi_i^{\text{AT0}}.$$

I look at the numerator of estimator $\hat{\rho}_1$. Let $\phi_i^{YDZW} = Y_i D_i Z_i W_i - E[YDZW]$ for any random variables (Y, D, Z, W). According to the definition of $\hat{r}_1(1), \hat{r}_1(0)$, their influence

functions $\phi_i^{r_{11}}$ and $\phi_i^{r_{10}}$ are shown as

$$\phi_i^{r_{11}} = \frac{\phi_i^{YDZW}}{E[ZW]} - \frac{E[YDZW]\phi_i^{ZW}}{(E[ZW])^2} - \left\{ \frac{\phi_i^{YD\check{Z}W}}{E[\check{Z}W]} - \frac{E[YD\check{Z}W]\phi_i^{\check{Z}W}}{(E[\check{Z}W])^2} \right\},$$

$$\phi_i^{r_{10}} = \frac{\phi_i^{YDZ\tilde{W}}}{E[Z\tilde{W}]} - \frac{E[YDZ\tilde{W}]\phi_i^{Z\tilde{W}}}{(E[Z\tilde{W}])^2} - \left\{ \frac{\phi_i^{YD\check{Z}\tilde{W}}}{E[\check{Z}\tilde{W}]} - \frac{E[YD\check{Z}\tilde{W}]\phi_i^{\check{Z}\tilde{W}}}{(E[\check{Z}\tilde{W}])^2} \right\}.$$

Let ϕ_i^{n1} denote the influence function of the numerator of $\hat{\rho}_1$, derived as

$$\phi_i^{n1} = r_1(1)\phi_i^{\mathrm{CP0}} + \Pr(\mathrm{CP} \mid 0)\phi_i^{r_{11}} - r_1(0)\phi_i^{\mathrm{CP1}} - \Pr(\mathrm{CP} \mid 1)\phi_i^{r_{10}}.$$

Now we are ready to derive the influence function for estimator $\hat{\rho}_1$. Let $\phi_i^{\rho_1}$ denote the influence function for estimator $\hat{\rho}_1$, derived as follows:

$$\phi_i^{\rho_1} = \frac{\phi_i^{n1} - \rho_1\phi_i^{d1}}{\Pr(\mathrm{AT} \mid 1)\Pr(\mathrm{CP} \mid 0) - \Pr(\mathrm{AT} \mid 0)\Pr(\mathrm{CP} \mid 1)}.$$

Therefore, the asymptotic property of estimator $\hat{\rho}_1$ is derived as

$$\sqrt{N}(\hat{\rho}_1 - \rho_1) \to \mathcal{N}(0, \mathrm{Var}(\phi_i^{\rho_1})).$$

The analysis for the estimators $\hat{\rho}_0$ and $\hat{w}_0^\rho$ can be shown similarly, so it is omitted here. Let $\phi_i^{\rho_0}$ and $\phi_i^{w_0^\rho}$ denote the influence function for $\hat{\rho}_0$ and $\hat{w}_0^\rho$ respectively. Then the influence function ϕ_i^{LATE} for estimator $\widehat{\mathrm{LATE}}$ is expressed as

$$\phi_i^{\mathrm{LATE}} = \phi_i^{\mathrm{IV}_1} - \rho_1\phi_i^{w_1^\rho} - w_1^\rho\phi_i^{\rho_1} - \rho_0\phi_i^{w_0^\rho} - w_0^\rho\phi_i^{\rho_0}.$$

The asymptotic property for $\widehat{\mathrm{LATE}}$ is shown as

$$\sqrt{N}(\widehat{\mathrm{LATE}} - \mathrm{LATE}) \to \mathcal{N}(0, \mathrm{Var}(\phi_i^{\mathrm{LATE}})).$$

A.3   Proof of Proposition 1

Proof. Following the proof in A.1, LATE can be divided into two groups $Z = 1$ and $Z = 0$ as follows:

$$\text{LATE} = E[Y_1 - Y_0 \mid \text{CP}]$$
$$= E[Y_{1,1} - Y_{0,1} \mid \text{CP}, Z = 1] \Pr(Z = 1) + E[Y_{1,0} - Y_{0,0} \mid \text{CP}, Z = 0] \Pr(Z = 0)$$
$$= E[Y_{1,1} - Y_{0,1} \mid \text{CP}] \Pr(Z = 1) + E[Y_{1,0} - Y_{0,0} \mid \text{CP}] \Pr(Z = 0).$$

Using $\rho_{\text{CP},d} = E[Y_{d,1} - Y_{d,0} \mid \text{CP}]$ to substitute $E[Y_{0,1} \mid \text{CP}]$ and $E[Y_{1,0} \mid \text{CP}]$ has the following implication:

$$\text{LATE} = E[Y_{1,1} - Y_{0,0} \mid \text{CP}] - \rho_{\text{CP},1} \Pr(Z = 0) - \rho_{\text{CP},0} \Pr(Z = 1).$$

As shown in A.1, using variation in $r_1(w)$ with respect to w can identify $\rho_{\text{AT},1}$ and $E[Y_{1,1} \mid \text{CP}]$ as follows:

$$\rho_{\text{AT},1} = \frac{r_1(1) \Pr(\text{CP} \mid 0) - r_1(0) \Pr(\text{CP} \mid 1)}{\Pr(\text{AT} \mid 1) \Pr(\text{CP} \mid 0) - \Pr(\text{AT} \mid 0) \Pr(\text{CP} \mid 1)},$$
$$E[Y_{1,1} \mid \text{CP}] = \frac{r_1(1) - \rho_1 \Pr(\text{AT} \mid 1)}{\Pr(\text{CP} \mid 1)}.$$

Similarly, the direct effect $\rho_{\text{NT},0}$ and $E[Y_{0,0} \mid \text{CP}]$ are identified by exploiting variation in $r_0(w)$:

$$\rho_{\text{NT},0} = \frac{r_0(1) \Pr(\text{CP} \mid 0) - r_0(0) \Pr(\text{CP} \mid 1)}{\Pr(\text{NT} \mid 1) \Pr(\text{CP} \mid 0) - \Pr(\text{NT} \mid 0) \Pr(\text{CP} \mid 1)},$$
$$E[Y_{0,0} \mid \text{CP}] = -\frac{r_0(1) - \rho_0 \Pr(\text{NT} \mid 1)}{\Pr(\text{CP} \mid 1)}.$$

Under Assumption 4 about the difference in direct effects across subgroups, $\rho_{\text{CP},d}$ for $d \in \{0, 1\}$ can be bounded as

$$\rho_{\text{AT},1} - k_1 \le \rho_{\text{CP},1} \le \rho_{\text{AT},1} + k_1,$$
$$\rho_{\text{NT},0} - k_0 \le \rho_{\text{CP},0} \le \rho_{\text{NT},0} + k_0.$$

Therefore, the bounds on LATE is established as follows:

$$\text{LATE} \geq E[Y_{1,1} - Y_{0,0} \mid \text{CP}] - (\rho_{\text{AT},1} + k_1)\Pr(Z = 0) - (\rho_{\text{NT},0} + k_0)\Pr(Z = 1),$$

$$\text{LATE} \leq E[Y_{1,1} - Y_{0,0} \mid \text{CP}] - (\rho_{\text{AT},1} - k_1)\Pr(Z = 0) - (\rho_{\text{NT},0} - k_0)\Pr(Z = 1).$$

Plugging into the formulas for $(\rho_{\text{AT},1}, \rho_{\text{NT},0})$ and $(E[Y_{1,1} \mid \text{CP}], E[Y_{0,0} \mid \text{CP}])$ lead to the results in Proposition 1. The lower bound is achieved when $\rho_{CP,1} = \rho_{\text{AT},1} + k_1$ and $\rho_{CP,0} = \rho_{\text{NT},0} + k_0$, and the upper bound is achieved when $\rho_{CP,1} = \rho_{\text{AT},1} - k_1$ and $\rho_{CP,0} = \rho_{\text{NT},0} - k_0$. Therefore, the bounds are sharp.

$\square$

A.4 More Simulation Results

This section presents the simulation results of the two estimators $(\hat{\theta}_{zw}, \hat{\theta}_z)$ when the size of the subgroups changes. Consider that the potential treatments (D_1, D_0) is given as follows:

$$D_1 = \mathbb{1}\{\epsilon \leq 1\},$$

$$D_0 = \mathbb{1}\{\epsilon \leq k\}.$$

The value of k determines the size of the three subgroups, and the size of compliers $\Pr(k < \epsilon < 1)$ decreases when the value of k increases. I consider three specifications of k: $k \in \{-0.25, 0, 0.25\}$.

Table 3 and 4 show the results of the two estimators $(\hat{\theta}_{zw}, \hat{\theta}_z)$ under different values of k and different values of direct effects ρ with the sample size $N = 1000$ and $N = 4000$, respectively. The two estimators $(\hat{\theta}_{zw}, \hat{\theta}_z)$ both perform better when the size of compliers increases, while the estimator $\hat{\theta}_{zw}$ uniformly performs better than the estimator $\hat{\theta}_z$ with nonzero direct effects regardless of the size of compliers. Therefore, the robustness feature of the estimator $\hat{\theta}_{zw}$ to nonzero direct effects still holds under different probabilities of subgroups.

Table 3: Performance Comparisons of $\hat{\theta}_{zw}$ and $\hat{\theta}_z$ ($N = 1000$)

ρ	$\hat{\theta}_{zw}$				$\hat{\theta}_z$			
	Bias	SD	rMSE	MAD	Bias	SD	rMSE	MAD
$k = -0.25$								
0	0.016	0.459	0.459	0.362	0.004	0.142	0.142	0.113
0.5	0.016	0.459	0.459	0.362	1.145	0.173	1.158	1.145
1	0.016	0.459	0.459	0.362	2.285	0.225	2.296	2.285
-1	0.016	0.459	0.459	0.362	-2.277	0.178	2.284	2.277
$k = 0$								
0	0.025	0.504	0.505	0.396	0.006	0.183	0.184	0.146
0.5	0.025	0.504	0.505	0.396	1.481	0.245	1.501	1.481
1	0.025	0.504	0.505	0.396	2.956	0.342	2.976	2.956
-1	0.025	0.504	0.505	0.396	-2.944	0.269	2.956	2.944
$k = 0.25$								
0	0.057	0.915	0.917	0.508	0.010	0.262	0.262	0.207
0.5	0.057	0.915	0.917	0.508	2.100	0.403	2.139	2.100
1	0.057	0.915	0.917	0.508	4.190	0.615	4.235	4.190
-1	0.057	0.915	0.917	0.508	-4.169	0.498	4.199	4.169

Table 4: Performance Comparisons of $\hat{\theta}_{zw}$ and $\hat{\theta}_z$ ($N = 4000$)

ρ	$\hat{\theta}_{zw}$				$\hat{\theta}_z$			
	Bias	SD	rMSE	MAD	Bias	SD	rMSE	MAD
$k = -0.25$								
0	0.007	0.214	0.214	0.170	-0.000	0.072	0.072	0.057
0.5	0.007	0.214	0.214	0.170	1.136	0.086	1.139	1.136
1	0.007	0.214	0.214	0.170	2.273	0.110	2.275	2.273
-1	0.007	0.214	0.214	0.170	-2.274	0.089	2.275	2.274
$k = 0$								
0	0.009	0.230	0.230	0.183	-0.001	0.092	0.092	0.073
0.5	0.009	0.230	0.230	0.183	1.466	0.119	1.470	1.466
1	0.009	0.230	0.230	0.183	2.932	0.165	2.936	2.932
-1	0.009	0.230	0.230	0.183	-2.933	0.133	2.936	2.933
$k = 0.25$								
0	0.013	0.273	0.273	0.217	-0.002	0.130	0.130	0.104
0.5	0.013	0.273	0.273	0.217	2.064	0.193	2.073	2.064
1	0.013	0.273	0.273	0.217	4.129	0.291	4.140	4.129
-1	0.013	0.273	0.273	0.217	-4.133	0.242	4.140	4.133