Sensitivity of LATE Estimates to Violations of the Monotonicity Assumption

Claudia Noack*

Job Market Paper

Abstract

In models with heterogeneous treatment effects, instrumental variables are often used to estimate local treatment effects. One key assumption of this approach is monotonicity, which implies that the effect of the instrument on the treatment status is monotone across the entire population. However, the validity of this assumption is questionable in many empirical applications. In this paper, I develop a method to assess the sensitivity of local treatment effect estimates to potential violations of the monotonicity assumption. I parameterize the degree to which monotonicity is violated using two sensitivity parameters: the first one determines the share of defiers in the population, and the second one measures differences in the distributions of outcomes between compliers and defiers. For each value of these two sensitivity parameters, I derive sharp bounds on the compliers' outcome distributions in the first-order stochastic dominance sense. I identify the robust region that is the set of all values of sensitivity parameters for which a given empirical conclusion, e.g. that the local average treatment effect is positive, is valid. Researchers can assess the credibility of their conclusion by checking whether all the plausible sensitivity parameters lie in the robust region. I obtain confidence sets for the robust region through a bootstrap procedure and illustrate the sensitivity analysis in an empirical application. I also extend this framework to analyze treatment effects of the entire population.

*This version: October 15, 2020. Please click here for the latest version. Department of Economics, University of Mannheim; E-mail: claudia.noack@gess.uni-mannheim.de. I am indepted to my advisor Christoph Rothe for invaluable support on this project. I am grateful for insightful discussions with Tim Armstrong, Matthew Masten, Ed Vytlacil, and Yoshiyasu Rai. Furthermore, I thank Erich Battistin, Michael Dobrew, Jasmin Fliegner, Xavier D'Haultfœuille, Martin Huber, Paul Goldsmith-Pinkham, Sacha Kapoor, Lukas Laffers, Thierry Magnac, Enno Mammen, Tomasz Olma, Vitor Possebom, Alexandre Poirier, Jonathan Roth, Pedro S'Antanna, Konrad Stahl, Matthias Stelter, Jörg Stoye, Philipp Wangner, and seminar and conference participants at the University of Mannheim, University of Heidelberg, University of Bonn, Yale University, ESWC 2020, ESEM 2019, IAAE 2019, and 1st International Ph.D. Conference at Erasmus University Rotterdam for helpful comments and suggestions.

1. INTRODUCTION

The local average treatment effect framework introduced in Imbens and Angrist (1994) is one of the most popular econometric frameworks for instrumental variable analysis in setups of heterogeneous treatment effects. The canonical setup consists of a binary instrumental variable and a binary treatment variable. In these settings, the Wald estimand equals the treatment effect of *compliers*, individuals for which the instrument influences the treatment status, given the well-known assumptions: monotonicity, independence, and relevance.

Monotonicity states that the instrument's effect on the treatment decision is monotone across all individuals. In the canonical example in which the instrument encourages individuals to take up the treatment, monotonicity thus rules out the existence of *defiers*, i.e. individuals that receive the treatment only if the instrument discourages them. Researchers might question the validity of this assumption in empirical applications. In these settings, the resulting local treatment effect estimates might be biased and might lead the researchers to draw wrong conclusions about the treatment effect.

To give an example of a setup in which monotonicity could plausibly be violated, I consider the often cited study of Angrist and Evans (1998), who analyze the effect of having a third child on mother's labor market outcomes. As the decision to have a third child is endogenous, the authors use as an instrument whether the first two children are of the same sex. The underlying reasoning is that some of the parents would only decide to have a third child if their first two children were of the same sex; these parents are compliers. Monotonicity seems to be questionable in this setting[1] as parents, who have a strong preference for one specific sex, might act as a defier within this setup. Consider for example parents who want to have at least two boys and their first child is a boy. Contrary to the instrument, they have two children if their second child is a boy, and three children if their second child is a girl. As the monotonicity assumption might therefore be questionable in this example, one should question the validity of the empirical conclusions drawn from this analysis.

In this paper, I provide a framework to evaluate the sensitivity of treatment effect estimates to a potential violation of the monotonicity assumption. As noted in Angrist, Imbens, and Rubin (1996), a violation of the monotonicity assumption always has two dimensions: The first dimension represents the heterogeneous effect of the instrumental variable on the treatment variable, the presence of defiers. The second dimension represents the heterogeneous effect of the treatment variable on the outcome variable, the outcome heterogeneity between defiers and compliers. I induce the degree to which monotonicity is violated by parameterizing these two dimensions. To keep things as simple as possible, I parameterize the existence of defiers

[1]The other local average treatment effect assumptions seem to be plausible in this setting. The independence assumption and the relevance assumption seem to be satisfied The independence assumption seems to be plausible by the following reasoning: First, the sex of a child is determined by nature, and therefore the random assignment assumption seems to be satisfied in this setting. Second, the exclusion restriction seems to hold as well as only the number of and not the sex of the child arguably influences the labor market outcome. The relevance assumption can be tested.

by their population size and the outcome heterogeneity by the Kolmogorov-Smirnov (KS) norm, which bounds the difference of the cumulative distribution functions of compliers and defiers. Based on these two sensitivity parameters, I identify sharp bounds of the compliers' outcome distributions in a first-order stochastic dominance sense. These bounds therefore also imply sharp bounds on various different treatment effects, e.g. the average treatment effect.

In a first step, I identify the *sensitivity region*. The sensitivity region defines the set of sensitivity parameters for which there exists a data generating process, which is consistent with the independence and relevance assumption and implies both the observed probabilities and the sensitivity parameters. Sensitivity parameters lying in the complement of the sensitivity region are not consistent with the model, and therefore I do not analyze them further. A byproduct of the derivation of the sensitivity region is a formal proof for sharp bounds of the population size of defiers.

In the second step, I identify the *robust region*, which is the set of sensitivity parameters that imply treatment effects to be consistent with a certain empirical conclusion (e.g., that the compliers' treatment effect has a specific sign or a certain order of magnitude).[2] Parameters lying in the complement of the robust region, the *nonrobust region*, imply treatment effects that are not, or may not be, consistent with the given empirical conclusion. The robust region and the nonrobust region are separated from each other by the *breakdown frontier*, (following the terminology of Masten and Poirier (2020)). The breakdown frontier identifies for each defiers' population size the weakest assumption about outcome heterogeneity, which is necessary to be imposed in order to imply treatment effects being consistent with the particular empirical conclusion.

Additionally, while the main focus of this paper lies on treatment effects of compliers, I also show how this framework can be exploited to analyze treatment effects of defiers. Under further support assumptions, treatment effects of even the entire population are identified as well, which complements known results in the literature (see Balke and Pearl, 1997; Kitagawa, 2020; Machado, Shaikan, and Vytlacil, 2019). Furthermore, I show how covariates can be used to tighten the sensitivity region. As the explicit expressions of the sensitivity and robust regions are rather complicated and difficult to interpret, I also provide simplified analytical expressions of these regions in the case of a binary outcome.

To construct confidence sets for both the sensitivity and the robust region, I show that both regions are determined through mappings of some underlying parameters. However, these mappings are not Hadamard-differentiable and inference methods relying on standard Delta-method arguments are therefore not applicable. I show how to construct alternative mappings which are smooth and which bound the parameters of interest. This construction leads to mappings for which standard Delta-method arguments are applicable, and I use the

[2] See Masten and Poirier (2020) for a detailed exposition of this approach.

nonparametric bootstrap to construct valid confidence sets for the parameters of interest. Considering a binary outcome variable, the mappings resulting in the sensitivity and robust region are considerably simpler. I can therefore use a Delta-method for directionally differentiable mappings to show asymptotic distributional results and apply a bootstrap procedure to construct asymptotically valid confidence sets.

To conclude this framework, I show in a Monte Carlo study that my proposed inference method has good finite sample properties. I further apply my method to the setup studied by Angrist and Evans (1998) introduced above. I show that relatively strong assumptions on either the population size of the defiers or the treatment effect heterogeneity have to be imposed in order to solely preserve the sign of the estimated treatment effect. This result demonstrates that the monotonicity assumption is key in the local treatment effect framework.

This approach is useful in at least four ways. First, by evaluating the size of the sensitivity region, one can determine the plausibility of the model. If this set is empty, the model is refuted, which implies that even though one would allow for an arbitrary violation of the monotonicity assumption, the independence assumption has to be violated. Second, researchers can analyze the sensitivity of their estimates with respect to the degree to which the monotonicity assumption is violated by varying the sensitivity parameters within the sensitivity region. Third, by evaluating the plausibility of the parameters within the robust region, researchers can assess the sign or the order of magnitude of the treatment effect. Moreover, they can be transparent about the assumptions they are imposing in order to arrive at certain empirical conclusions. By doing so, they can still arrive at empirical conclusions of interest while imposing assumptions which might be more credible in the given situation. Fourth, one can assess to which degree monotonicity has to be violated to overturn a particular empirical conclusion. Within my framework, researchers can use their economic insights about the analyzed situation to judge the severity of a violation of the monotonicity assumption.

The remainder is structured as follows: A literature review follows and Section 2 illustrates the setup in a simplified setting. Section 3 introduces the sensitivity parameters. The main sensitivity analysis is presented in Section 4. Section 5 also discusses extensions and Section 6 derives estimation and inference results. Section 6 contains a simulation study and Section 8 an empirical example. Section 9 concludes. Proofs and additional materials are deferred to the appendix.

Literature. This paper relates to several strands of the literature. First, this paper follows the growing strand of the literature, which considers sensitivity analysis in various applications. These applications include among many others, violations of parametric assumptions, violations of moment conditions, and various examples within the treatment effect literature (see among others Andrews, Gentzkow, and Shapiro, 2017, 2018, 2020; Andrews and Shapiro, 2020; Armstrong and Kolesár, 2018; Bonhomme and Weidner, 2018, 2019; Chen, Tamer, and

Torgovitsky, 2015; Christensen and Connault, 2019; Conley, Hansen, and Rossi, 2012; Imbens, 2003; Kitamura, Otsu, and Evdokimov, 2013; Mukhin, 2018; Rambachan and Roth, 2019). My paper is very closely related to Masten and Poirier (2018, 2020), who generalize the idea of breakdown points developed in Horowitz and Manski (1995); Imbens (2003); Kline and Santos (2013); Stoye (2005, 2010). These papers consider several assumptions in the treatment effect literature, but not the monotonicity assumption.

Second, it is related to the local average treatment effect framework literature, which is formally introduced in Imbens and Angrist (1994) and further in Vytlacil (2002). Several papers consider violations of the monotonicity assumption through different types of assumptions. Balke and Pearl (1997); Huber (2015); Huber, Laffers, and Mellace (2017); Huber and Mellace (2015); Machado et al. (2019); Manski (1990) consider a binary and Kitagawa (2020) a continuous outcome variable and partially identify the average treatment effect. (Dahl, Huber, and Mellace, 2019; Huber et al., 2017; Manski and Pepper, 2000; Small, Tan, Ramsahai, Lorch, and Brookhart, 2017) propose alternative assumptions on the data generating process, which are strictly weaker than monotonicity and obtain bounds on various treatment effects.

De Chaisemartin (2017) shows that in the presence of defiers, the Wald estimand still identifies a convex combination of causal treatment effects under certain assumptions of only a subpopulation of compliers. In a policy context, the compliers' treatment effect might be of particular interest because the compliers' treatment status is most likely to change with a small policy change. However, the same reasoning does not apply to the subpopulation of compliers. Klein (2010) evaluates the sensitivity of the compliers' treatment effect to random departures from monotonicity. Fiorini and Stevens (2016) give examples of analyzing the sensitivity of the monotonicity, and Huber (2014) considers a violation of monotonicity in a specific example. They do not provide sharp identification results of the compliers' treatment effect in the presence of defiers, nor do they derive the robust region. A violation of the monotonicity assumption with a non-binary instrumental variable is considered, and alternative assumption and testing procedures are proposed in Frandsen, Lefgren, and Leslie (2019); Mogstad, Torgovitsky, and Walters (2019); Norris (2019); Norris, Pecenco, and Weaver (2019). This paper contributes to this literature by presenting an effective tool to analyze the severity of a potential violation of the monotonicity assumption without answering whether a violation is present. It thus gives applied researchers a new tool to evaluate the robustness of their estimates to a violation of the monotonicity assumption and their estimates may thereby gain in credibility.

My proposed inference procedure builds on seminal work about Delta-methods for non-differentiable mappings by Dümbgen (1993); Fang and Santos (2018); Hong and Li (2018); Shapiro (1991) and it further exploits ideas of smoothing population parameters by Chernozhukov, Férnandez–Val, and Galichon (2010); Masten and Poirier (2020).

Table 1: Groups defined by their treatment status and instrument

	Always Takers (AT)	Never Takers (NT)	Compliers (CO)	Defiers (DF)
Z=1	$D_1 = 1$	$D_1 = 0$	$D_1 = 1$	$D_1 = 0$
Z=0	$D_0 = 1$	$D_0 = 0$	$D_0 = 0$	$D_0 = 1$

2. SETUP AND ILLUSTRATIVE EXAMPLE

In this section, I illustrate the sensitivity analysis in a simplified example.[3] I introduce the sensitivity parameters, the sensitivity region and the robust region.

2.1. **Setup.** Let Y_0 and Y_1 denote the potential outcome, D_1 and D_0 the binary potential treatment status and Z the binary instrument. I assume that the random variables (Y, D, Z) are observed, where $Y = DY_1 + (1 - D)Y_0$ and $D = ZD_1 + (1 - Z)D_0$. Based on the effect of the instrument on the treatment status, one can distinguish four different groups: always takers, never takers, compliers and defiers. The groups are displayed in Table 1 and π_{AT}, π_{NT}, π_{CO}, and π_{DF} denote the population sizes of the respective group. To simplify the notation, I denote by dT the group of always takers if $d = 1$ and otherwise the never takers. I denote the outcome distribution of a variable Y by F_Y, its density function by f_Y if it exists, and its support by $\mathbb{Y}$. I assume that all necessary moments of all outcomes exist. The key parameters of interest in this analysis are compliers' treatment effects, where the compliers' average treatment effect is denoted by[4]

$$\Delta_{CO} = \mathbb{E}[Y_1 - Y_0 | CO].$$

I set without loss of generality $\mathbb{P}(D = 1|Z = 1) \geq \mathbb{P}(D = 1|Z = 0)$. Throughout the paper, I impose the following identifying assumptions.

Assumption 1. *(i) Independence: $(Y_1, Y_0, D_1, D_0) \perp Z$; (ii) Relevance: $\mathbb{P}(D = 1|Z = 1) > \mathbb{P}(D = 1|Z = 0)$.*

I refer to an extensive discussion of these assumptions to Angrist et al. (1996).[5]

2.2. **Illustrative Example.** To illustrate the sensitivity analysis, I provide a simple example to show the main mechanisms of the model.

2.2.1. *Sensitivity Parameter Space.* In the presence of defiers, the compliers' treatment effect is not point identified. Angrist et al. (1996) show that the Wald estimand equals a weighted

[3] I do not derive any sharp identification results here as this example is purely for illustration.

[4] Similarly, the defier's average treatment effect is denoted by $\Delta_{DF} = \mathbb{E}[Y_1 - Y_0|DF]$ and the compliers' quantile treatment by $\Delta_{DF}(\tau)$ for a given quantile τ.

[5] The relevance assumption is imposed for the main sensitivity analysis of the paper for the sake of simplification of notation. It could be replaced by $\pi_{\text{CO}} > 0$.

difference of the compliers' and defiers' potential outcomes

$$\beta^{IV} = \frac{\text{Cov}(Y, Z)}{\text{Cov}(D, Z)} = \frac{1}{\pi_{\text{CO}} - \pi_{\text{DF}}} \left(\pi_{\text{CO}}\Delta_{CO} - \pi_{\text{DF}}\Delta_{DF}\right). \tag{1}$$

If $\pi_{\text{DF}} = 0$, implying the absence of defiers, heterogeneous treatment effects of compliers and defiers do not alter the identification of the compliers' treatment effect. Vice versa, if compliers and defiers have the same average treatment effect, the compliers' treatment effect is identified despite the potential presence of defiers.

As the population size of compliers is identified for any given population size of defiers[6], it follows that three parameters of Equation 1 are generally not identified: the population size of defiers π_{DF}, the treatment effect of compliers Δ_{CO} and of defiers Δ_{DF}. To bound the compliers' treatment effect, I introduce two sensitivity parameters. The first one determines the defiers' population size and the second one treatment effect heterogeneity of compliers and defiers. These two parameters measure the degree to which monotonicity is violated and represent the two dimensions of heterogeneity: (i) heterogeneous effects of the instrument on the treatment status and (ii) heterogeneous effects of the treatment on the outcome.

The heterogenous impact of the instrument on the treatment status, is parameterized as simple as possible by the population size of defiers

$$\pi_{\text{DF}} = \mathbb{P}(D_1 = 0 \text{ and } D_0 = 1).$$

The absence of defiers is denoted by $\pi_{\text{DF}} = 0$ in which case monotonicity holds. A greater value of the parameter π_{DF} implies a more severe violation of monotonicity. For a given value of the parameter π_{DF} the population size of each group is point identified.[7]

Even if one knew the population size of defiers, the identification of the compliers' average treatment effect would be still infeasible if compliers and defiers have heterogeneous treatment effects. To further parameterize the second dimension of heterogeneity, I introduce a sensitivity parameter δ_a which equals the absolute differences in treatment effects of both groups

$$\delta_a = |\Delta_{CO} - \Delta_{DF}|.$$

2.2.2. *The Sensitivity Region.* The *sensitivity region* is the set of sensitivity parameters which could have induced the observed probabilities without violating the model assumptions. On the contrary, sensitivity parameters lying in its complement imply a violation of the model assumptions. The sensitivity region is therefore not identified by imposing any additional assumptions but solely because the observed probabilities cannot be explained by any data generating process, which is consistent with the specific sensitivity parameters and the

[6]The population size of compliers is given by $\mathbb{P}(D = 1|Z = 1) - \mathbb{P}(D = 0|Z = 0) + \pi_{\text{DF}}$.

[7]It follows that $\pi_{\text{AT}} = \mathbb{P}(D = 1|Z = 0) - \pi_{\text{DF}}$, $\pi_{\text{NT}}(\pi_{\text{DF}}) = \mathbb{P}(D = 0|Z = 1) - \pi_{\text{DF}}$ and $\pi_{\text{CO}} = \mathbb{P}(D = 1|Z = 1) - \pi_{\text{AT}}(\pi_{\text{DF}})$.

independence assumption. This reasoning implies that if the sensitivity region is empty, the model is rejected: Even though the monotonicity assumption may be violated, the independence assumption has to be violated as well. In this illustrative example, I set the sensitivity region for simplicity just to the entire sensitivity parameter space such that

$$\mathrm{SR}_a = [0, 0.5) \times \mathbb{R}_+.$$

2.2.3. *The Robust Region and the Breakdown Frontier.* Given some imposed values of the sensitivity parameters, it follows that the average treatment effect of the compliers is partially identified given any pair of sensitivity parameters by

$$\Delta_{CO} \in \left[\beta^{IV} - \frac{\pi_{\mathrm{DF}}}{\pi_{\mathrm{CO}} - \pi_{\mathrm{DF}}}\delta_a,\ \beta^{IV} + \frac{\pi_{\mathrm{DF}}}{\pi_{\mathrm{CO}} - \pi_{\mathrm{DF}}}\delta_a\right].$$

In a typical sensitivity analysis, one now considers different values of the sensitivity parameters and evaluates the identified sets of the treatment effect of interest. Based on these identified sets, one can judge the robustness of the LATE estimates to a potential violation.

In many empirical applications, one is, however, not interested in the precise treatment effect but rather in its sign or order of magnitude. Within a sensitivity analysis, it is therefore natural to start with the empirical conclusion of interest and to ask which sensitivity parameters imply treatment effects which are consistent with this conclusion. I therefore follow the approach of breakdown frontiers (see Kline and Santos (2013); Masten and Poirier (2020)).

I identify the *robust region*, which is the set of sensitivity parameters implying treatment effects being consistent with a particular conclusion. In this setting, I focus on one-dimensional empirical conclusions, e.g., the treatment effect has a specific sign or is of a certain order of magnitude. Parameters lying in the complement of the robust region, the *nonrobust region*, imply treatment effects, which are not or might not be consistent with the empirical conclusion. The two regions are separated by the *breakdown frontier*. As the degree to which monotonicity is violated is increasing in the sensitivity parameters, the breakdown frontier identifies for each value of the population size of defiers the weakest assumption on outcome heterogeneity between compliers and defiers, which implies that the compliers' treatment effect is still consistent with the particular empirical conclusion.

If Assumption 1 holds in this illustrative example and $\beta^{IV} \geq \mu$, the breakdown frontier (BF) and the robust region (RR) for the claim $\Delta CO \geq \mu$ are given by

$$BF(\mu) = \left\{(\pi_{\mathrm{DF}}, \delta_a) \in \mathrm{SR}_a : \delta_a = \frac{\pi_{\mathrm{CO}} - \pi_{\mathrm{DF}}}{\pi_{\mathrm{DF}}}(\beta^{IV} - \mu)\right\}$$
$$RR(\mu) = \left\{(\pi_{\mathrm{DF}}, \delta_a) \in \mathrm{SR}_a : \delta_a \leq \frac{\pi_{\mathrm{CO}} - \pi_{\mathrm{DF}}}{\pi_{\mathrm{DF}}}(\beta^{IV} - \mu)\right\}.$$

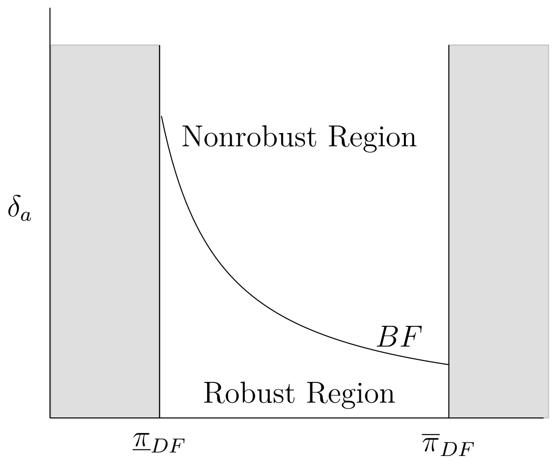


Figure 1: Illustration of a sensitivity and a robust region. Non-shaded area represents sensitivity region. $[\underline{\pi}_{DF}, \overline{\pi}_{DF}]$ represent some bounds on the population size of defiers.

Due to the functional form of the breakdown frontier, the nonrobust region is a convex set. An illustrative example is shown in Figure 1.

Based on my framework, researchers can evaluate the robustness of their empirical findings to a potential violation of monotonicity. First, by analyzing the size of the sensitivity region, researchers can judge the model assumption, e.g., if this region is empty, the model is rejected. Second, they can qualitatively evaluate the robustness of the estimated by the size of the robust region. If all sensitivity parameters, which the research considers to be plausible, are lying in the robust region, the empirical conclusion is robust to a potential violation of monotonicity. While if any sensitivity parameter pair, which the researcher considers to be reasonable, lies in the nonrobust region, the empirical conclusion is not robust. Third, they can quantify the impact of a potential violation of the monotonicity while naming the sensitivity parameters they are imposing to arrive at a certain empirical conclusion. By this analysis, the empirical estimates may gain in credibility.[8]

In the specific example of the average treatment effect, neither the sensitivity region nor the robust regions are sharp, as the reduction of one parameter in comparison to the entire distribution necessarily leads to a loss of information. Even though a parameter pair may ly within the sensitivity region, it might not imply a well-defined data generating process, which would be consistent with the model assumptions and the observed probabilities. In order to avoid this critique the entire counterfactual outcome distributions are considered in the remainder of the paper.

[8]Within a different context, Masten and Poirier (2018) propose to consider the point in the sensitivity parameter space, which imposes the weakest violation of the underlying assumptions but lies in the sensitivity region which is $\pi_{\mathrm{DF}} = \underline{\pi}_{DF}$ in the context of this sensitivity parameter. The outcome distributions of compliers and defiers are point identified for this specific sensitivity parameter, and thus the treatment effect is point identified as well. However, in the context of the violations of the monotonicity assumption regarding this point is equivalent to the assumption imposed in (Huber and Mellace, 2012), implying that the support of the outcome variable of defiers and compliers is distinct.

3. SENSITIVITY PARAMETERS

I present the sensitivity parameters for the main sensitivity analysis of this paper. Similar to the Wald estimand, the following functions

$$G_1(y) = \frac{\mathrm{Cov}(\mathbb{1}[Y \leq y], D)}{\mathrm{Cov}(Z, D)} \qquad \text{and} \qquad G_0(y) = \frac{\mathrm{Cov}(\mathbb{1}[Y \leq y], 1-D)}{\mathrm{Cov}(Z, 1-D)}. \tag{2}$$

equal the outcome distribution of the compliers in the absence of defiers. For any population size of defiers it holds analog to Equation 1 that

$$G_d(y) = \frac{1}{\pi_{\mathrm{CO}} - \pi_{\mathrm{DF}}} \left(\pi_{\mathrm{CO}} F_{Y_d^{CO}}(y) - \pi_{\mathrm{DF}} F_{Y_d^{DF}}(y) \right).$$

The compliers' outcome distribution functions are thus identified up to the presence of defiers and the heterogeneity of the compliers' and defiers' outcome distributions. Similarly as above, I introduce two sensitivity parameters parameterizing these two dimensions of heterogeneity. First, the presence of defiers is parameterized by π_{DF} as defined above. Second, outcome heterogeneity is represented by δ, which bounds the maximal difference of the outcome cumulative distribution functions of both groups by the Kolmogorov-Smirnov (KS) norm

$$\max_{d \in \{0,1\}} \sup_{y \in \mathbb{Y}} \{ |F_{Y_d^{CO}}(y) - F_{Y_d^{DF}}(y)| \} = \delta,$$

where $\delta \in [0, 1]$. If $\delta = 1$, the outcome distributions might be arbitrarily different. If $\delta = 0$, the outcome distributions are restricted the most as both distribution functions coincide. A greater value of δ thus indicates a more severe violation of monotonicity.

There are clearly many different possibilities how to parameterize hetereogentiy between distributions. In this paper, I choose the Kolomogrov-Smirnov norm, as it leads to tractable analytical solutions of the bounds of the compliers outcome distribution and more importantly this parameterizing is simple enough to be interpretable in an empirical conclusion. I give an intuition about one possible interpretation of this restriction within an empirical context in Section 8. A similar parameterization in a different context is chosen in Kline and Santos (2013).[9] By interpreting the results in an empirical context, one should thus bear in mind that the conclusion are conservative as the use of the KS norm necessarily leads to a loss of information.

4. SENSITIVITY ANALYSIS

The main theoretical result of this paper is the identification of the sharp sensitivity and robust region for a certain empirical conclusion based on the sensitivity parameters $(\pi_{\mathrm{DF}}, \delta)$.

[9] A disadvantage of this parameterization is that it is very weak on the tails of the distributions and thus bounds on the tails are likely to be uninformative. To overcome this weakness, one could also impose a *weighted* KS assumption penalizing deviations at the tails of the two distributions more. However, despite its weakness but due to its tractability, I keep the unweighted KS assumption.

A key step within this derivation is to identify sharp bounds of the compliers' outcome distribution for any given sensitivity parameter pair $(\pi_{\mathrm{DF}}, \delta)$.

4.1. **Derivation of sharp bounds on the outcome distributions.** In this section, I derive the explicit characterization of sharp bounds on the compliers' outcome distribution for any given sensitivity parameter pair in the sensitivity region $(\pi_{\mathrm{DF}}, \delta)$[10]. These bounds are sharp in a first order stochastic dominance sense. An explicit characterization of these bounds allows to understand the influence of the sensitivity parameters on the bounds and more importantly, it implies sharp bounds on the parameters of interest, e.g.the compliers' average treatment effect. Moreover, the explicit characterization allows that the inference procedure can be directly based on these sharp bounds.

Identification of the compliers' outcome distribution is based on the following thought experiment in this framework: consider that the model is correct (i.e. independence assumption is satisfied) and both the population distribution of (Y, D, Z) and the sensitivity parameters $(\pi_{\mathrm{DF}}, \delta)$ are known, what can be learned about the data generating process, which is given by (Y_1, Y_0, D_1, D_0, Z) by the model assumptions. In this framework, there are two possible states of the world. First, there exists at least one such data generating process in which case the compliers' outcome distributions can be (partially) identified. Second, there does not exist such a data generating process such that the sensitivity parameters are refuted. One can then consider a different pair of sensitivity parameters. If all sensitivity parameters can be refuted, the model is refuted, which implies that even though the monotonicity assumption might be arbitrarily violated, any of the other model assumptions has to be violated too. In the following, I proceed by backwards engineering. I assume that the data is compatible with the sensitivity parameter, but if there is no such compliers' outcome distribution this pair of sensitivity parameters is refuted.

The identification of the compliers' outcome distribution is based on the following reasoning: Based on the independence assumption and the definition of the types, the data generating process is represented by the set of marginal outcome distributions of all four groups in the presence and absence of treatment. Considering the marginal distributions is sufficient in this analysis, as the data are non-informative about the joint distributions beyond the populations sizes. However, the population sizes are already point identified through the sensitivity parameters, where I already use this information. As the data are also non-informative about the always takers in the absence of treatment and of the never takers in the presence of treatment, these distributions are left unrestricted as well. Therefore, there are six marginal distributions left on which the observed distribution poses restrictions. The problem is further simplified by observing that based on the sensitivity parameters and the model assumptions, the observed probabilities impose two separate sets of restrictions about the

[10]Parameters in the complement of the sensitivity region refute the mode. I can therefore not bound the compliers outcome distribution for these sensitivity parameters.

marginal distributions in the absence and presence of treatment.

The following outcome probabilities are observed

$$Q_{dd}(y) \equiv \mathbb{P}(Y \leq y, D = d | Z = d) = \pi_{\mathrm{CO}} F_{Y_d^{CO}}(y) + \pi_{dT} F_{Y_d^{dT}}(y) \tag{3}$$

$$Q_{d(1-d)}(y) \equiv \mathbb{P}(Y \leq y, D = d | Z = 1 - d) = \pi_{\mathrm{DF}} F_{Y_d^{DF}}(y) + \pi_{dT} F_{Y_d^{dT}}(y). \tag{4}$$

For a given value on the sensitivity parameter π_{DF} and a given compliers' distribution, the always takers and defiers distributions are point identified in the presence of treatment. This reasoning implies that the set of all potential compliers' outcome distributions is partially identified. This set is the set of all distribution functions which imply through these two equations distribution functions of the always takers and defiers, and that the compliers' and defiers' outcome distributions are consistent with the sensitivity parameter δ. This identified set is in general relative complicated to describe. However, in this setting the interest lies in treatment effects of the compliers which are mappings of these functions. It therefore suffices to derive sharp bounds in a first order stochastic dominance sense. This construction is possible as Equations 3 and 4 show that due to the linear relation between the outcome functions and the linear restriction imposed by the sensitivity parameter δ that a bound of the compliers' outcome distribution implies a bound on the always or never takers outcome distribution and thereby a bound on the defiers outcome distribution.

To construct sharp bounds of the compliers' outcome distribution, I note that Equations 3 and 4 implies a set of constraints, which has to be satisfied by the compliers' outcome distribution. This set of constraints include the the outcome distributions of all four subgroups are nondecreasing, lie between zero and one, imply the imposed sensitivity parameters, and have left and right limits equal to 0 and 1. Any function, which satisfies these constraints, and is right-continuous and has left-limits can represent the compliers' outcome distribution, as it implies a data generating process, which is consistent with the model assumptions and the sensitivity parameters.

The analytical expressions of the bounds are rather complicated and I provide them in the following. Let[11]

$$G_d^+(y) = \frac{1}{\pi_\Delta} \sup_{B \in \mathscr{B}} \{(\mathbb{P}(Y \in B, Y \leq y, D = d | Z = d) - \mathbb{P}(Y \in B, Y \leq y, D = d | Z = 1 - d))\}.$$

Intuitively, $G_d^+(y)$ gives a weighted lower bound on the compliers distribution. Let[12]

$$\underline{H}(y, \pi_{\mathrm{DF}}, \delta) = \frac{1}{\pi_{\mathrm{CO}}} \cdot \max\{0, Q_{\mathrm{dd}}(y) - \pi_d, \pi_\Delta G_d^{\sup}(y), \pi_{\mathrm{CO}} G_d^{\sup}(y) - \pi_{\mathrm{CO}} \frac{\pi_{\mathrm{DF}}}{\pi_\Delta} \delta\}, \tag{5}$$

[11]Under regularity conditions, the conditional densities q_{dd} and $q_{d(1-d)}$ exists and are continuous for $d \in \{0, 1\}$, it holds that $G_d^+(y) = \int_{\underline{y}}^y \max\{0, g_d(z)\} dz$, where $g_d(y) = \frac{\partial}{\partial y} G_d(y)$.

[12]It is clear that the population sizes π_{AT}, π_{NT}, and π_{CO} are functions of the sensitivity parameter π_{DF}. To simplify the notation, I let this dependency implicit.

where $\pi_\Delta = \pi_{\text{CO}} - \pi_{\text{DF}}$ and $G_d^{\sup}(y) = \sup_{\hat{y}\leq y} G_d(\hat{y})$. $G_d^{\sup}(y)$ is the smallest envelope function, which is nondecreasing and greater than $G_d(y)$ for any $y \in \mathbb{R}$. Let the lower bound of the compliers' outcome distribution be denoted by

$$\underline{F}_{Y_d^{CO}}(y, \pi_{\text{DF}}, \delta) = \\ \frac{1}{\pi_{\text{CO}}}\left(Q_{\text{dd}}(y) - \inf_{\tilde{y}\geq y}\left(Q_{dd}(\tilde{y}) - \left(\pi_\Delta G_d^+(\tilde{y}) - \inf_{\hat{y}\leq\tilde{y}}\left(\pi_\Delta G_d^+(\hat{y}) - \pi_{\text{CO}}\underline{H}(\hat{y}, \pi_{\text{DF}}, \delta)\right)\right)\right)\right). \quad (6)$$

Consider now the upper bound. Let

$$\overline{H}(y, \pi_{\text{DF}}, \delta) = \frac{1}{\pi_{\text{CO}}}\min\{1, Q_{dd}(y), \pi_\Delta(G_d^{\inf}(y) + \pi_{\text{DF}}), \pi_{\text{CO}}G_d^{\inf}(y) + \pi_{\text{CO}}\frac{\pi_{\text{DF}}}{\pi_\Delta}\delta\}, \quad (7)$$

where $G_d^{\inf} = \inf_{\hat{y}\leq y} G_d(\hat{y})$ denotes the greatest envelope function, which is nondecreasing and satisfies $G_d^{\inf}(y) \leq G_d(y)$ for all $y \in \mathbb{R}$. Let the upper bound of the compliers' outcome distribution be given by

$$\overline{F}_{Y_d^{CO}}(y, \pi_{\text{DF}}, \delta) = \\ \frac{1}{\pi_{\text{CO}}}\left(\pi_\Delta G_d^+(y) - \sup_{\tilde{y}\geq y}\left(\pi_\Delta G_d^+(\tilde{y}) - \left(Q_{dd}(\tilde{y}) - \sup_{\hat{y}\leq\tilde{y}}\left(Q_{dd}(\hat{y}) - \pi_{\text{CO}}\overline{H}(\hat{y}, \pi_{\text{DF}}, \delta)\right)\right)\right)\right). \quad (8)$$

Theorem 4.1 summarizes the results.

Theorem 4.1. *Suppose the instrument satisfies Assumption 1, and the data generating process is compatible with the sensitivity parameters* $(\pi_{\text{DF}}, \delta)$. *Then, for* $d \in \{0, 1\}$

$$\underline{F}_{Y_d^{CO}}(y, \pi_{\text{DF}}, \delta) \leq F_{Y_d^{CO}}(y) \leq \overline{F}_{Y_d^{CO}}(y, \pi_{\text{DF}}, \delta).$$

Moreover, there exist DGPs which are consistent with the above assumptions such that the compliers' outcome distribution equals either $\underline{F}_{Y_d^{CO}}(y, \pi_{\text{DF}}, \delta)$, $\overline{F}_{Y_d^{CO}}(y, \pi_{\text{DF}}, \delta)$, *or any convex combination of these bounds. Thus, the bounds are sharp.*

Theorem 4.1 shows not only that the proposed bounds are valid bounds but also that without imposing further assumptions the bounds cannot be tightened in a first order stochastic dominance sense.

Remark 1. Consider the set of all distribution functions, which are bounded in a first order stochastic dominance sense by the bounds of Theorem 4.1. This set includes functions which might not represent the compliers' outcome distribution, as these functions do not necessarily imply nondecreasing distribution functions of the other groups. However, the interest of the sensitivity analysis does not lie in the compliers' outcome distribution function bounds, but rather in the robust region. The conditions derived in Theorem 4.1 are sufficient to derive sharp bounds on many compliers' treatment effects of interest and therefore on sharp bounds on the robust and sensitivity regions.

Remark 2. Without imposing any assumption on the sensitivity parameter δ, the bounds simplify substantially, and the upper bound is give by

$$\overline{F}_{Y_d^{CO}}(y, \pi_{\mathrm{DF}}, 1) = \frac{1}{\pi_{\mathrm{CO}}} \max\{\pi_{\mathrm{CO}}, Q_{\mathrm{dd}}(y) - \pi_d, G_d^+(y)\}$$

and the lower bound by

$$\underline{F}_{Y_d^{CO}}(y, \pi_{\mathrm{DF}}, 1) = \frac{1}{\pi_{\mathrm{CO}}} \min\{0, G_d^+(y) - \pi_{\mathrm{DF}}, Q_{dd}(y)\}.$$

In the case of $\pi_{\mathrm{DF}} = \underline{\pi}_{DF}$, these bounds coincide with the bounds derived in Kitagawa (2020).

4.2. **Sensitivity Region.** In this section, I derive the sensitivity region of the estimator. I first show how the defiers' population size is sharply bounded, and second, how the sensitivity parameter about outcome heterogeneity between compliers and defiers is bounded for each population size of defiers.

4.2.1. *Bounding the Population Size of Defiers.* Both the average and the quantile compliers' treatment effects crucially depend on the population size of defiers. The population size of defiers is generally not point identified, but it is sharply bounded from below and from above, as I show in Lemma 4.2. The upper bound is denoted by

$$\overline{\pi}_{DF} = \min\{\mathbb{P}(D = 1|Z = 0), \mathbb{P}(D = 0|Z = 1)\}. \tag{9}$$

The intuition of this lower bound is simple: Suppose Assumption 1 holds, the first element of the minimum represents the sum of the population size of always takers and defiers, whereas the second one of never takers and defiers. The population size of defiers is smaller than both of these quantities. Let $\mathscr{B}$ denote the Borel σ-algebra. The lower bound on the population size of defiers is denoted by

$$\underline{\pi}_{DF} = \max_{s \in \{0,1\}} \{\sup_{B \in \mathscr{B}} \{\mathbb{P}(Y \in B, D = s|Z = 1 - s) - \mathbb{P}(Y \in B, D = s|Z = s)\}\}. \tag{10}$$

The reasoning is similar to the one above. The supremum is taken over all sets which imply a minimal bound on the population size of defiers. For example, the supremum is taken over the differences in the population sizes of defiers and compliers, which bounds the population size of defiers from below. This lower bound is similar to the instrument validity test statistic in Kitagawa (2015) and bounds presented in Balke and Pearl (1997). This reasoning is summarized in the following lemma.[13]

Lemma 4.2. *Suppose Assumption 1 holds. Then the population size of defiers is sharply bounded by* $\underline{\pi}_{DF} \leq \pi_{\mathrm{DF}} \leq \overline{\pi}_{DF}$.

[13] Huber et al. (2017) also present bounds on the population size of defiers, but as they remark their bounds are not sharp.

If $\underline{\pi}_{DF} > 0$, at least on of the classical LATE assumptions including monotonicity is violated, and this result follows from Kitagawa (2015). However, if the above inequalities contradict, i.e., $\underline{\pi}_{DF} > \overline{\pi}_{DF}$, the sensitivity region is empty. This implies that even though one allows for a violation of monotonicity, the independence assumption must be violated as well.

4.2.2. *Bounding Outcome Heterogeneity.* Based on Theorem 4.1, the sensitivity region of the sensitivity parameter space can be derived. Theorem 4.2 shows that the parameter space of π_{DF} is bounded from below and from above.

If for given sensitivity parameter values π_{DF} and δ the lower and upper bounds intersect, the two sensitivity parameters cannot be justified by the model as there would not exist any compliers' outcome distribution being consistent with the assumptions and satisfying the sensitivity parameter constraint. The domain of the sensitivity parameter δ is therefore bounded for a given value of the sensitivity parameter π_{DF} from below by

$$\underline{\delta}(\pi_{\text{DF}}) = \min_{d} \inf\{\delta : \inf_{y} \underline{F}_{Y_d^{CO}}(y, \pi_{\text{DF}}, \delta) - \overline{F}_{Y_d^{CO}}(y, \pi_{\text{DF}}, \delta) \geq 0\}. \tag{11}$$

The identified set of δ is further bounded from above. If the induced bounds of the compliers' and defiers' outcome distribution are such that for a given value of π_{DF} the maximal imposed difference of these two outcome distribution functions is smaller than the imposed outcome heterogeneity sensitivity parameter, this parameter is not binding. Furthermore, there cannot exists an outcome distribution of compliers and defiers justifying this heterogeneity and being consists with Assumption 1, the imposed value on the population size of defiers and the observed probabilities. Based on this observation, one can calculate for each given value of the population size of defiers the maximal implied outcome heterogeneity δ, denoted by $\overline{\delta}$. These bounds are sharp by first order stochastic dominance and are given by

$$\overline{\delta}(\pi_{\text{DF}}) = \max_{d\in\{0,1\}} \sup_{y}\{|\overline{F}_{Y_d^{CO}}(y, \pi_{\text{DF}}, 1) - \overline{F}_{Y_d^{DF}}(y, \pi_{\text{DF}}, 1)|, |\underline{F}_{Y_d^{CO}}(y, \pi_{\text{DF}}, 1) - \underline{F}_{Y_d^{DF}}(y, \pi_{\text{DF}}, 1)|\}. \tag{12}$$

If these conditions are violated for the weakest assumptions about the sensitivity parameters, $\pi_{\text{DF}} = \overline{\pi}_{DF}$ and $\delta = 1$, the sensitivity region is empty implying that there cannot be a compliers' outcome distribution satisfying the imposed assumptions. The independence assumption must be violated. The sensitivity region is therefore given by

$$SR = \{(\pi_{\text{DF}}, \delta) \in [\underline{\pi}_{DF}, \overline{\pi}_{DF}] \times [0, 1] : \underline{\delta}(\pi_{\text{DF}}) \leq \pi_{\text{DF}} \leq \overline{\delta}(\pi_{\text{DF}})\}$$

4.3. **Robust Region.** In the following, I derive the robust region which depends in contrast to the derived sensitivity region on the particular treatment effect of interest. To avoid to distinguish the degenerate case, I assume in the following that $\Delta_{CO}(\underline{\pi}_{DF}, \underline{\delta}(\underline{\pi}_{DF})) \geq \mu$ and that the sensitivity region is nonempty.[14]

By first order stochastic dominance of the bounds, sharp bounds of parameters of these bounds are obtained by evaluating the functionals characterized in Theorem 4.1 (Stoye, 2010). To give an example, the average treatment effects is given by $[\underline{\Delta}_{CO}(\pi_{\mathrm{DF}}, \delta), \overline{\Delta}_{CO}(\pi_{\mathrm{DF}}, \delta)]$ with

$$[\int_{\mathbb{Y}} y \; d\overline{F}_{Y_1^{CO}}(y, \pi_{\mathrm{DF}}, \delta) - \int_{\mathbb{Y}} y \; d\underline{F}_{Y_0^{CO}}(y, \pi_{\mathrm{DF}}, \delta), \int_{\mathbb{Y}} y \; d\underline{F}_{Y_1^{CO}}(y, \pi_{\mathrm{DF}}, \delta) - \int_{\mathbb{Y}} y \; d\overline{F}_{Y_0^{CO}}(y, \pi_{\mathrm{DF}}, \delta)].$$

The robust region of the claim that the average treatment effect is larger than a certain value μ is therefore characterized by the following set

$$RR_{CO}(\mu) = \{(\pi_{\mathrm{DF}}, \delta) \in \mathrm{SR} : \underline{\Delta}_{CO}(\pi_{\mathrm{DF}}, \delta) \geq \mu\}. \tag{13}$$

The breakdown point is given by

$$BP_{CO}(\pi_{\mathrm{DF}}, \mu) = \sup\{\delta : (\pi_{\mathrm{DF}}, \delta) \in \mathrm{SR} \text{ and } \underline{\Delta}_{CO}(\pi_{\mathrm{DF}}, \delta) \geq \mu\}.$$

It identifies the weakest assumption on outcome heterogeneity of compliers and defiers, such that the empirical conclusion holds. In contrast to the simple example of Section 2, the breakdown point as a function of the defiers population size is not necessarily decreasing in the defiers population size.

The breakdown frontier of the average treatment effect is the boundary of the robust region and given by the set of all breakdown points

$$BF_{CO}(\mu) = \{(\pi_{\mathrm{DF}}, \delta) \in \mathrm{SR} : \delta = BP_{CO}(\pi_{\mathrm{DF}}, \mu)\}. \tag{14}$$

The nonrobust region of the quantile treatment effect contains pairs of sensitivity parameters which only may not imply treatment effects being consistent with the empirical conclusion. I Figure 2 illustrates how the sets may look like.

5. EXTENSIONS

I present additional extensions of the model in this section. First, I show how this framework incorporates an analysis of other population parameters. Second, I discuss how covariates can be used within this framework, and third I illustrate a model of a binary outcome variable.

5.1. **Treatment Effects of other Groups.** Depending on the empirical question, the treatment effect of compliers might not be the only object of interest, and the researcher might be instead interested in other population parameters as well. As the proof of Theorem 4.1 requires the derivation of the outcome distributions of all four groups, the defiers' outcome distribution is sharply bounded by this theorem as well. Thus, the same framework can be used for a sensitivity analysis of the defiers treatment effect.

[14] If $\Delta_{CO}(\underline{\pi}_{DF}, \underline{\delta}(\underline{\pi}_{DF})) < \mu$, the robust region would be empty.

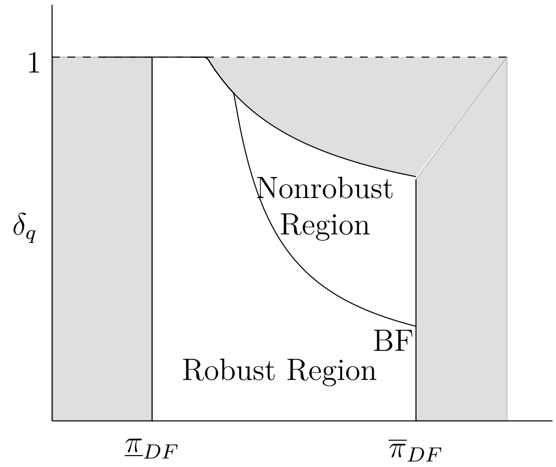


Figure 2: Sensitivity and Robust Region. Non-shaded region represents sensitivity region.

If one is rather interested in population treatment effects and is willing to assume that the support of the outcome variables is bounded, e.g., by $[\underline{y}, \overline{y}]$ sharp bounds of the potential outcomes are given in Proposition 5.1. Let therefore for $y \in [\underline{y}, \overline{y}]$

$$\underline{F}_{Y_d}(y, \pi_{\mathrm{DF}}, \delta) = \pi_{\mathrm{CO}} \cdot \underline{F}_{Y_0^{CO}}(y, \pi_{\mathrm{DF}}, \delta) + Q_{d(1-d)}$$
$$\overline{F}_{Y_d}(y, \pi_{\mathrm{DF}}, \delta) = \pi_d + \pi_{\mathrm{CO}} \cdot \overline{F}_{Y_0^{CO}}(y, \pi_{\mathrm{DF}}, \delta) + Q_{d(1-d)}.$$

Proposition 5.1. *Let Y_0 and Y_1 be distributed on $[\underline{y}, \overline{y}]$. Suppose the instrument satisfies the independence and relevance assumption, and the DGP is compatible with the sensitivity parameters $(\pi_{\mathrm{DF}}, \delta)$. Then, for $d \in \{0, 1\}$*

$$\underline{F}_{Y_d}(y, \pi_{\mathrm{DF}}, \delta) \leq F_{Y_d}(y, \pi_{\mathrm{DF}}, \delta) \leq \overline{F}_{Y_d}(y, \pi_{\mathrm{DF}}, \delta).$$

Moreover, there exist data generating processes which are consistent with the above assumptions such that the potential outcome distributions equal either $\overline{F}_{Y_d}(y, \pi_{\mathrm{DF}}, \delta)$, $\underline{F}_{Y_d}(y, \pi_{\mathrm{DF}}, \delta)$, or any convex combination of these bounds. Thus, the bounds are sharp.

To interpret these bounds The population size of compliers increases with the population size of defiers and the always and never takers decreases. As the bounds of the compliers outcome distribution lie between zero and one, it follows that the bounds for a smaller value of the population size of defiers first order stochastic dominate the bounds for greater values.

Based on these outcome distributions, sharp bounds on the average treatment effects are given by $[\underline{\Delta}(\pi_{\mathrm{DF}}, \delta), \overline{\Delta}(\pi_{\mathrm{DF}}, \delta)]$ with

$$[\int_{\mathbb{Y}} y \; d\underline{F}_{Y_d}(y, \pi_{\mathrm{DF}}, \delta) - \int_{\mathbb{Y}} y \; d\underline{F}_{Y_d}(y, \pi_{\mathrm{DF}}, \delta), \int_{\mathbb{Y}} y \; d\underline{F}_{Y_d}(y, \pi_{\mathrm{DF}}, \delta) - \int_{\mathbb{Y}} y \; d\overline{F}_{Y_d}(y, \pi_{\mathrm{DF}}, \delta)].$$

In contrast to the previous analysis, but based on the reasoning from above, the size of the identified set of the average treatment effect decreases in the population size of defiers.

Intuitively this is the case, as the bounds on the average treatment effect increase if the population size of always takers and respectively never takers is the largest, as in these cases both groups' outcome distributions are represented by the boundary of the support.

This reasoning aligns with results of Kitagawa (2020), who showed that imposing the monotonicity assumptions (e.g. $\pi_{\mathrm{DF}} = 0$) does not imply a smaller identified of the populations' average treatment effect. Similarly, it follows that in the case of $\pi_{\mathrm{DF}} = \underline{\pi}_{DF}$, these bounds equal the bounds derived in Kitagawa (2020) and for the special case of a binary outcome variable bounds derived in Balke and Pearl (1997); De Chaisemartin (2017); Machado et al. (2019) as shown in Appendix D.

In some empirical application, one might be willing to impose assumptions not only on outcome heterogeneity between compliers and defiers but also between other groups. To be precise, it might be reasonable to impose a sensitivity parameter, which is such that

$$\max_d \sup_y \{|F_{Y_d^T}(y) - F_{Y_d^{T'}}(y)|\}| \leq \delta_p \quad \forall T, T' \in \{AT, NT, CO, DF\},$$

where obviously $\delta_p \in [0, 1]$. If one is willing to impose this assumption across all or some groups, the bounds on population treatment effects might be substantially reduced. Based on this alternative sensitivity parameter, bounds on the outcome distributions of all four groups can be derived similarly as above.

5.2. **Other Treatment Effects.** In some applications, the parameter of interest is not only the average treatment effect but also e.g., quantile and distribution treatment effects. As Theorem 4.1 identifies the entire outcome distributions of all groups, these treatment effects are identified as well. Moreover, they are sharp by first order stochastic dominance of the bounds on the distribution functions.

To give an example, consider quantile treatment effects. Let define the lower and upper bounds of the quantile functions by the respectively left and right inverse of the bounds of the outcome distributions

$$\underline{Q}_{Y_d^{CO}}(\tau, \pi_{\mathrm{DF}}, \delta) = \sup\{y \in \mathbb{Y} : \overline{F}_{Y_d^{CO}}(y, \pi_{\mathrm{DF}}, \delta) \leq \tau\}$$
$$\overline{Q}_{Y_d^{CO}}(\tau, \pi_{\mathrm{DF}}, \delta) = \inf\{y \in \mathbb{Y} : \underline{F}_{Y_d^{CO}}(y, \pi_{\mathrm{DF}}, \delta) \geq \tau\}.$$

The quantile treatment effect of quantile τ is given by $[\underline{\Delta}_{CO}(\tau, \pi_{\mathrm{DF}}, \delta), \overline{\Delta}_{CO}(\tau, \pi_{\mathrm{DF}}, \delta)]$

$$[\underline{Q}_{Y_1^{CO}}(\tau, \pi_{\mathrm{DF}}, \delta) - \overline{Q}_{Y_0^{CO}}(\tau, \pi_{\mathrm{DF}}, \delta), \overline{Q}_{Y_1^{CO}}(\tau, \pi_{\mathrm{DF}}, \delta) - \underline{Q}_{Y_0^{CO}}(\tau, \pi_{\mathrm{DF}}, \delta)].$$

The construction of the compliers' outcome distribution bounds implies that these bounds might not be strictly increasing. The quantile functions are therefore not necessarily continuous even if one imposes continuity assumptions on the potential outcome distributions.

5.3. **Covariates.** In the previous analysis, I abstracted from covariates. In some empirical settings, the random assignment and exclusion restriction may only be valid conditional on covariates. Covariates can be used in this setup to tighten the identified bounds. Considering estimation and inference, the use of covariates might also increase the precision of the estimation. Assume that the covariates are discrete and given by $\{x_1, \ldots, x_K\}$, which splits the population in non-overlapping groups. I further impose the following assumptions.

Assumption 2. *(i) Conditional independence assignment:* $(Y_1, Y_0, D) \perp Z|X = x$*, (ii) Conditional relevance:* $\mathbb{P}(D = 1|Z = 1, X = x) > \mathbb{P}(D = 1|Z = 0, X = x)$*, (iii) Common support:* $0 < \mathbb{P}(Z = 1|X = x) < 1$.

I construct the sensitivity parameters such that for all $x \in \{x_1, \ldots, x_K\}$

$$\pi_{\mathrm{DF}}(x) \leq \pi_{\mathrm{DF}}.$$

This parameterization implies that the population size of defiers is bounded from above for each value of the covariate. I note that this kind of parameterization implies without further assumptions conservative bounds. The implied bounds are only non conservative if $\pi_{\mathrm{DF}}(x) = \pi_{\mathrm{DF}}$ for all values of x. However, this kind of assumption is very restrictive and often implausible. Alternatively, one could also consider a parameterization such that for all values of the covariates $\pi_{\mathrm{DF}}(x) = \pi^x_{\mathrm{DF}}$. However, this parametrization has the disadvantage that the parameter space might be very large and therefore difficult to interpret. By similar reasoning the heterogeneity in the outcome distribution is restricted by

$$|F_{Y_d^{CO}|X=x}(y|X = x) - F_{Y_d^{DF}|X=x}(y|X = x)| \leq \delta.$$

Based on the covariates one can calculate for each k, $\underline{\pi}_{DF}(x_k), \overline{\pi}_{DF}(x_k)$. The bounds on the sensitivity parameters can then be calculated based on the definition of the sensitivity parameters by $\underline{\pi}_{DF} = \max_k(\underline{\pi}_{DF}(x_k))$ and $\overline{\pi}_{DF} = \max_k(\overline{\pi}_{DF}(x_k))$. Let

$$\underline{F}^x_{Y_d^{CO}}(y, \pi_{\mathrm{DF}}, \delta) = \frac{1}{\pi_{\mathrm{DF}}} \sum_{k=1}^{K} \mathbb{P}(X = x_k, \pi^k_{\mathrm{DF}})\, \underline{F}^x_{Y_d^{CO}}(y, \pi_{\mathrm{DF}}, \delta|X = x_k)$$
$$\overline{F}^x_{Y_d^{CO}}(y, \pi_{\mathrm{DF}}, \delta) = \frac{1}{\pi_{\mathrm{DF}}} \sum_{k=1}^{K} \mathbb{P}(X = x_k, \pi^k_{\mathrm{DF}})\, \overline{F}^x_{Y_d^{CO}}(y, \pi_{\mathrm{DF}}, \delta|X = x_k).$$

Proposition 5.2. *Suppose the instrument satisfies conditional random assignment, conditional exclusion restriction and conditional relevance, and the DGP is compatible with the sensitivity parameters* $(\pi_{\mathrm{DF}}, \delta)$*. Then, for* $d \in \{0, 1\}$

$$\underline{F}_{Y_d}(y, \pi_{\mathrm{DF}}, \delta) \leq F_{Y_d}(y, \pi_{\mathrm{DF}}, \delta) \leq \overline{F}_{Y_d}(y, \pi_{\mathrm{DF}}, \delta).$$

Moreover, there exist DGPs which are consistent with the above assumptions such that the *compliers' outcome distribution equals either $\overline{F}_{Y_d^{CO}}(y, \pi_{\text{DF}}, \delta)$, $\underline{F}_{Y_d^{CO}}(y, \pi_{\text{DF}}, \delta)$, or any convex combination of these bounds, if for all $x \in \{x_1, \ldots, x_K\}$ it holds that $\pi_{\text{DF}}(x) = \pi_{\text{DF}}$ and there exists at least one y for each x such that $|F_{Y_d^{CO}|X=x}(y|X = x) - F_{Y_d^{DF}|X=x}(y|X = x)| = \delta$. Thus, the bounds are sharp.*

Based on this result, the derivation of the sensitivity and robust region follows the same arguments as in the unconditional case.

5.4. **Binary Outcome Variable.** The outcome of interest in many empirical applications is binary. I therefore consider in this section how sharp sets of the sensitivity regions and robust regions can be constructed for a binary outcome variable. In comparison to the general case, these bounds simplify substantially, and they therefore might be applicable in many empirical applications. Let $P_d^T = \mathbb{P}(Y_d^T = 1)$ and let the conditional joint probability of the outcome and the treatment status be given by $P_{ds} = \mathbb{P}(Y = 1, D = d|Z = s)$.

I bound the presence of defiers as above by π_{DF}. Outcome heterogeneity of compliers and defiers by the maximal difference in the expected outcomes of compliers and defiers, where the Kolmogorov -Smirnov of a binary outcome simplifies to

$$\delta_b = \max_{d \in \{0,1\}} |P_d^{CO} - P_d^{DF}|.$$

5.4.1. *Identification of sharp bounds of the marginal outcome distributions.* Following the same arguments as above, the sensitivity and robust region depend on the marginal outcome distributions of the compliers. I therefore specify the implied outcome distributions for all four groups in the presence and absence of treatment in the following[15]. Under treatment status d the following two equalities restrict the respectively population sizes are analog to Equation 3 $P_{dd} = \pi_{\text{CO}} P_d^{CO} + \pi_{\text{d}} P_d^{dT}$ and to Equation 4 $P_{d(1-d)} = \pi_{\text{DF}} P_d^{DF} + \pi_{\text{d}} P_d^{dT}$. It is clear from the same reasoning as above that if $\pi_{\text{DF}} = 0$ or $\delta_b = 0$, the compliers' outcome distribution is point identified in the presence and absence of treatment. For a given value of P_d^{CO} and the population size of defiers, it follows that the always and never takers as well as the defiers' outcome probabilities are point identified. Any potential compliers' outcome distribution has to fulfill that these probabilities are between zero and one. These conditions imply that the compliers' outcome probabilities are bounded from below and from above by

$$\underline{P}_d^{CO}(\delta, \pi_{\text{DF}}) = \max\{0, \frac{P_{dd} - \pi_{\text{d}}}{\pi_{\text{CO}}}, \frac{P_{dd} - P_{d(1-d)}}{\pi_{\text{CO}}}, \frac{P_{dd} - P_{d(1-d)} - \pi_{\text{DF}}\delta}{\pi_{\text{CO}} - \pi_{\text{DF}}}\}$$
$$\overline{P}_d^{CO}(\delta, \pi_{\text{DF}}) = \min\{1, \frac{P_{dd}}{\pi_{\text{CO}}}, \frac{P_{dd} - P_{d(1-d)} + \pi_{\text{DF}}}{\pi_{\text{CO}}}, \frac{P_{dd} - P_{d(1-d)} + \pi_{\text{DF}}\delta}{\pi_{\text{CO}} - \pi_{\text{DF}}}\}.$$

[15] The outcome distributions of the always takers in the absence of treatment and the one of the never takers in the presence of treatment are not identified and are therefore ignored in the following considerations.

If the data generating process satisfy the sensitivity parameters, then $\underline{P}_d^{CO}(\delta, \pi_{\mathrm{DF}}) \leq \underline{P}_d^{CO}(\delta, \pi_{\mathrm{DF}})$.

Corollary 5.3. *If the model assumptions and the sensitivity parameter pairs are satisfied, the compliers' outcome distributions are sharply bounded by* $\underline{P}_d^{CO} \leq P_d^{CO} \leq \overline{P}_d^{CO}$. *Moreover, the compliers' outcome distributions may attain any value in between these bounds.*

The proof of the Corollary 5.3 follows from similar arguments as above. The bounds of the outcome distributions of the other groups follow similarly.

5.4.2. *Sensitivity region.* The lower bound of the defiers' population size is given by

$$\underline{\pi}_{DF} = \max_{d \in \{0,1\}} \left\{ \sum_{y=0}^{1} \max\{0, \mathbb{P}(Y = y, D = d | Z = 1 - d) - P(Y = y, D = d | Z = d)\} \right\}.$$

The upper bound on the population size of defiers is given by Equation 9.

Based on Corollary 5.3, the sensitivity region can be derived. For any population size of defiers, the minimal and maximal bounds on the difference in the compliers' and defiers' outcome probability are such that the compliers' outcome probabilities are well-defined probabilities. It follows from basic algebra that the minimal bound on the outcome heterogeneity is given by[16]

$$\delta_b^{min}(\pi_{\mathrm{DF}}) = \frac{\underline{\pi}_{DF}}{\pi_{\mathrm{DF}}},$$

where $\delta_b^{min}(\pi_{\mathrm{DF}}) = 0$ if $\underline{\pi}_{DF} = 0$. The lower bound takes a very simple form and it decreases with the population size of defiers. The maximal bound on the outcome heterogeneity is given by the maximal difference of compliers and defiers' outcome probabilities

$$\delta_b^{max}(\pi_{\mathrm{DF}}) = \max_{d \in \{0,1\}} \max\{|\underline{P}_d^{CO}(\pi_{\mathrm{DF}}, 1) - \underline{P}_d^{DF}(\pi_{\mathrm{DF}}, 1)|, |\overline{P}_d^{CO}(\pi_{\mathrm{DF}}, 1) - \overline{P}_d^{DF}(\pi_{\mathrm{DF}}, 1)|\}.$$

The sensitivity parameter space is given by

$$\mathrm{SR}_b = \{(\pi_{\mathrm{DF}}, \delta_b) \in [\underline{\pi}_{DF}, \overline{\pi}_{DF}] \times [0, 1] : \delta_b^{min}(\pi_{\mathrm{DF}}) \leq \delta_b \leq \delta_b^{max}(\pi_{\mathrm{DF}})\}.$$

By construction this sensitivity region is such that any parameter pair within the sensitivity region is implied by some data generating process which is consistent with the model assumptions and implies the observed probabilities.

5.4.3. *Robust region.* The robust region for the claim that $\Delta_{CO} \geq \mu$ is given by all sensitivity parameter pairs within the sensitivity region that imply outcome distributions being consistent

[16] $\delta_b^{min} \in [0, 1]$ for all $\pi_{\mathrm{DF}} \in [\underline{\pi}_{DF}, \overline{\pi}_{DF}]$ by the definition of the lower bound of the defiers' population size $\underline{\pi}_{DF}$.

with this empirical conclusion. This reasoning implies that the robust region is given by

$$RR_b = \{(\pi_{\mathrm{DF}}, \delta_b) \in \mathrm{SR}_b : \underline{P}_1^{CO}(\pi_{\mathrm{DF}}, \delta_b) - \overline{P}_0^{CO}(\pi_{\mathrm{DF}}, \delta_b) \geq \mu\},$$

if the robust region is non-degenerated. Based on the simple algebra structure of the bounds of the outcome probabilities, a close form expression for the breakdown frontier can be easily derived. As this expression is rather complicated without providing much intuition, I show it in Appendix B.5.

6. ESTIMATION AND INFERENCE

While the main focus in this paper is on deriving the sensitivity and robust region for a certain empirical conclusion about the compliers' treatment effect, I discuss some methods for estimation and inference of these two regions in this section.[17] A more detailed discussion is deferred to Appendix A. I first present the goal of inference and then show how to obtain point estimates of the sensitivity and robust region. I propose two different inference procedures for a continuously and a binary distributed outcome variable.[18]

6.1. **Goal of Inference.** The parameter of interests in this sensitivity analysis are the sensitivity and the robust regions. I propose to construct confidence sets for both regions such that the confidence set for the sensitivity region is an outer confidence sets and the confidence set for the robust region is an inner confidence sets. These confidence sets should therefore jointly satisfy with probability approaching $1-\alpha$ that there does not exist any sensitivity parameter pair (i) of the sensitivity region, which does not lie within the confidence set for the sensitivity region and (ii) of the nonrobust region, which does not lie within the confidence set for the nonrobust region.[19] Denote by $\widehat{\mathrm{SR}}_B$ and $\widehat{RR}_B$ two sets of the sensitivity parameters. They are confidence sets of the sensitivity and respectively robust region if they imply that

$$\lim_{n\to\infty} \mathbb{P}(\mathrm{SR} \subseteq \widehat{\mathrm{SR}}_B \text{ and } \widehat{RR}_B(\mathrm{SR}) \subseteq RR(\mathrm{SR})) \geq 1-\alpha.$$

The identification argument of the sensitivity and robust regions are constructive. It follows from Section 4 that the population bounds of both the sensitivity and robust regions

[17]If one is rather interested in the identified set of the treatment effect for specific values of $(\pi_{\mathrm{DF}}, \delta)$, one can adopt a similar procedure. Conservative confidence sets for these sets can be directly applied from this presented procedure. To obtain nonconservative confidence sets one can howeever follow the ideas of Imbens and Manski (2004).

[18]If the outcome variable is distributed as an intermediate form, similar results could be obtained as the one discussed here under appropriate assumptions and corresponding estimators. I do not discuss this further here.

[19]If one is rather interested in the implied treatment effect of a particular choice of sensitivity parameters $(\pi_{\mathrm{DF}}, \delta)$, following the proposed methodology would imply conservative confidence sets of the parameter of interest. However, a non conservative procedure can be straightforwardly deduced from the presented procedure of this section (see also Imbens and Manski (2004) and the literature about partially identified parameters).

are given by the following mapping $\phi_{TE}(\cdot,\cdot) = (\underline{\pi}_{DF}(\cdot), -\overline{\pi}_{DF}(\cdot), \underline{\delta}(\cdot,\cdot), -\overline{\delta}(\cdot,\cdot), BP(\cdot,\cdot))$[20], which are mappings evaluated at $\pi_{\text{DF}} \in [0, 0.5)$ and $\theta(\cdot,\cdot,\cdot) = (Q_{(\cdot,\cdot)}(\cdot), G^+_{(\cdot)}(\cdot), P_{(\cdot)})$, where $Q_{(\cdot,\cdot)}(\cdot)$ denotes the conditional joint density $Q_{ds}(y)$ as a function of $(y, d, s) \in \mathbb{R} \times \{0,1\}^2$; $G^+_{(\cdot)}(\cdot)$ denotes $G^+_d(y)$ as a function of $(y, d) \in \mathbb{R} \times \{0,1\}$; and $P_{(\cdot)}$ denotes the probability of treatment $\mathbb{P}(D = 1|Z = z)$ as a function of the instrument $z \in \{0,1\}$. Let e_l be the l-th unit vector of length five. By the definition of the robust and sensitivity region, it suffices to find a joint lower confidence of the parameters in $\phi_{TE}(\widehat{\theta}, \pi_{\text{DF}})$, which implies to find a lower bound $\phi^L_{TE}(\widehat{\theta}, \pi_{\text{DF}})$ of the mapping $\phi_{TE}(\theta, \pi_{\text{DF}}))$ such that

$$\lim_{n\to\infty} \inf_{\pi_{\text{DF}}\in[0,0.5)} \mathbb{P}(\max_{l\leq 5} e'_l(\phi^L_{TE}(\widehat{\theta}, \pi_{\text{DF}}) - \phi_{TE}(\theta, \pi_{\text{DF}})) \leq 0) \geq 1 - \alpha.$$

6.2. **Estimation.** Estimates of the underlying parameters θ are simply obtained by replacing unknown population quantities by their corresponding nonparametric sample counterparts and by standard nonparametric kernel methods.[21] Point estimates of the sensitivity and robust regions can then be derived by simple plug-in. I defer a detailed description to Appendix A.1. I denote these estimated

6.3. **Inference for a continuous outcome variable.** I impose two sets of assumptions. The first set of assumptions imposes regularity conditions on the data generating process, which are common in the nonparametric literature. The second set of assumptions imposes conditions on parameters chosen by the researcher.

Assumption 3. *(i) For $z \in \{0,1\}$ $\{(Y^z_i, D^z_i)\}^{n_s}_{i=1}$ are identically and independently distributed according to the distribution of (Y^z, D^z) which is drawn conditional on $Z = z$ with support $\mathbb{Y} \times \{0,1\}$. It holds that n_0/n_1 converges to a nonzero constant as $n_o + n_1$ converges to infinity. (ii) $\mathbb{Y}_d$ is given by $[\underline{y}_d, \overline{y}_d]$ for $\infty < \underline{y}_d < \overline{y}_d < \infty$ for $d \in \{0,1\}$. (iii) $\forall d, z \in \{0,1\}$, the functions $q_{dz}(y)$ are bounded, absolutely continuous and two times continuously differentiable with uniformly bounded derivatives. (iv) For $d \in \{0,1\}$ $q_{dd}(z)$ and $q_{d(1-d)}(z)$ cross at a finite number of times.*

Assumption *(i)* is a standard. Assumption *(ii)* assumes compact support of the outcome variable. It simplifies the analysis, as the derivation of the bounds implies that mappings of the type $\sup_{z\leq y} F(z)$ have to be bounded from below and from above, where F is e.g. a continuous and nondecreasing real valued function. The compact support assumption simplifies this analysis. Assumption *(iii)* imposes smoothness conditions on the joint densities,

[20]The sign changes between the parameters is a pure normalization and it simplifies the notation in the following.

[21]Note that under mild regularity conditions on the densities of q_{ds} for $d, z \in \{0,1\}$ (e.g. the conditional joint outcome and treatment densities exist and are continuous), $G^+_d(y)$ can be equivalently represented by $\int_{\underline{y}}^{v} \max\{0, g_d(v)\}dv$, which can be estimated by replacing $g_d(v)$ by standard nonparametric kernel estimators. $\underline{y}$ is the left boundary of the support.

which are standard in the nonparametric literature. Assumption (vi) is imposed for simplicity and can be replaced by weaker assumptions.[22]

Assumption 4. (i) *The kernel is a 2nd order kernel function, being symmetric around zero, integrates to 1, twice continuously differentiable, of bounded variation and zero-valued off, say* $[-0.5, 0.5]$. (ii) *The bandwidth satisfies: (a)* $nh^4 \to 0$, *(b)* $nh^2 \to \infty$, *(c)* $nh/\log(n) \to \infty$.

Assumption (i) and (ii) impose conditions on the choice of kernel and the bandwidth, which can be satisfied by construction.

I first show that the estimators of the underlying parameters $\widehat{\theta}$ converge in $\sqrt{n}$ to a tight mean-zero Gaussian process indexed at $(y, d, s) \in \mathbb{R} \times \{0, 1\}^2$ under Assumptions 3. This result is summarized in the following Theorem.

Theorem 6.1. *Let Assumptions 3 and 4 hold. It follows that* $\sqrt{n}(\widehat{\theta}(y, d, z) - \theta(y, d, z)) \to \mathcal{Z}_1(y, d, z)$, *where* $\mathcal{Z}_1(y, d, z)$ *is a tight mean-zero Gaussian process in* $\ell^\infty(\mathbb{R} \times \{0, 1\}^2, \mathbb{R}^6)$.[23]

I give a characterization of this process in its proof. The mapping ϕ_{TE} is not Hadamard-differentiable, as it depends on minimum, maximum, supremum and infimum of unknown functions. Standard Delta-method arguments do therefore not apply in this setup (see Fang and Santos, 2018). I consider an inference procedure, which is asymptotically conservative, but valid. It is based on ideas that have been suggested in Chernozhukov et al. (2010); Haile and Tamer (2003); Masten and Poirier (2020) (see population smoothing).

In contrast to considering the mapping ϕ_{TE}, which yields sharp bounds of the sensitivity and robust region, I construct a smooth mapping, which yields outer sets of the sensitivity region and inner sets of the robust region. By construction, this smooth mapping is uniformly Hadamard-differentiable such that the standard Delta-method is applicable. Denote this smooth mapping by $\phi_{TE,\kappa}$, where $\kappa \in \mathbb{N}$ denotes a smoothing parameter. If the smoothing parameter increases the smoothed sensitivity and robust region are becoming indistinguishable from the original regions. This reasoning implies that confidence sets based on the smoothed versions are less conservative for the original sets as κ increases. In finite sample, one might, however face a trade-off between being too conservative and being willing to choose a smaller value of the smoothing parameter to ensure better approximation of the final estimator to normality. In the Appendix A.2, I show how the smoothed mappings can be constructed.

[22] This assumption ensures that my proposed estimator of $G_d^+(y)$ is well behaved. It is satisfied if the densities of the defiers weighted by the population size of defiers and the density of the compliers weighted by the population size of compliers equal each other only finitely many times. If one is not willing to impose this assumption, my proposed estimator of $\widehat{G}_d^+(y)$ is a biased estimator of $G_d^+(y)$. Following the arguments of Anderson, Linton, and Whang (2012), one can construct a debiased estimator of $G_d^+(\overline{y})$ and which converges in $\sqrt{n}$ to a mean-zero normal distribution. Based on similar arguments, one could now construct a debiased estimator of $G_d^+(y)$, such that this estimator converges to a tight mean-zero Gaussian process under relative mild regularity conditions. As this is a rather tedious exercise and not the purpose of this paper, I impose this stronger assumption, which might be reasonable in many relevant empirical examples.

[23] Let A be some arbitrary set and B a Banach space. Then $\ell^\infty(A, B)$ denotes the set of all mappings $f : A \to B$, which satisfy that $\sup_{a \in A} ||f(a)||_B \leq \infty$.

It then follows that plug-in estimators of the smoothed mappings converge in $\sqrt{n}$ to a Gaussian process. Instead of estimating the covariance structure of this process, I apply the nonparametric bootstrap to simulate this distribution. Consistency of this bootstrap procedure then follows from arguments of Fang and Santos (2018).

Choosing the critical values in the construction of the confidence sets involves a trade-off between the size of the sensitivity region and the robust region. I propose to choose the critical values such that the size of the sensitivity region is minimized and the size of the robust region is maximized. A more detailed analysis is deferred to Appendix A.

Let us denote by $\widehat{RR}_B(\kappa)$ a confidence set for the smoothed robust region and by $\widehat{\mathrm{SR}}_B(\kappa)$ a confidence set for the smoothed sensitivity region constructed through the described procedure. The following proposition then shows that these confidence sets are valid for the parameter of interest.

Proposition 6.2. *If Assumption 3 holds,* $\lim_{n\to\infty} \mathbb{P}(\widehat{RR}_B(\kappa) \subseteq RR,\ \mathrm{SR} \subseteq \widehat{\mathrm{SR}}_B(\kappa)) \geq 1-\alpha.$

6.4. **Inference for a binary outcome variable.** For the binary outcome model, it holds that for $\theta^b = (P_{11}, P_{10}, P_{01}, P_{00}, P_0, P_1) \in [0,1]^6$, the mapping yielding the sensitivity and the robust region is given by

$$\phi^b_{TE}(\theta^b, \pi_{\mathrm{DF}}) = (\underline{\pi}^b_{DF}(\pi_{\mathrm{DF}}), \overline{\pi}^b_{DF}(\pi_{\mathrm{DF}}), \delta^{min}_b(\theta^b, \pi_{\mathrm{DF}}), \delta^{max}_b(\theta^b, \pi_{\mathrm{DF}}), BP_p(\theta^b, \pi_{\mathrm{DF}})),$$

where its precise definition is given in Appendix A.3. The interpretation of ϕ^b follows the one for a continuously distributed outcome variable. Considering a binary outcome variable, it then follows from standard central limit arguments that the estimator of the underlying parameters is $\sqrt{n}$ normally distributed with mean zero. The mapping $\phi^b_{TE}(\theta^b, \pi_{\mathrm{DF}})$ is not Hadamard-differentiable at $(\theta^b, \pi_{\mathrm{DF}})$, as it consist among others of minimum and maximum of random functions. Standard Delta-method arguments are therefore not applicable. In this setting, one could simple use projection arguments to obtain valid confidence sets, which yield conservative confidence sets in general. I show instead that these mappings are uniformly Hadamard directionally differentiable for any finite set $\{\pi^k_{\mathrm{DF}}\}_{k=1}^K$.[24] Following the arguments of Fang and Santos (2018), the estimator of the boundaries of the sensitivity and robust region converge to some tight random process, which is a continuous transformation of Gaussian process, indexed at any finite set $\{\pi^k_{\mathrm{DF}}\}_{k=1}^K$. As this limiting distribution is rather complicated, I do not construct my inference procedure directly on its limiting distribution. Instead, one can choose various modified bootstraps to perform asymptotically valid inference e.g., subsampling or numerical-Delta-method (see Dümbgen, 1993; Hong and Li, 2018). However, as these procedures are relatively sensitivity to the choice of tuning parameters (as argued by Dümbgen, 1993), I follow a bootstrap method which relies on ideas based Andrews and Soares (2010); Bugni (2010); Fang and Santos (2018). I explain the procedure in detail

[24] Indeed, $\phi^b(\theta^b, \pi_{\mathrm{DF}})$ is not uniform Hadamard directionally differentiable with respect to π_{DF}.

in Appendix A.3. This procedure gives valid lower confidence sets for $\phi^b_{TE}(\theta^b, \pi_{\text{DF}})$ indexed at any finite set of sensitivity parameters $\{\pi^k_{\text{DF}}\}_{k=1}^K$. I obtain uniform convergence across π_{DF}, by interpolating the finite set through exploiting the functional form of the mappings. Let $\widehat{RR}^b_B$ and $\widehat{SR}^b_B$ denote these confidence sets. Proposition 6.3 shows that these confidence sets are valid.

Proposition 6.3. *Let Assumption 3(i) hold. Under the above construction,* $\lim_{n\to\infty} \mathbb{P}(\widehat{RR}^b_B \subseteq RR^b,\ \text{SR}^b \subseteq \widehat{\text{SR}}^b_B) \geq 1-\alpha$.

7. SIMULATIONS

I now study the finite sample performance of the proposed estimators of the sensitivity and robust regions through a Monte Carlo Study. I consider a number of different data generating processes with varying degrees of violations of monotonicity, which imply different sizes and shapes of both the sensitivity and robust regions. I show that this procedure leads to confidence sets which simulated coverage equal to its nominal level. Specifically, I consider the following population sizes $(\pi_{\text{CO}}, \pi_{\text{DF}}) \in \{(0.35, 0.05), (0.25, 0.15)\}$, where $\pi_{\text{DF}} = \pi_{\text{AT}} = 0.3$. I further assume that $\mathbb{P}(Z=1) = 0.5$ and generate the outcome as follows

$$Y_1^{CO} \sim \mathcal{B}(1, 0.5+\Delta_{CO}) \quad Y_1^{DF} \sim \mathcal{B}(1, 0.5+\Delta_{DF}) \quad Y_1^{AT}, Y_0^{NT}, Y_0^{DF}, Y_0^{CO} \sim \mathcal{B}(1, 0.5),$$

where $\Delta_{CO} \in \{0.2, 0.1\}$, $\Delta_{DF} \in \{0, -0.4\}$ and $\mathcal{B}(1,p)$ denotes the Bernoulli distribution with parameter p. All these data generating process imply that the independence and relevance assumption are satisfied. The empirical conclusion of interest it that the compliers' treatment effect is greater than zero. The Wald estimand is positive in each application and therefore overstates the empirical conclusion of interest. This reasoning implies that the robust region is nonempty in each of the DGPs. The bootstrap procedure requires to choose the smoothing parameter η, which is explained in Appendix A.3. I consider different values of η given by $\{0.2, 0.5, 1, 1.5, 2\}/\sqrt{N}$. The results are based on 10,000 Monte Carlo draws. I consider $\eta = 0.5$, $\eta = 1$, and $\eta = 1.5$ as reasonable choices in applications.

7.1. **Results.**

Table 2 shows the results of the simulated coverage rates at which the confidence sets I consider cover the population sensitivity and nonrobust region for the different data generating process and choices of tuning parameters. For the choices of tuning parameters, which I consider the simulated coverage is close to the nominal one in almost all data generating processes. However, if the tuning parameter is arbitrarily large or small chosen, the simulated coverage rate can deviate substantially from the nominal one. These results show that the confidence method performs reasonable well in finite samples.

Table 2: Simulated coverage rates of the sensitivity and robust region for a positive treatment effect.

π_{CO}	Δ_{CO}	Δ_{TE}	$\eta = 0.2$	$\eta = 0.5$	$\eta = 1$	$\eta = 1.5$	$\eta = 2$
0.35	0.3	0	99.1	97.8	95.1	91.3	90.8
	0.3	-0.3	96.9	94.3	91.2	92.5	91.6
	0.1	0	99.3	98.5	96.0	92.9	91.2
	0.1	-0.3	99.3	98.5	95.6	93.1	91.3
0.25	0.3	0	98.9	98.1	95.2	92.4	91.2
	0.3	-0.3	99.1	98.0	94.3	92.9	91.4
	0.1	0	99.3	98.6	96.0	93.5	91.2
	0.1	-0.3	99.0	97.7	94.3	92.9	90.9

The data generating process and the expressions follow the description of the text. Results are based on 10,000 Monte Carlo draws.

8. EMPIRICAL APPLICATION

In this section, I illustrate this sensitivity analysis in the empirical application of Angrist and Evans (1998). It is shown that even small violations of the monotonicity assumption may have a large impact on the robustness of the estimated treatment effects such that even the sign of the treatment effects may be indeterminate.

The same sex instrument in Angrist and Evans (1998) arguably satisfies Assumption 1: The independence assumption seems to be plausible by the following reasoning: First, the sex of a child is determined by nature, and therefore the random assignment assumption seems to be satisfied in this setting. Second, the exclusion restriction seems to hold as well as only the number of and not the sex of the child arguably influences the labor market outcome. The relevance assumption can be tested. However, monotonicity might be violated. To evaluate the robustness of the estimated treatment effects to a potential violation of monotonicity in this setting I apply the proposed sensitivity analysis. For simplicity, I focus on two outcome variables: the mother's labor market participation and the mother's annual wage.[25] The binary decision to treat therefore represent the extensive margin and the continuous outcome variable a mix of extensive and intensive variable. I use the same data and data restrictions as Angrist and Evans (1998).[26] The sample size is 211,983. The point estimated difference of the population sizes of compliers and defiers is given by 0.06.

8.1. **Sensitivity analysis for the labor market participation.** I consider labor market participation of the mother as the outcome variable. The Wald estimate is given by -0.13. Figure 3 illustrates the 95% confidence set for the sensitivity and the robust region for the claim that the compliers' treatment effect is negative. The formal definition of these sets is

[25] The annual wage is basically a continuously distributed running variable with a point mass at zero.

[26] Data are taken from Joshua D. Angrist's website `www.economics.mit.edu/faculty/angrist` from 1980. To have a comparable group of interest, the sample is restricted to women at the age of 20-36, having at least 2 children, being white and having their first child at the age of 19-25.

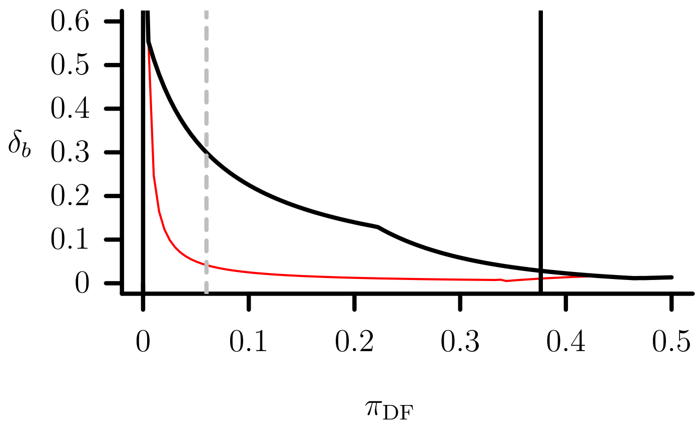


Figure 3: Confidence sets for the sensitivity and robust region for a negative compliers' treatment effect. The confidence level is 95 %. The compliers treatment effect is the effect of getting a third child on the labor market participation of mothers complying with the same sex instrument. The sensitivity region is bounded by the black lines and the red line indicates the boundary of the robust region.

given in Section 6.1. The population size of defiers is on the horizontal axis and outcome heterogeneity between compliers and defiers on the vertical axis.

In this example, a (conservative) 95% confidence set for π_{DF} is given by $[0, 0.37]$. One can therefore not conclude that monotonicity is violated in this example, which follows results in the literature Small et al. (2017). The sensitivity parameter pairs below the red line represent the robust region, which is the estimated set of sensitivity parameters implying a negative compliers' treatment effect. This figure shows that concerns about the validity of the monotonicity assumption haven be taken seriously. Without imposing any assumptions on the data generating process the hypothesis that the treatment effect is negative cannot be rejected, as $BP(0.37)$ is almost zero. If the population size of defiers increases, the breakdown frontier is relatively steeply declining and thus the robust region is rather small. This implies that as soon as defiers are present relatively strong assumptions on the compliers' and defiers' outcome distributions have to be imposed to imply a negative treatment effect. In contrast, if the population size of defiers is small, it is not necessary to impose strong assumptions about heterogeneity in the outcome variables to imply a negative treatment effect.

This example shows, that without imposing any assumptions on the data generating process only non informative conclusions can be drawn in this example. This is especially the case as the population size of defiers is not much restricted. To give some interpretation of these numbers, it is important to note that the upper bound is a rather conservative estimate, as if roughly 37% of the population were a defier, then approximately 43% of the population would have been a complier. This reasoning implies that roughly 90% of the population would base their decision to have a third child on the sex composition of the first two children. This reasoning seems to be in plausible.

One therefore might be willing to impose further assumptions to arrive at more interesting result and I show how one could plausibly proceed. These assumptions should only serve as an example and obviously they have to be always adapted to the analyzed situation. I adopt the approach of De Chaisemartin (2017). One of the most important inherently unknown quantity of interest is the population size of defiers. Imposing a smaller upper bound of this quantity, which is based on economic reasoning, allows to derive sharper results. Based on a survey conducted in the US, De Chaisemartin (2017) states that it seems reasonable that 5% of defiers is a conservative upper bound of the population size of defiers in this setting[27]. If one is willing to impose this assumption, one would still has to assume that the differences in the Kolmogorov-Smirnov norm is less than 0.05, which is quite a strong assumption. In this specific example, I would therefore rather conclude that the treatment effect is not robust to a potential violation.

8.2. **Sensitivity analysis for log annual wage.** I now consider the log annual income of the mother. This variable has a point mass at zero, representing all women who do not work, but is otherwise continuously distributed. The Wald estimate is given by -1.23. Figure 4 shows the corresponding 95% confidence sets for the robust and sensitivity region. If the monotonicity assumption were not violated, this estimate would imply that women who get a third child have a log annual wage reduced by 1.23.

Figure 4 shows 95 % confidence sets for both the sensitivity and robust region. The same interpretation applies as in the case of a binary outcome variable. One can see that without imposing any assumption about the population size of defiers, the empirical conclusion of a negative treatment effect is not robust to a potential violation of monotonicity. However, applying the same reasoning as above and imposing a maximal population size of 5% as an upper bound of the population size of defiers, one can see that the empirical conclusion is now robust to a potential violation of monotonicity.

To conclude, this sensitivity analysis is of interest, as one can identify the sign and the order of magnitude of the treatment effects by imposing further assumptions. These imposed assumptions are substantially weaker than the monotonicity assumption. The estimates therefore gain in credibility.

[27] *"In the 2012 Peruvian wave of the Demographic and Health Surveys, women were asked their ideal sex sibship composition. Among women whose first two kids is a boy and a girl, 1.8% had 3 children or more and retrospectively declare that their ideal sex sibship composition would have been two boys and no girl, or no boy and two girls. These women seem to have been induced to having a third child because their first two children were a boy and a girl. To my knowledge, similar questions have never been asked in a survey in the U.S. 1,8% could under or overestimate the share of defiers in the U.S. population. But this figure is, as of now, the best piece of evidence available to assess the percentage of defiers in Angrist & Evans (1998). 5% therefore sounds like a reasonably conservative upper bound."De Chaisemartin (2017) pp. 22-23*

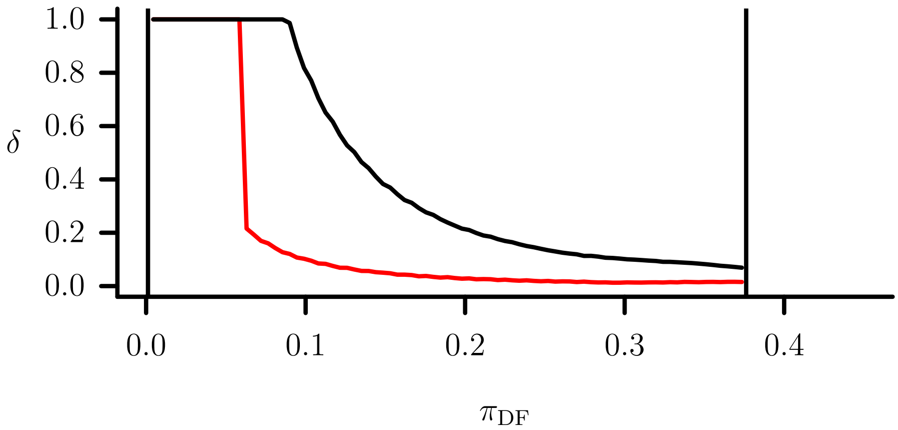


Figure 4: Confidence sets for the sensitivity and robust region for negative compliers' treatment effects. The confidence level is 95 %. The compliers treatment effect is the effect of getting a third child on the log annual wage of mothers complying with the same sex instrument. The sensitivity region is bounded by the black lines and the red line indicates the boundary of the robust region.

9. CONCLUSION

The local average treatment effect framework is popular to evaluate heterogeneous treatment effects in settings of endogenous treatment decisions and instrumental variables. In some empirical settings, one might doubt the validity of one of its key identifying assumptions the monotonicity assumption. Conducting a sensitivity analysis of the estimates in these settings improves the reliability of the results. This paper therefore proposes a new framework, which allows researchers to assess the robustness of the treatment effect estimates to a potential violation of monotonicity. It parameterize a violation of monotonicity by two parameters, the presence of defiers and heterogeneity of defiers and compliers. The former parameter is represented by the population size of defiers and the latter by the Kolmogorov-Smirnov norm bounding the outcome distributions of both groups. Based on these two parameters, I derive sharp identified sets for the compliers' average treatment effect and for any other group under further mild support assumption on the outcome variable. This identification results allows me to identify the set of sensitivity parameters which imply conclusions of treatment effect being consistent with the empirical conclusion. The empirical example of Angrist and Evans (1998) same sex instruments underlines the importance of the validity of the monotonicity assumptions as small violations of monotonicity may already lead to uninformative results.

REFERENCES

Anderson, G., O. B. Linton, and Y.-J. Whang (2012): "Nonparametric estimation and inference about the overlap of two distributions," *Journal of Econometrics*, 171, 1 – 23.

Andrews, D. W. and G. Soares (2010): "Inference for parameters defined by moment inequalities using generalized moment selection," *Econometrica*, 78, 119–157.

Andrews, I., M. Gentzkow, and J. M. Shapiro (2017): "Measuring the sensitivity of parameter estimates to estimation moments," *The Quarterly Journal of Economics*, 132, 1553–1592.

——— (2018): "On the informativeness of descriptive statistics for structural estimates," Tech. rep., National Bureau of Economic Research.

——— (2020): "Transparency in structural research," Tech. rep., National Bureau of Economic Research.

Andrews, I. and J. M. Shapiro (2020): "A Model of Scientific Communication," Tech. rep., National Bureau of Economic Research.

Angrist, J. and W. Evans (1998): "Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size," *American Economic Review*, 88, 450–77.

Angrist, J. D., G. W. Imbens, and D. B. Rubin (1996): "Identification of Causal Effects Using Instrumental Variables," *Journal of the American Statistical Association*, 91, 444–455.

Armstrong, T. B. and M. Kolesár (2018): "Sensitivity analysis using approximate moment condition models," *arXiv preprint arXiv:1808.07387*.

Balke, A. and J. Pearl (1997): "Bounds on Treatment Effects from Studies with Imperfect Compliance," *Journal of the American Statistical Association*, 92, 1171–1177.

Bonhomme, S. and M. Weidner (2018): "Minimizing sensitivity to model misspecification," *arXiv preprint arXiv:1807.02161*.

——— (2019): "Posterior average effects," *arXiv preprint arXiv:1906.06360*.

Bugni, F. A. (2010): "Bootstrap inference in partially identified models defined by moment inequalities: Coverage of the identified set," *Econometrica*, 78, 735–753.

Chen, X., E. Tamer, and A. Torgovitsky (2015): "Sensitivity Analysis in Semiparametric Likelihood Models," .

Chernozhukov, V., I. Férnandez–Val, and A. Galichon (2010): “Quantile and Probability Curves Without Crossing,” *Econometrica*, 78, 1093–1125.

Christensen, T. and B. Connault (2019): “Counterfactual sensitivity and robustness,” *arXiv preprint arXiv:1904.00989*.

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

Dahl, C. M., M. Huber, and G. Mellace (2019): “It’s never too LATE: A new look at local average treatment effects with or without defiers,” *Working Paper*.

De Chaisemartin, C. (2017): “Tolerating defiance? Local Average Treatment Effects without Monotonicity,” *Forthcoming: Quantitative Economics*.

Dümbgen, L. (1993): “On nondifferentiable functions and the bootstrap,” *Probability Theory and Related Fields*, 95, 125–140.

Fang, Z. and A. Santos (2018): “Inference on directionally differentiable functions,” *The Review of Economic Studies*, 86, 377–412.

Fiorini, M. and K. Stevens (2016): “Assessing the Monotonicity Assumption in IV and fuzzy RD designs,” *Working Papers: University of Sydney, School of Economics:2014-13*.

Frandsen, B. R., L. J. Lefgren, and E. C. Leslie (2019): “Judging Judge Fixed Effects,” *Working Paper*.

Freyberger, J. and Y. Rai (2018): “Uniform confidence bands: Characterization and optimality,” *Journal of Econometrics*, 204, 119–130.

Giné, E. and R. Nickl (2008): “Uniform central limit theorems for kernel density estimators,” *Probability Theory and Related Fields*, 141, 333–387.

Haile, P. A. and E. Tamer (2003): “Inference with an incomplete model of English auctions,” *Journal of Political Economy*, 111, 1–51.

Hong, H. and J. Li (2018): “ "The numerical delta method",” *Journal of Econometrics*, 206, 379 – 394.

Horowitz, J. L. and C. F. Manski (1995): “Identification and Robustness with Contaminated and Corrupted Data,” *Econometrica*, 63, 281–302.

Huber, M. (2014): “Sensitivity checks for the local average treatment effect,” *Economics Letters*, 123, 220 – 223.

——— (2015): "Testing the Validity of the Sibling Sex Ratio Instrument," *Labour*, 29, 1–14.

Huber, M., L. Laffers, and G. Mellace (2017): "Sharp IV Bounds on Average Treatment Effects on the Treated and Other Populations Under Endogeneity and Noncompliance," *Journal of Applied Econometrics*, 32, 56–79.

Huber, M. and G. Mellace (2012): "Relaxing monotonicity in the identification of local average treatment effects," *Discussion Paper 2012-12, Department of Economics, University of St. Gallen.*

——— (2015): "Testing instrument validity for LATE identification based on inequality moment constraints," *Review of Economics and Statistics*, 97, 398–411.

Imbens, G. W. (2003): "Sensitivity to Exogeneity Assumptions in Program Evaluation ," *American Economic Review*, 93, 126–132.

Imbens, G. W. and J. D. Angrist (1994): "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 62, 467–475.

Imbens, G. W. and C. F. Manski (2004): "Confidence intervals for partially identified parameters," *Econometrica*, 72, 1845–1857.

Kitagawa, T. (2015): "A Test for Instrument Validity," *Econometrica*, 83, 2043–2063.

——— (2020): "Identification region of the potential outcome distributions under instrument independence," *CeMMAP working papers.*

Kitamura, Y., T. Otsu, and K. Evdokimov (2013): "Robustness, infinitesimal neighborhoods, and moment restrictions," *Econometrica*, 81, 1185–1201.

Klein, T. J. (2010): "Heterogeneous treatment effects: Instrumental variables without monotonicity?" *Journal of Econometrics*, 155, 99 – 116.

Kline, P. and A. Santos (2013): "Sensitivity to missing data assumptions: Theory and an evaluation of the U.S. wage structure," *Quantitative Economics*, 4, 231–267.

Machado, C., A. M. Shaikan, and E. J. Vytlacil (2019): "Instrumental Variables, Monotonicity Restrictions, and the Sign of the average Treatment Effect," *forthcoming: Journal of Econometrics.*

Manski, C. F. (1990): "Nonparametric Bounds on Treatment Effects," *The American Economic Review*, 80, 319–323.

Manski, C. F. and J. V. Pepper (2000): "Monotone Instrumental Variables: With an Application to the Returns to Schooling," *Econometrica*, 68, 997–1010.

Masten, M. A. and A. Poirier (2018): "Salvaging Falsified Instrumental Variable Models," *arXiv:1812.11598.*

——— (2020): "Inference on Breakdown Frontiers," *Quantitative Economics*, 11, 41–111.

Mogstad, M., A. Torgovitsky, and C. R. Walters (2019): "Identification of Causal Effects with Multiple Instruments: Problems and Some Solutions," Working Paper 25691, National Bureau of Economic Research.

Mukhin, Y. (2018): "Sensitivity of regular estimators," *arXiv preprint arXiv:1805.08883.*

Norris, S. (2019): "Examiner Inconsistency: Evidence from Refugee Appeals," *Working Paper.*

Norris, S., M. Pecenco, and J. Weaver (2019): "The Effects of Parental and Sibling Incarceration: Evidence from Ohio," *Working Paper.*

Rambachan, A. and J. Roth (2019): "An honest approach to parallel trends," Tech. rep.

Shapiro, A. (1991): "Asymptotic analysis of stochastic programs," *Annals of Operations Research*, 30, 169–186.

Small, D. S., Z. Tan, R. R. Ramsahai, S. A. Lorch, and M. A. Brookhart (2017): "Instrumental Variable Estimation with a Stochastic Monotonicity Assumption," *Statist. Sci.*, 32, 561–579.

Stoye, J. (2005): "Essays on partial identification and satistical decisions," *Ph.D. thesis, Northewstern University.*

——— (2010): "Partial Identification of spread parameters," *Quantitative Economics*, 1, 323–357.

van der Vaart, A. W. (1998): *Asymptotic Statistics*, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press.

Vytlacil, E. (2002): "Independence, monotonicity, and latent index models: An equivalence result," *Econometrica*, 70, 331–341.

Appendix

Appendix A presents more details on how the sensitivity and robust region can be estimated and outlines in detail how inference can be conducted. Appendix B presents the proofs to the main results of the main text. Additional results are collected in Section C and D

A. MORE DETAILS ON ESTIMATION AND INFERENCE

In this section, I present more details on the estimation and inference method proposed in the main text. I first provide details about the estimation of the sensitivity and robust regions. I then proceed by showing results for conducting inference in a model of a continuous and a binary outcome variable separately.

A.1. **Estimation.** In this section, I give further details on the construction of the estimators described in Section 6. As explained in Section 6, it suffices to study the mapping $\phi_{TE}(\theta, \pi_{\text{DF}})$ when studying the sensitivity and robust regions.

I estimate the conditional treatment probabilities by $\widehat{\mathbb{P}}(D = d|Z = z) = \sum_{i=1}^{n_z} \mathbb{1}[D_i^z = d]$. The difference in compliers and defiers population sizes as well as any population sizes given a sensitivity parameter value π_{DF} are simple linear combinations of these quantities and estimated by plug-in. The conditional distribution functions are estimated by $\widehat{Q}_{dz}(y) = \sum_{i=1}^{n} \mathbb{1}[Y_i^z \leq y] \cdot \mathbb{1}[D_i^z = d]$. Moreover, let $K_h(\cdot) = K(\cdot/h)/h$, where $K(\cdot)$ denotes a kernel function and $h > 0$ a bandwidth. I define $\widehat{G}_d^+(y) = \sum_{i=1}^{n} \mathbb{1}[Y_i^z \leq y]\widehat{g}_d(y)$, where $\widehat{g}_d(y) = \widehat{q}_{dd}(y) - \widehat{q}_{d(1-d)}(y)$. I estimate the conditional joint densities by standard nonparametric kernel estimators

$$\widehat{q}_{dz}(y) = \frac{1}{n_z h} \sum_{i=1}^{n_z} K_h(Y_i^z - y) \cdot \mathbb{1}[D_i^z = d].$$

Based on these estimators, the parameters of θ can be estimated and therefore by plug-in the mapping $\phi_{TE}(\theta, \pi_{\text{DF}})$.

A.1.1. *Binary.* In the binary outcome model, the mapping of interest is $\phi^b(\theta^b, \pi_{\text{DF}})$. As shown in Section 5.4, θ^b is given by $(P_{11}, P_{10}, P_{01}, P_{00}, P_0, P_1)$. I estimate the probabilities by simple replacing them with their sample counterparts, e.g. $P_{d,s} = \frac{1}{n_s} \sum_{i=1}^{n_s} \mathbb{1}[Y^s = 1, D^s = d]$ and correspondingly $P_s = \frac{1}{n_s} \sum_{i=1}^{n_s} \mathbb{1}[D^s = 1]$. Correspondingly, I estimate $\phi^b(\widehat{\theta}^b, \pi_{\text{DF}})$ by simple plug-in estimates.

A.2. **Inference for a continuous outcome variable.** I explain in this section in detail how to construct joint lower confidence intervals for the parameters of the continuous outcome model $\phi_{TE}(\theta, \pi_{\text{DF}})$.

A.2.1. *Smoothing Functions.* The definition of the smooth mappings under considerations is presented in the following definition.

Definition A.1 (Definition 1, Masten and Poirier (2020).)**.** *Let* $(\Theta, \|\cdot\|_\Theta)$ *and* $(\mathcal{H}, \|\cdot\|_\mathcal{H})$ *be Banach spaces. Let* $\leq$ *be a partial order on* $\mathcal{H}$*. Let* $h : \Theta \to \mathcal{H}$ *be a function. Consider a function* $H_\kappa : \Theta \to \mathcal{H}$*, where* $\kappa \in \mathbb{R}_+^{\dim(\kappa)}$ *is a vector of smoothing parameters. Then* H_κ *denotes a smooth lower approximation of* H *if*

1. *Lower envelope:* $H_\kappa(\theta) \leq H(\theta)$ *for all* $\theta \in \Theta$ *and* $\kappa \in \mathbb{R}_+^{\dim(\kappa)}$.

2. *Approximating: For each* $\theta \in \Theta$*,* $H_\kappa(\theta) \to H(\theta)$ *for* $\kappa \to \infty$ *(pointwise).*

3. *Smoothing:* H_κ *is Hadamard-differentiable.*

Let $(\Theta, \|\cdot\|_\Theta)$ and $(\mathcal{H}, \|\cdot\|_\mathcal{H})$ be Banach spaces, where $\leq$ is a partial order on $\mathcal{H}$. I consider two mappings $f, g : \Theta \to \mathcal{H}$ in the following, which are both Hadamard-differentiable. Let $\kappa > 1$ be the smoothing parameter. I first consider the function $\psi_{av}(f) = |f|$, where a smooth upper and lower approximation is given by

$$\psi_{av}^U(f;\kappa) = \sqrt{f^2 + 1/\kappa} \qquad \psi_{av}^L = \frac{f^2}{\sqrt{f^2 + 1/\kappa}}.$$

Lemma A.2. $\psi_{av}^L(f;\kappa)$ *is a smooth lower approximation to the mapping* $\psi_{av}(f)$ *and* $\psi_{av}^U(f;\kappa)$ *a smooth upper approximation.*

By fundamental algebra $\psi_{\min}(f,g) = \min(f,g) = \frac{1}{2}(f+g-|f-g|)$ and that $\psi_{\max}(f,g) = \max(f,g) = \frac{1}{2}(f+g+|f-g|)$. By the reasoning from above, these functions are bounded from below and from above by replacing $|f-g|$ by ψ_{av}^L and $\psi_{av}^U(f;\kappa)$, respectively, e.g. a smooth lower approximation of $\psi_{\max}(f,g)$ is given by $\psi_{\max}^L(f,g;\kappa) = f+g+\psi_{av}^L(f-g;\kappa)$ and a smooth upper approximation is given by $\psi_{\max}^U(f,g;\kappa) = f+g+\psi_{av}^U(f-g;\kappa)$. It follows from a simple induction argument, that one can generalize this procedure to the maximum of a set of finitely many mappings, e.g. $\psi_{\max}^U(\{f_j\}_{j=1}^J;\kappa) = f_J + \psi_{\max}^U(\{f_j\}_{j=1}^{J-1};\kappa) + \psi_{av}^U(f - \psi_{\max}^U(\{f_j\}_{j=1}^{J-1};\kappa);\kappa)$.

In the following, I consider the mapping $\psi_{\sup,\leq}(f,g)(\cdot) = \sup_{z\leq\cdot} f(z) - g(z)$. I consider the equally binned set $\mathbb{Y} = \bigcup_{k=1}^{\kappa}[\underline{y} + (k-1)d_Y, kd_Y]$, where $d_Y = \frac{1}{\kappa}(\overline{y} - \underline{y})$. Let $k_j = \underline{y} + j \cdot d_y$, where $j \in \{0, 1, 2, \ldots, \kappa\}$.

$$\begin{aligned}
\psi_{\sup,\leq}^L(f,g;\kappa)(\cdot) &= \psi_{\max}^L(\{g(k_j) - f(k_j)\}_{j:k_j\leq\cdot};\kappa). \\
\psi_{\sup,\leq}^U(f,g;\kappa)(\cdot) &= \psi_{\max}^U(\{g(k_j) - f(\min(y, k_{j+1}))\}_{j:k_j<\cdot}\,;\kappa).
\end{aligned}$$

I similarly define for the mapping $\psi_{\inf,\leq}(f,g)(\cdot) = \inf_{z\leq\cdot} f(z) - g(z)$ that

$$\begin{aligned}
\psi_{\inf}^U(f,g;\kappa)(\cdot) &= \psi_{\min}^U(\{g(k_j) - f(k_{j+1}))\}_{j:k_j\leq\cdot};\kappa). \\
\psi_{\inf}^L(f,g;\kappa)(\cdot) &= \psi_{\min}^L(\{g(\min(\cdot, k_{j+1}) - f(k_j))\}_{j:k_j<\cdot};\kappa).
\end{aligned}$$

Lemma A.3. *If f and g are monotone increasing, $\psi^L_{\inf,\leq}(f,g;\kappa)$ is a smooth lower approximation and $\psi^U_{\sup,\leq}(f,g;\kappa)$ a smooth upper approximation to the function $\psi_{\sup,\leq}(f,g)$, and $\psi^U_{\inf,\leq}(f,g;\kappa)$ a smooth upper approximation and $\psi^L_{\inf,\leq}(f,g;\kappa)$ a smooth lower approximation to the function $\psi_{\inf,\leq}(f,g)$.*[28]

A.2.2. *Smoothing the sensitivity and robust regions.* In this section, I explain in detail how to construct the smoothed mappings, which yield the smoothed version of $\phi_{TE}(\theta,\pi_{\mathrm{DF}})$ which is denoted by $\phi_{TE,\kappa}(\theta,\pi_{\mathrm{DF}})$. First, I illustratively show how to smooth the lower bound from above. I note that $G_d^{\sup}(y) = \sup_{z\leq y} G_d^+(z) - G_d^-(z)$. I denote respectively the upper and lower bounds by

$$G_d^{\sup,U}(y;\kappa) = \psi^U_{\sup,\leq}(G_d^+ - G_d^-;\kappa)(y)$$

The bounds on the compliers' outcome distributions are therefore bounded by

$$\underline{H}^U(y,\pi_{\mathrm{DF}},\delta;\kappa) = \frac{1}{\pi_{\mathrm{CO}}}\cdot\psi^U_{\max}\left(\left\{0,\pi_\Delta G_d^{\sup,U}(y;\kappa), Q_{\mathrm{dd}}(y)-\pi_d, \frac{\pi_{\mathrm{CO}}}{\pi_\Delta}(G_d^{\sup,U}(y;\kappa)-\pi_{\mathrm{DF}}\delta)\right\};\kappa\right)$$

A smooth upper approximation of $\underline{F}_{Y_d^{CO}}(y,\pi_{\mathrm{DF}},\delta)$ is given by $\underline{F}^U_{Y_d^{CO}}(y,\pi_{\mathrm{DF}},\delta;\kappa) = \frac{1}{\pi_{\mathrm{CO}}}\cdot$

$$\left(Q_{\mathrm{dd}}(y) + \psi^L_{\sup,\geq}\left(\pi_\Delta G_d^+(\tilde{y}) - Q_{dd}(\tilde{y}) - \psi^L_{\inf,\leq}\left(\pi_\Delta G_d^+(\hat{y}) - \pi_{\mathrm{CO}}\underline{H}^U(y,\pi_{\mathrm{DF}},\delta;\kappa);\kappa\right)(\hat{y})\;;\kappa\right)(\tilde{y})\right).$$

A smooth lower approximation of $\overline{F}_{Y_d^{CO}}(y,\pi_{\mathrm{DF}},\delta)$ can be similarly constructed as well as a smooth lower and upper approximation of $\overline{F}_{Y_d^{CO}}(y,\pi_{\mathrm{DF}},\delta)$.

The sensitivity region can be similarly derived, e.g.

$$\overline{\pi}^U_{DF} = \psi^U_{\min}(\{\mathbb{P}(D=1|Z=0),\mathbb{P}(D=0|Z=1)\};\kappa)$$

and similarly

$$\overline{\pi}^L_{DF} = \frac{\pi_\Delta}{\pi_{\mathrm{CO}}}\psi^L_{\max}(\{G_d^+(\overline{y}),G_d^+(\overline{y})\};\kappa) - 1.$$

The same reasoning applies to the mappings of $\underline{\delta}$ and $\overline{\delta}$.

I can smoothly bound $\underline{\delta}$ from above. I therefore note that just by construction $\underline{F}^U_{Y_d^{CO}}(y,\pi_{\mathrm{DF}},\delta)$ and $\overline{F}^L_{Y_d^{CO}}(y,\pi_{\mathrm{DF}},\delta)$ are Hadamard-differentiable in their arguments. Moreover, they are strictly increasing in δ if δ is small and therefore taking the inf is a Hadamard-differentiable by Lemma 21.4 van der Vaart (1998). A smooth lower approximation is therefore given by

$$\underline{\delta}^L(\pi_{\mathrm{DF}}) = \psi^L_{\min}(\{\delta : \inf_y \overline{F}^U_{Y_d^{CO}}(y,\pi_{\mathrm{DF}},\delta) - \underline{F}^L_{Y_d^{CO}}(y,\pi_{\mathrm{DF}},\delta)\geq 0\};\kappa).$$

[28]By similar reasoning, the functions $\psi_{\sup,\geq}(f,g)$ and $\psi_{\inf,\geq}(f,g)$ can be smoothly approximated.

A smooth upper approximation of $\overline{\delta}(\pi_{\text{DF}})$ is given by

$$\begin{aligned}\overline{\delta}^{U}(\pi_{\text{DF}}) =&\psi_{\max}(\Big\{\psi_{\sup,\leq}(\overline{F}_{Y_d^{CO}}(y,\pi_{\text{DF}},1),\overline{F}_{Y_d^{DF}}(y,\pi_{\text{DF}},1))(\overline{y};\kappa),\\ &\psi_{\sup,\leq}(\overline{F}_{Y_d^{DF}}(y,\pi_{\text{DF}},1),\overline{F}_{Y_d^{CO}}(y,\pi_{\text{DF}},1))(\overline{y},\kappa),\\ &\psi_{\sup,\leq}(\underline{F}_{Y_d^{CO}}(y,\pi_{\text{DF}},1),\underline{F}_{Y_d^{DF}}(y,\pi_{\text{DF}},1))(\overline{y},\kappa),\\ &\psi_{\sup,\leq}(\underline{F}_{Y_d^{DF}}(y,\pi_{\text{DF}},1),\underline{F}_{Y_d^{CO}}(y,\pi_{\text{DF}},1))/(\overline{y},\kappa)\Big\};\kappa))\end{aligned}$$

Based on these definitions, I can smoothly bound the compliers' treatment effect, where the lower bound is given by

$$\underline{\Delta}^{L}_{CO}(\pi_{\text{DF}},\delta)=\int_{\mathbb{Y}} y\; d\overline{F}^{U}_{Y_1^{CO}}(y,\pi_{\text{DF}},\delta)-\int_{\mathbb{Y}} y\; d\underline{F}^{L}_{Y_0^{CO}}(y,\pi_{\text{DF}},\delta).$$

$\Delta^{L}_{CO}(\pi_{\text{DF}},\delta)$ is thus by construction a smooth lower approximation to $\Delta_{CO}(\pi_{\text{DF}},\delta)$. The breakdown point can be derived by

$$BP^{L}(\pi_{\text{DF}})=\sup\{\delta\leq\delta_{\max}(\pi_{\text{DF}}):\underline{\Delta}^{L}_{CO}(\pi_{\text{DF}},\delta)\geq\mu\}.$$

As $\Delta^{L}_{CO}(\pi_{\text{DF}},\delta)$ is strictly increasing in δ, $BP(\pi_{\text{DF}})$ is Hadamard-differentiable. It follows that $BP^{L}(\pi_{\text{DF}})$ is a smooth lower approximation of $BP(\pi_{\text{DF}})$.

The following Theorem summarizes the result.

Theorem A.4. *Let Assumption 3 and 4 hold. It holds that uniformly over $\pi_{\text{DF}}\in[0,0.5]$, $\sqrt{n}(\phi_{TE,\kappa}(\widehat{\theta},\pi_{\text{DF}})-\phi_{TE,\kappa}(\theta,\pi_{\text{DF}}))\to\mathcal{Z}_2(\pi_{\text{DF}})$, where $\mathcal{Z}_2(\pi_{\text{DF}})$ is a tight Gaussian process in $\ell^{\infty}([0,0.5],\mathbb{R}^5)$, for which $\phi_{TE,\kappa}(\widehat{\theta},\pi_{\text{DF}})$ lies in the interior of the sensitivity parameter space.*

It follows by Theorem Fang and Santos (2018) that in this case the nonparametric bootstrap is consistent and therefore, if $\theta^{\star}$ denotes a draw from the nonparameric boostrap, it follows that

$\sqrt{n}(\phi_{TE,\kappa}(\theta^{\star},\pi_{\text{DF}})-\phi_{TE,\kappa}(\widehat{\theta},,\pi_{\text{DF}}))\to\mathcal{Z}_2(\pi_{\text{DF}})$, where $\mathcal{Z}_2(\pi_{\text{DF}})\to\phi'_{\theta}(\mathcal{Z}_1),\phi)$

$$z_{1-\alpha}=\inf\{z\in\mathbb{R}:\mathbb{P}(\sup_{\pi_{\text{DF}}\in[0,0.5],l\leq 5}\sqrt{n}e_l(\phi_{TE,\kappa}(\theta^{\star},\pi_{\text{DF}})-\phi_{TE,\kappa}(\widehat{\theta},\pi_{\text{DF}}))\leq z|\{D_i,Y_i\})\geq 1-\alpha\}$$

If quantile is strictly increasing, then $\widehat{z}_{1-\alpha}=z_{1-\alpha}+o_p(1)$, which then in turn implies that

$$\mathbb{P}(\widehat{R}R(\kappa)\subset RR(\kappa)\text{ and }SR(\kappa)\subset\widehat{S}R(\kappa))\geq 1-\alpha.$$

I therefore consider a random set of $\{\pi^{k}_{\text{DF}}\}_{k=1}^{K}\in[0,0.5)^K$ implying that the critical values are chosen such that $\min \iota' \sum_{k=1}^{K} cv(\pi_{\mathrm{DF}}^{k})$

$$\mathbb{P}(\max_{k\leq K, l\leq 5} e_l' \left(\phi_{TE,\kappa}(\widehat{\theta},\pi_{\mathrm{DF}}^{k}) - cv(\pi_{\mathrm{DF}}^{k}) - \phi_{TE,\kappa}(\widehat{\theta},\pi_{\mathrm{DF}}^{k})'\right) \leq 0) = 1-\alpha + o_P(1),$$

$cv(\pi_{\mathrm{DF}}^{k})$ can either be chosen to be constant in each of its elements or might depend on π_{DF}^{k}.

A.3. **Inference for a binary outcome variable.** In this section, I present more details on how to conduct inference for a binary outcome variable. For this purpose, I will explicitly write down the expression for the mapping of interest, and propose the bootstrap method. Based on the derivation in Section 5.4, it follows that for some functions $\psi_{i,j}(\theta^b,\pi_{\mathrm{DF}})$, which are uniformly Hadamard differentiable with respect to both θ^b and π_{DF}

$$\begin{aligned}\phi^b(\theta^b,\pi_{\mathrm{DF}}) = \Big(&\max(\{\psi_{1,j}(\theta^b,\pi_{\mathrm{DF}})\}_{j=1}^{4}), \min(\{\psi_{2,j}(\theta^b,\pi_{\mathrm{DF}})\}_{j=1}^{2}), \max(\{\psi_{3,j}(\theta^b,\pi_{\mathrm{DF}})\}_{j=1}^{2}),\\ &\min(\{\psi_{4,j}(\theta^b,\pi_{\mathrm{DF}})\}_{j=1}^{2}), \min(\{\psi_{5,j}(\theta^b,\pi_{\mathrm{DF}})\}_{j=1}^{16})\Big)\end{aligned}$$

In the following, I consider the mapping $\psi_i(\theta^b,\pi_{\mathrm{DF}}) = \max(\{\psi_{i,j}(\theta^b,\pi_{\mathrm{DF}})\}_{j=1}^{4})$. It follows that this mapping is Hadamard-directionally-differentiable for any fix π_{DF}, with directional derivative given by

$$\psi_{i,h}'(\theta^b,\pi_{\mathrm{DF}}) = \max_{j:\psi_{i,j}(\theta^b,\pi_{\mathrm{DF}})>\max_{k\neq j}\{\psi_{i,k}(\theta^b,\pi_{\mathrm{DF}})\}} h_{i,j},$$

where $h_{i,j}$ is the

I denote an estimator of this directional derivative by $\widehat{\psi}_h'(x) = \max_{j:x_j>\max_{k\neq j}\{x_k\}+\kappa} h_j$. where $\kappa > 0$ converges to zero as n goes to infinity. In the application and the Monte Carlo simulation, I consider $\kappa \sim 1/\sqrt{n}$. The algorithm for the bootstrap is as follows.

1. Get estimate of θ^b and $\phi(\theta^b,\pi_{\mathrm{DF}}^{k})$ from the original sample.
2. Generate B bootstrap sample $\{(Y_i^{zb},D_i^{zb})\}_{i=1}^{n_z}$ by drawing n_z observations with replacements from the original data $\{Y_i^{z},D_i^{z}\}_{i=1}^{n_z}$ for $z\in\{0,1\}$.
3. Calculate $\widehat{\phi}_{\theta}^{b,\prime}$ for each bootstrap iteration.
4. Take
$$cv_{1-\alpha} = \inf(z : \{\phi_{\widehat{\theta}}'(\sqrt{n}(\widehat{\theta}^b - \widehat{\theta}^b),\pi_{\mathrm{DF}}^{k}) - z/\sigma(\pi_{\mathrm{DF}})_i \leq 0\}).$$
5. Take as lower confidence set $\phi(\hat{\theta}) + \phi_\theta'/\sqrt{n}$.

Theorem A.5. *For any finite set $\{\pi_{\mathrm{DF}}^{k}\}_{i=1}^{I}$, it holds that*

$$\lim_{n\to\infty}\mathbb{P}(\forall k\in\{1,\ldots,K\} : \phi(\widehat{\theta},\pi_{\mathrm{DF}}^{k}) + cv_{1-\alpha}(\pi_{\mathrm{DF}}^{k}) \geq \phi(\theta,\pi_{\mathrm{DF}}^{k})) \geq 1-\alpha,$$

where the inequality is meant to hold for each row.

It is important to note that the number of points under consideration K is some large number, which is fix as the sample size increases. To obtain a confidence band of $\phi_{TE,\kappa}$, which is valid uniformly in π_{DF} , I exploit the functional form of ϕ.I construct the mapping in the following way. To simplify the analysis, I consider the mappings individually. First note that only $BP(\theta^b, \pi_{\mathrm{DF}})$, $\underline{\delta_b}(\theta^b, \pi_{\mathrm{DF}})$ and $\overline{\delta_b}(\theta^b, \pi_{\mathrm{DF}})$ depend on π^k_{DF}. As $\underline{\delta_b}(\theta^b, \pi_{\mathrm{DF}})$ is a simple function of $\underline{\pi}_{DF}$ in π_{DF}, I can construct a valid confidence set by simple projections. It is easy to see that $\overline{\delta_b}(\theta^b, \pi_{\mathrm{DF}})$ is strictly decreasing in π_{DF} and a valid confidence set can thus be easily constructed by exploring monotonicity. To consider $BP(\theta^b, \pi_{\mathrm{DF}})$, I note that $BP(\theta^b, \pi_{\mathrm{DF}})$ is a quasi-convex function of π_{DF} and constructing a lower confidence set for this function is therefore possible as well. By constructing $\tilde{cv}_{1-\alpha}(\pi_{\mathrm{DF}})$ in such a manner, e.g. $\tilde{cv}_{1-\alpha}(\pi^k_{\mathrm{DF}}) = cv_{1-\alpha}(\pi^k_{\mathrm{DF}})$ for all $k \leq K$ and otherwise according to the description above. It therefore holds that

$$\lim_{n\to\infty} \inf_{\pi_{\mathrm{DF}}\in[0,0.5)} \mathbb{P}(\phi(\widehat{\theta}, \pi_{\mathrm{DF}}) + cv_{1-\alpha}(\pi_{\mathrm{DF}}) \geq \phi(\theta, \pi_{\mathrm{DF}})) \leq 1 - \alpha.$$

and more importantly

$$\lim_{n\to\infty} \mathbb{P}(\widehat{RR}^b_{\kappa,B} \subseteq RR^b,\ \mathrm{SR}^b \subseteq \widehat{\mathrm{SR}}^b_{\kappa,B}) \geq 1 - \alpha.$$

Following the arguments of Andrews and Soares (2010); Fang and Santos (2018); Hong and Li (2018) I want to emphasize that these results only hold pointwise in the underlying data generating process as our directional derivative is non sub-additive[29].

B. PROOFS

B.1. **Proof of Theorem 4.1.** I proceed in two steps to prove Theorem 4.1. In Part 1, I present the set of necessary and sufficient conditions, which have to be satisfied by the compliers' outcome distribution. These conditions are based on Assumption 1, the sensitivity parameter constraints and the observation that the outcome distributions of all groups are distribution functions. Part 2 shows that the proposed bounds of the compliers' outcome distribution satisfy all these constraints and additionally that for all $y \in \mathbb{Y}$ at least one of the inequality constraints holds with equality. This implies that any function taking values outside the bounds cannot be the compliers' outcome distribution. Thus, $\underline{F}_{Y^{CO}_d}$ and $\overline{F}_{Y^{CO}_d}$ bound the compliers' outcome distribution sharply as described in the Theorem.

I give first the expression of the outcome distributions of all groups given the compliers

[29]Sub-additive means that $\phi(\theta_1 + \theta_2) \leq \phi(\theta_1) + \phi(\theta_2)$.

outcome distribution. Let $d \in \{0,1\}$

$$F_{Y_d^{DF}}(y) = \frac{1}{\pi_{\mathrm{DF}}}\Big(\pi_{\mathrm{CO}} F_{Y_d^{CO}}(y) - \pi_\Delta G_d(y)\Big)$$
$$F_{Y_d^{d}}(y) = \frac{Q_{dd}(y) - \pi_{\mathrm{CO}} F_{Y_d^{CO}}(y)}{\pi_d}$$

I denote by $\underline{y}$ and $\overline{y}$ the respectively left and right limits of $\mathbb{Y}$, which might equal $\pm\infty$.

Part 1: The compliers' outcome distribution trivially satisfies

$$(i)\ F_{Y_d^{CO}}(y) \in [0,1].$$

As the same reasoning applies to the defiers' outcome distribution and as $F_{Y_d^{CO}}(y) = \frac{\pi_\Delta}{\pi_{\mathrm{CO}}}(G_d(y) + \frac{\pi_{\mathrm{DF}}}{\pi_\Delta} F_{Y_d^{DF}}(y))$, it follows that

$$(ii)\ F_{Y_d^{CO}}(y) \in \left[\frac{\pi_\Delta}{\pi_{\mathrm{CO}}} G_d(y), \frac{\pi_\Delta}{\pi_{\mathrm{CO}}}(G_d(y) + \frac{\pi_{\mathrm{DF}}}{\pi_\Delta})\right].$$

Consider the always and never takers. By the same reasoning as above and as $Q_{dd}(y) = \pi_{\mathrm{CO}} F_{Y_d^{CO}}(y) + \pi_d F_{Y_d^{d}}(y)$, it follows that

$$(iii)\ F_{Y_d^{CO}}(y) \in \left[\frac{Q_{dd}(y) - \pi_d}{\pi_{\mathrm{CO}}}, \frac{Q_{dd}(y)}{\pi_{\mathrm{CO}}}\right].$$

The sensitivity parameter, which identifies the outcome heterogeneity, implies that
$\delta \geq |F_{Y_d^{CO}}(y) - F_{Y_d^{DF}}(y)| = |F_{Y_d^{CO}}(y) - \frac{\pi_\Delta}{\pi_{\mathrm{DF}}}(G_d(y) + \frac{\pi_{\mathrm{CO}}}{\pi_\Delta} F_{Y_d^{CO}}(y))| \Rightarrow |F_{Y_d^{CO}}(y) - G_d(y)| \leq \frac{\pi_{\mathrm{DF}}}{\pi_\Delta}\delta.$
This reasoning implies the fourth constraint

$$(iv)\ F_{Y_d^{CO}}(y) \in [G_d(y) - \frac{\pi_{\mathrm{DF}}}{\pi_\Delta}\delta, G_d(y) + \frac{\pi_{\mathrm{DF}}}{\pi_\Delta}\delta].$$

Let $y, y' \in \mathbb{Y}\ \forall y \geq y'$. As distribution functions are nondecreasing, it has to hold that

$$(v)\ F_{Y_d^{CO}}(y) \geq F_{Y_d^{CO}}(y').$$

The same reasoning applies to the group of defiers and therefore

$$F_{Y_d^{CO}}(y) = \frac{\pi_\Delta}{\pi_{\mathrm{CO}}}(G_d(y) + \frac{\pi_{\mathrm{DF}}}{\pi_\Delta} F_{Y_d^{DF}}(y)) \geq \frac{\pi_\Delta}{\pi_{\mathrm{CO}}}(G_d(y) + \frac{\pi_{\mathrm{DF}}}{\pi_\Delta} F_{Y_d^{DF}}(y'))$$
$$= \frac{\pi_\Delta}{\pi_{\mathrm{CO}}}(G_d(y) - (G_d(y') - \frac{\pi_{\mathrm{CO}}}{\pi_\Delta} F_{Y_d^{CO}}(y'))).$$

This reasoning implies that

$$(vi^\star)\ F_{Y_d^{CO}}(y) - F_{Y_d^{CO}}(y') \geq \frac{\pi_\Delta}{\pi_{\mathrm{CO}}}\left(G_d(y) - G_d(y')\right).$$

$G_d(y)$ is not necessarily nondecreasing. The inequality constraints of conditions (v) and $(vi^\star)$ jointly imply that $F_{Y_d^{CO}}(y) - F_{Y_d^{CO}}(y') \geq \max\{0, \frac{\pi_\Delta}{\pi_{\text{CO}}}(G_d(y) - G_d(y'))\}$ and therefore

$$(vi)\ F_{Y_d^{CO}}(y) - F_{Y_d^{CO}}(y') \geq \frac{\pi_\Delta}{\pi_{\text{CO}}}\left(G_d^+(y) - G_d^+(y')\right).$$

As the always and never takers outcome distribution are nondecreasing, it follows that

$$\begin{aligned}
F_{Y_d^{CO}}(y) &= \frac{1}{\pi_{\text{CO}}}(Q_{dd}(y) - \pi_d F_{Y_d^d}(y)) \leq \frac{1}{\pi_{\text{CO}}}(Q_{dd}(y) - \pi_d F_{Y_d^d}(y')) \\
&= \frac{1}{\pi_{\text{CO}}}(Q_{dd}(y) - \frac{\pi_d}{\pi_d}(Q_{dd}(y') - \pi_{\text{CO}} F_{Y_d^{CO}}(y'))) = \frac{1}{\pi_{\text{CO}}}(Q_{dd}(y) - Q_{dd}(y')) - F_{Y_d^{CO}}(y'),
\end{aligned}$$

and therefore

$$(vii)\ F_{Y_d^{CO}}(y) - F_{Y_d^{CO}}(y') \leq \frac{Q_{dd}(y) - Q_{dd}(y')}{\pi_{\text{CO}}}.$$

Constraints (vi) and (vii) obviously do not contradict each other. The compliers distribution functions have to further satisfy the limit conditions

$$(viii)\ \lim_{y\to\underline{y}} F_{Y_d^{CO}}(y) = 0 \text{ and } \lim_{y\to\overline{y}} F_{Y_d^{CO}}(y) = 1.$$

I further note that it holds trivially that $\lim_{y\to\underline{y}} G_d(y) = \lim_{y\to\underline{y}} Q_{ds}(y) = 0$ and $\lim_{y\to\overline{y}} G_d(y) = 1$ and $\lim_{y\to\overline{y}} Q_{dd}(y) = \pi_d + \pi_{\text{CO}}$ and $\lim_{y\to\overline{y}} Q_{d(1-d)}(y) = \pi_d + \pi_{\text{DF}}$ for any $d, s \in \{0, 1\}$. The other groups fulfill these limit conditions as well if condition $(viii)$ holds as

$$\begin{aligned}
\lim_{y\to\underline{y}} F_{Y_d^{DF}}(y) &= \lim_{y\to\underline{y}} \frac{1}{\pi_{\text{DF}}}\left(\pi_{\text{CO}} F_{Y_d^{CO}}(y) - \pi_\Delta G_d(y)\right) = 0 \\
\lim_{y\to\overline{y}} F_{Y_d^{DF}}(y) &= \lim_{y\to\overline{y}} \frac{1}{\pi_{\text{DF}}}\left(\pi_{\text{CO}} F_{Y_d^{CO}}(y) - \pi_\Delta G_d(y)\right) = \frac{1}{\pi_{\text{DF}}}(\pi_{\text{CO}} - \pi_\Delta) = 1 \\
\lim_{y\to\underline{y}} F_{Y_d^d}(y) &= \lim_{y\to\underline{y}} \frac{Q_{dd}(y) - \pi_{\text{CO}} F_{Y_d^{CO}}(y)}{\pi_d} = 0 \\
\lim_{y\to\overline{y}} F_{Y_d^d}(y) &= \lim_{y\to\overline{y}} \frac{Q_{dd}(y) - \pi_{\text{CO}} F_{Y_d^{CO}}(y)}{\pi_d} = \frac{\pi_{\text{CO}} + \pi_d - \pi_{\text{CO}}}{\pi_d} = 1.
\end{aligned}$$

Any real-valued function, which is defined on $\mathbb{Y}$, and right-continuous, which left-limits exists and which satisfies conditions (i) - $(viii)$, induces by construction potential outcome distributions for all four groups, which are consistent with the imposed assumption, the sensitivity parameter constraints, and the observed probability functions. It can thus represent the compliers' outcome distribution. It is moreover clear that the simple structure of constraints $(i) - (viii)$ implies that if there are two such functions any convex combinations of these functions satisfy these constraints as well.

Part 2: I show in the following both that the proposed bounds satisfy constraints

$(i)-(viii)$ and that any function which takes values outside these bounds contradict one of the constraints and can therefore not be the compliers' outcome distribution. As the sensitivity parameters under consideration lie within the sensitivity region by assumption, bounds on the compliers' outcome distributions exist and are therefore non-intersecting by construction. The lower bound therefore has to satisfy the left and the upper bound the right constraints of $(i)-(iv)$. $\underline{H}(y, \pi_{\mathrm{DF}}, \delta)$ and $\overline{H}(y, \pi_{\mathrm{DF}}, \delta)$ take the maximum and respectively the minimum of the constraints $(i)-(iv)$ and imposes constraint (v) by construction. They therefore satisfy these constraints by construction, where at least one of these constraints is binding. Consider

$$\underline{H}_1(y) = \pi_\Delta G_d^+(y) - \inf_{\tilde{y} \leq y} \left(\pi_\Delta G_d^+(\tilde{y}) - \pi_{\mathrm{CO}} \underline{H}(\tilde{y}, \pi_{\mathrm{DF}}, \delta)\right). \tag{15}$$

It clearly holds that $\underline{H}_1(y) \geq \pi_{\mathrm{CO}} \underline{H}(y, \pi_{\mathrm{DF}}, \delta)$ and $\underline{H}_1(y)$ therefore satisfies constraints $(i)-(v)$ as well. If $\underline{H}_1(y) = \pi_{\mathrm{CO}} \underline{H}(y, \pi_{\mathrm{DF}}, \delta)$, one of the constraints $(i)-(v)$ is binding. Consider again any $y, y' \in \mathbb{Y}$ such that $y' \leq y$. Based on this reasoning $\underline{H}_1(y)$ satisfies constraint (vi) as

$$\begin{aligned}
&\underline{H}_1(y) - \underline{H}_1(y') \\
&= \pi_\Delta G_d^+(y) - \pi_\Delta G_d^+(y') - \left(\inf_{\tilde{y} \leq y} \left(\pi_\Delta G_d^+(\tilde{y}) - \underline{H}(\tilde{y}, \pi_{\mathrm{DF}}, \delta)\right) - \inf_{\tilde{y} \leq y'} \left(\pi_\Delta G_d^+(\tilde{y}') - \underline{H}(\tilde{y}', \pi_{\mathrm{DF}}, \delta)\right)\right) \\
&\geq \pi_\Delta G_d^+(y) - \pi_\Delta G_d^+(y').
\end{aligned}$$

If $\underline{H}_1(y) \neq \underline{H}(y, \pi_{\mathrm{DF}}, \delta)$, the inequality holds with equality for some $y' < y$ such that constraint (vi) is binding. As this reasoning applies to all $y' \leq y$, it follows that $\underline{H}_1(y)$ is such that one of the constraints $(i)-(vi)$ is binding for all $y \in \mathbb{Y}$. Now consider the function $\underline{H}_2(y) = \pi_{\mathrm{CO}} \overline{F}_{Y_d^{CO}}(y, \pi_{\mathrm{DF}}, \delta)$, which is given by

$$\underline{H}_2(y) = Q_{dd}(y) - \inf_{\tilde{y} \geq y} \left(Q_{dd}(\tilde{y}) - \underline{H}_1(\tilde{y})\right).$$

It is clear that $\underline{H}_2(y) \geq \underline{H}_1(y)$ and that $Q_{dd}(y) - Q_{dd}(y') \geq \underline{H}_2(y) - \underline{H}_2(y')$. I further show that $\underline{H}_2(y)$ satisfies constraint (vi). Note that for any $y, y' \in \mathbb{Y}$ there are three possibilities: Either $\underline{H}_2(y) - \underline{H}_2(y') = \underline{H}_1(y) - \underline{H}_1(y')$, or $\underline{H}_2(y) - \underline{H}_2(y') = Q_{dd}(y) - Q_{dd}(y')$ or a combination of both and as

$$Q_{dd}(y) - Q_{dd}(y') \geq G_d^+(y) - G_d^+(y')$$

constraint (vi) is satisfied and one of the constraints is always binding.

The lower bound is sharp if it satisfies condition $(viii)$. The same reasoning applies to the upper bound. To briefly sketch this reasoning, let

$$\overline{H}_1(y) = Q_{dd}(y) - \sup_{\hat{y} \leq y} \left(Q_{dd}(\hat{y}) - \pi_{\mathrm{CO}} \overline{H}(\hat{y}, \pi_{\mathrm{DF}}, \delta)\right).$$

It follows similarly that constrain (vii) as well as constraints $(i)-(v)$,$\overline{H}_1(y) \geq \pi_{\text{CO}}\overline{H}(y, \pi_{\text{DF}}, \delta)$, are fulfilled such that at least one of them is binding. Consider now

$$\overline{H}_2(y) = \pi_\Delta G_d^+(y) - \sup_{\tilde{y} \geq y} \left(\pi_\Delta G_d^+(\tilde{y}) - \overline{H}_1(y)\right).$$

It is clear that $\overline{H}_2(y)$ fulfills constraints $(i)-(vii)$ and the upper bound is therefore sharp, if constraint $(viii)$ is satisfied.

It is therefore left to show that the bounds satisfy condition $(viii)$. Consider first the lower bound. By construction, $\underline{F}_{Y_d^{CO}}(y, \pi_{\text{DF}}, \delta) \in [0, 1]$. I therefore show that $\lim_{y \to \underline{y}} \underline{F}_{Y_d^{CO}}(y, \pi_{\text{DF}}, \delta) \leq 0$ and $\lim_{y \to \overline{y}} \underline{F}_{Y_d^{CO}}(y, \pi_{\text{DF}}, \delta) \geq 1$. It follows that

$$\lim_{y \to \underline{y}} \pi_{\text{CO}}\ \underline{F}_{Y_d^{CO}}(y, \pi_{\text{DF}}, \delta) = \sup_{\tilde{y} \in \mathbb{R}} \left(\pi_\Delta G_d^+(\tilde{y}) - Q_{dd}(\tilde{y}) - \inf_{\hat{y} \leq \tilde{y}} \left(\pi_\Delta G_d^+(\hat{y}) + \pi_{\text{CO}}\underline{H}(\hat{y}, \pi_{\text{DF}}, \delta)\right)\right)$$

The equality follows as $\lim_{y \to \underline{y}} Q_{dd}(y) = 0$. I note that for all $y, y' \in \mathbb{Y}$ and $y' \leq y$

$$Q_{dd}(y) - Q_{dd}(y') \geq \pi_\Delta \left(G_d^+(y) - G_d^+(y')\right).$$

and therefore $\lim_{y \to \underline{y}} \pi_{\text{CO}}\underline{F}_{Y_d^{CO}}(y, \pi_{\text{DF}}, \delta)$ is bounded from above by

$$\sup_{\hat{y} \in \mathbb{Y}} \left(\max\{\underbrace{0}_{(1)}, \underbrace{Q_{dd}(\hat{y}) - \pi_d}_{(2)}, \underbrace{\pi_\Delta G_d^{\sup}(\hat{y})}_{(3)}, \underbrace{\pi_{\text{CO}} G_d^{\sup}(\hat{y}) - \pi_{\text{CO}} \frac{\pi_{\text{DF}}}{\pi_\Delta}\delta}_{(4)}\} - Q_{dd}(\hat{y}) \right) \leq 0$$

I show that the last inequality holds for all $\hat{y} \in \mathbb{Y}$. I therefore show that each of the four elements in the expression of max is smaller than $Q_{dd}(\hat{y})$. It is obvious that expressions (1) and (2) hold. (3) follows as $\pi_\Delta G_d^{\sup}(y) - Q_{dd}(\hat{y}) \leq Q_{dd}(y) - Q_{dd}(\hat{y}) \leq 0$. (4) follows as for $y^\star \leq \hat{y}$ satisfying $G_d^{\sup}(\hat{y}) \leq F_{Y_d^{CO}}(y^\star) + \frac{\pi_{\text{DF}}}{\pi_\Delta}\delta$

$$\begin{aligned}
(4)\ G_d^{\sup}(\hat{y}) - \frac{\pi_{\text{DF}}}{\pi_\Delta}\delta - \frac{Q_{dd}(\hat{y})}{\pi_{\text{CO}}} &\leq F_{Y_d^{CO}}(y^\star) + \frac{\pi_{\text{DF}}}{\pi_\Delta}\delta - \frac{\pi_{\text{DF}}}{\pi_\Delta}\delta - \frac{Q_{dd}(\hat{y})}{\pi_{\text{CO}}} \\
\leq F_{Y_d^{CO}}(y^\star) + \frac{\pi_{\text{DF}}}{\pi_\Delta}\delta - \frac{\pi_{\text{DF}}}{\pi_\Delta}\delta - \frac{\pi_{\text{CO}} F_{Y_d^{CO}}(\hat{y}) + \pi_d F_{Y_d^d}(\hat{y})}{\pi_{\text{CO}}} &\leq -\frac{\pi_d F_{Y_d^d}(\hat{y})}{\pi_{\text{CO}}} \leq 0
\end{aligned}$$

The first inequality follows from the reasoning above and the second by definition of $Q_{dd}(\hat{y})$. The third inequality follows as $F_{Y_d^{CO}}(\hat{y}) \geq F_{Y_d^{CO}}(y^\star)$ as $\hat{y} \geq y^\star$. The last equality holds, as $F_{Y_d^d}$ satisfies constraint (i).

Consider now the right limit. It holds from the reasoning above that $\underline{F}_{Y_d^{CO}}(y, \pi_{\text{DF}}, \delta) \geq$

$\underline{H}(y, \pi_{\mathrm{DF}}, \delta)$ and therefore

$$\lim_{y\to\overline{y}} \underline{F}_{Y_d^{CO}}(y, \pi_{\mathrm{DF}}, \delta) \geq \lim_{y\to\overline{y}} \max\{0, G_d^{\sup}(y) - \frac{\pi_{\mathrm{DF}}}{\pi_\Delta}\delta, \frac{\pi_\Delta}{\pi_{\mathrm{CO}}} G_d^{\sup}(y), \frac{Q_{dd}(y) - \pi_d}{\pi_{\mathrm{CO}}}\} \geq 1.$$

The second inequality follows as $\lim_{y\to\overline{y}} Q_{dd}(y) - \pi_d = \pi_{\mathrm{CO}} + \pi_d - \pi_d = \pi_{\mathrm{CO}}$.

I now consider the upper bound. It holds that

$$\lim_{y\to\underline{y}} \overline{F}_{Y_d^{CO}}(y, \pi_{\mathrm{DF}}, \delta) \leq \lim_{y\to\underline{y}} \min\{1, G_d^{\inf}(y) + \frac{\pi_{\mathrm{DF}}}{\pi_\Delta}\delta, \frac{\pi_{\mathrm{CO}}}{\pi_\Delta} G_d^{\inf}(y) + \frac{\pi_{\mathrm{CO}}}{\pi_{\mathrm{DF}}}, \frac{Q_{dd}(y)}{\pi_{\mathrm{CO}}}\} \leq 0,$$

where the second inequality follows by $\lim_{y\to\underline{y}} \frac{Q_{dd}(y)}{\pi_{\mathrm{CO}}} = 0$. Consider now $\lim_{y\to\overline{y}} \overline{F}_{Y_d^{CO}}(y, \pi_{\mathrm{DF}}, \delta)$, which equals

$$\frac{\pi_{\mathrm{CO}} + \pi_d}{\pi_{\mathrm{CO}}} - \sup_{\hat{y}\in\mathbb{Y}} \left(\frac{Q_{dd}(\hat{y})}{\pi_{\mathrm{CO}}} - \min\{\underbrace{1}_{(1)}, \underbrace{\frac{Q_{dd}(\hat{y})}{\pi_{\mathrm{CO}}}}_{(2)}, \underbrace{\frac{\pi_\Delta}{\pi_{\mathrm{CO}}} G_d^{\inf}(\hat{y}) + \frac{\pi_{\mathrm{DF}}}{\pi_{\mathrm{CO}}}}_{(3)}, \underbrace{G_d^{\inf}(\hat{y}) + \frac{\pi_{\mathrm{DF}}}{\pi_\Delta}\delta}_{(4)}\} \right).$$

I show that the expression in $\sup_{y\in\mathbb{Y}}(\dots)$ is bounded from above by $\frac{\pi_d}{\pi_{\mathrm{CO}}}$ such that

$$\lim_{y\to\overline{y}} \overline{F}_{Y_d^{CO}}(y, \pi_{\mathrm{DF}}, \delta) \geq \frac{\pi_{\mathrm{CO}} + \pi_d}{\pi_{\mathrm{CO}}} - \frac{\pi_d}{\pi_{\mathrm{CO}}} = 1.$$

It holds clearly that for all $y \in \mathbb{Y}$

$$(1)\ \frac{Q_{dd}(y)}{\pi_{\mathrm{CO}}} - 1 \leq \frac{\pi_{\mathrm{CO}} + \pi_d}{\pi_{\mathrm{CO}}} - 1 = \frac{\pi_d}{\pi_{\mathrm{CO}}} \qquad (2)\ \frac{Q_{dd}(y)}{\pi_{\mathrm{CO}}} - \frac{Q_{dd}(\hat{y})}{\pi_{\mathrm{CO}}} = 0 \leq \frac{\pi_d}{\pi_{\mathrm{CO}}}$$

There exists $y^\star \geq \hat{y}$ such that $G_d^{\inf}(\hat{y}) \geq G_d(y^\star)$. It holds that (iii) multiplied by π_{CO} implies

$$\begin{aligned}(3)\ & Q_{dd}(\hat{y}) - \pi_\Delta G_d^{\inf}(\hat{y}) - \pi_{\mathrm{DF}} \leq Q_{dd}(\hat{y}) - \pi_\Delta G_d(y^\star) - \pi_{\mathrm{DF}} \\ &= Q_{dd}(\hat{y}) - Q_{dd}(y^\star) + Q_{d(1-d)}(y^\star) - \pi_{\mathrm{DF}} \leq Q_{d(1-d)}(y^\star) - \pi_{\mathrm{DF}} \leq \pi_d + \pi_{\mathrm{DF}} - \pi_{\mathrm{DF}} = \pi_d.\end{aligned}$$

The first equality follows from the definition of $G_d(y)$. The second inequality follows as $Q_{\mathrm{dd}}(\hat{y}) \leq Q_{dd}(y^\star)$. The last inequality follows as $Q_{d(1-d)}(y^\star) \leq \pi_d + \pi_{\mathrm{DF}}$. It further holds that

$$\begin{aligned}(4)\ & \frac{Q_{dd}(\hat{y})}{\pi_{\mathrm{CO}}} - G_d^{\inf}(\hat{y}) - \frac{\pi_{\mathrm{DF}}}{\pi_\Delta}\delta = \frac{Q_{dd}(\hat{y})}{\pi_{\mathrm{CO}}} - G_d(y^\star) - \frac{\pi_{\mathrm{DF}}}{\pi_\Delta}\delta \\ &\leq F_{Y_d^{CO}}(\hat{y}) + \frac{\pi_d F_{Y_d^d}(\hat{y})}{\pi_{\mathrm{CO}}} - F_{Y_d^{CO}}(y^\star) - \frac{\pi_{\mathrm{DF}}}{\pi_\Delta}\delta + \frac{\pi_{\mathrm{DF}}}{\pi_\Delta}\delta \leq \frac{\pi_d F_{Y_d^d}(\hat{y})}{\pi_{\mathrm{CO}}} \leq \frac{\pi_d}{\pi_{\mathrm{CO}}}.\end{aligned}$$

The first inequality follows from the definition of $Q_{dd}(y)$ and as $-G_d(y) \leq -F_{Y_d^{CO}}(y) + \frac{\pi_{\mathrm{DF}}}{\pi_\Delta}\delta$.

The second inequality follows as $F_{Y_d^{CO}}(y) \leq F_{Y_d^{CO}}(y^\star)$.

This reasoning completes the proof that the proposed bounds induce proper distribution functions of the always takers, never takers, defiers, and compliers, which are consistent with Assumption 1, the parameter constraints and by construction with the observed probabilities. This is the case as both bounds of the compliers' outcome distribution satisfy constraints $(i)-(viii)$. Additionally, as both bounds are smooth mappings from distribution functions such that they preserve the existence of limits and continuity, the bounds are right-continuous and have left-limits. Therefore, the bounds are attainable as compliers' outcome distributions and are sharp as described in the Theorem. □

B.2. **Proof of Theorem 4.2.** As a reminder, I show that the compliers' population size is sharply bounded by

$$\begin{aligned}
\overline{\pi}_{CO} =& \min\{\mathbb{P}(D=1|Z=1), \mathbb{P}(D=0|Z=0)\} \\
\underline{\pi}_{CO} =& \max_{d\in\{0,1\}}\{\sup_{B\in\mathscr{B}}\{\mathbb{P}(Y\in B, D=d|Z=d) - \mathbb{P}(Y\in B, D=d|Z=1-d)\}\},
\end{aligned}$$

where $\mathbb{P}(Y\in B, D=d|Z=z)$ with $B\in\mathscr{B}$ and $d,z\in\{0,1\}$ denotes the conditional joint probability distribution of observable variables. The Lemma follows from this statement as $\pi_{DF} = \pi_{CO} - \mathbb{P}(D=1|Z=1) + \mathbb{P}(D=1|Z=0)$. Let $\mathbb{P}(Y_d^t \in B)$ denotes the unobserved probability distribution of the potential outcome variable of group t with treatment status d.[30]

I first show that $\underline{\pi}_{CO}$ is a valid lower bound of the compliers' population size. It follows from the definition of types and as probabilities are nonnegative that

$$\overline{\pi}_{CO} = \min\{\pi_{AT} + \pi_{CO}, \pi_{NT} + \pi_{CO}\} \geq \pi_{CO}.$$

Similarly, $\overline{\pi}_{CO}$ bounds the compliers' population size from above as $\underline{\pi}_{CO}$ equals

$$\begin{aligned}
\max\Big\{&\sup_{B\in\mathscr{B}}\{\mathbb{P}(Y_1\in B, T=AT) + \mathbb{P}(Y_1\in B, T=CO) \\
&\qquad -\mathbb{P}(Y_1\in B, T=AT) - \mathbb{P}(Y_1\in B, T=DF)\}, \\
&\sup_{B\in\mathscr{B}}\{\mathbb{P}(Y_0\in B, T=NT) + \mathbb{P}(Y_0\in B, T=CO) \\
&\qquad -\mathbb{P}(Y_0\in B, T=NT) - \mathbb{P}(Y_0\in B, T=DF)\}\Big\} \\
&\leq \max\{\sup_{B\in\mathscr{B}}\{\mathbb{P}(Y_1\in B, T=CO)\}, \sup_{B\in\mathscr{B}}\{\mathbb{P}(Y_0\in B, T=CO)\}\} = \pi_{CO}.
\end{aligned}$$

[30]In principle, Theorem 4.2 is a Corollary of Theorem 4.1. Considering the sharp lower bound on the population size of defiers, one could simply solve these bounds for the maximal minimal size of defiers for which there exists one value of outcome heterogeneity δ_q such that the bounds are non-intersecting. However, as the bounds are rather complex this exercise is tedious. I therefore propose a simpler and direct proof for this claim in this section.

The second equality follows from the independence assumption and the definition of the types. The last inequality follows as probabilities are nonnegative. For any $\tilde{\pi}_{CO} \in [\underline{\pi}_{CO}, \overline{\pi}_{CO}]$, I show that one can construct probability distributions of eight potential outcome variables $\tilde{Y}_d^T$ for $d \in \{0,1\}$ and $T \in \{CO, DF, AT, NT\}$, which imply $\tilde{\pi}_{CO}$ as the compliers population size.[31]

Let $B \in \mathscr{B}$ and $\mathscr{B}_B = \{A \cap B | A \in \mathscr{B}\}$. The observed outcome probabilities are uninformative about the outcome probabilities of the always takers in the absence and of the never takers in the presence of treatment. Thus, they are left unspecified in this analysis. In the following the six potential outcome probability distributions are defined for $\tilde{\pi}_{CO} \in [\underline{\pi}_{CO}, \overline{\pi}_{CO}]$[32]

$$\mathbb{P}(\tilde{Y}_d \in B, T = CO) = \mathbb{P}(Y \in B, D = d | Z = d) - \mathbb{P}(\tilde{Y}_d \in B, T = dT)$$
$$\mathbb{P}(\tilde{Y}_d \in B, T = DF) = \mathbb{P}(Y \in B, D = d | Z = 1-d) - \mathbb{P}(\tilde{Y}_d \in B, T = dT)$$
$$\mathbb{P}(\tilde{Y}_d \in B, T = dT)$$
$$= \frac{\mathbb{P}(D = d|Z = d) - \tilde{\pi}_{CO}}{\mathbb{P}(D = d|Z = d) - \sup_{C \in \mathscr{B}}(\mathbb{P}(Y \in C, D = d|Z = d) - \mathbb{P}(Y \in C, D = d|Z = 1-d))} \times$$
$$\left(\mathbb{P}(Y \in B, D = d|Z = d) - \sup_{C \in \mathscr{B}_B}(\mathbb{P}(Y \in C, D = d|Z = d) - \mathbb{P}(Y \in C, D = d|Z = 1-d))\right).$$

The specified outcome distribution of group dT consist of two factors. The first factor weights the probability distribution such that it integrates to the corresponding population size. The second factor takes the supremum of the outcome probability for each Borel set B such that the implied outcome probabilities of the compliers and defiers are positive. As the outcome distributions of the groups are defined independently of the random variable $\tilde{Z}$, they satisfy by construction the independence assumption and the relevance assumption is satisfied by the imposed population sizes. The proposed outcome probability distributions imply the observed outcome probability distributions: $\forall B \in \mathscr{B}$ and $\forall d, z \in \{0,1\}$ $\mathbb{P}(Y \in B, D = d|Z = z) = \mathbb{P}(\tilde{Y} \in B, \tilde{D} = d|\tilde{Z} = z)$. It thus has to be shown that the implied probability distributions are proper outcome distribution, which satisfy $\forall T \in \{CO, DF, AT, NT\}$, $d \in \{0,1\}$, and $B, B' \in \mathscr{B}$, where $B' \subseteq B$.

1.) $\mathbb{P}(\tilde{Y}_d^T \in B) \geq 0$,

2.) $\mathbb{P}(\tilde{Y}_d^T \in \mathbb{Y}) = 1$ and

[31] It suffices to specify the marginal potential outcome distribution of the groups rather than the full distribution of (Y_0, Y_1, T, Z). The reasoning is that by the independence assumption $(Y_0, Y_1, T) \perp Z$ and thus the probability distribution function of Z is left unspecified. Moreover, as the joint distribution of (Y_0^T, Y_1^T) is not restricted by observable - beyond their marginals - it is left unspecified as well conditional on population sizes.

[32] Otherwise $\mathbb{P}(\tilde{Y}_d \in B, T = dT)$ is defined to be zero if $\mathbb{P}(D = d|Z = d) = \sup_{C \in \mathscr{B}}(\mathbb{P}(Y \in C, D = d|Z = d) - \mathbb{P}(Y \in C, D = d|Z = 1-d))$. The other probability distributions stay the same.

3.) $\mathbb{P}(\tilde{Y}_d^T \in B) \geq \mathbb{P}(\tilde{Y}_d^T \in B')$

I first consider 1.) such that for any $B \in \mathcal{B}$

$$\mathbb{P}(\tilde{Y}_d \in B, T = dT) \geq 0.$$

First note that it follows from $\underline{\pi}_{CO}$ that the fraction is smaller than one and from $\overline{\pi}_{CO}$ that it is positive. Second note that

$$\begin{aligned}
\mathbb{P}(Y \in B, D = d|Z = d) - \sup_{C \in \mathscr{B}_B} \left(\mathbb{P}(Y \in C, D = d|Z = d) - \mathbb{P}(Y \in C, D = d|Z = 1-d)\right) \\
\geq \mathbb{P}(Y \in B, D = d|Z = d) - \sup_{C \in \mathscr{B}_B} \left(\mathbb{P}(Y \in C, D = d|Z = d)\right) = 0.
\end{aligned}$$

The inequality follows. Moreover, note that

$$\begin{aligned}
&\mathbb{P}(\tilde{Y}_d \in B, T = DF) = \mathbb{P}(Y \in B, D = d|Z = 1-d) - \mathbb{P}(\tilde{Y}_d \in B, T = dT) \\
&= \mathbb{P}(Y \in B, D = d|Z = 1-d) \\
&\quad - \frac{\mathbb{P}(D = d|Z = d) - \tilde{\pi}_{CO}}{\mathbb{P}(D = d|Z = d) - \sup_{C \in \mathscr{B}}(\mathbb{P}(Y \in C, D = d|Z = d) - \mathbb{P}(Y \in C, D = d|Z = 1-d))} \\
&\quad \left(\mathbb{P}(Y \in B, D = d|Z = d) - \sup_{C \in \mathscr{B}_B} \left(\mathbb{P}(Y \in C, D = d|Z = d) - \mathbb{P}(Y \in C, D = d|Z = 1-d)\right)\right),
\end{aligned}$$

which is greater or equal to zero. By the same reasoning as above the fraction is smaller than one and positive. The inequality follows as it further holds that for any $\omega \in (0, 1)$

$$\begin{aligned}
&\mathbb{P}(Y \in B, D = d|Z = 1-d) - \omega\mathbb{P}(Y \in B, D = d|Z = d) \\
&\qquad + \omega \sup_{C \in \mathscr{B}_B} \left(\mathbb{P}(Y \in C, D = d|Z = d) - \mathbb{P}(Y \in C, D = d|Z = 1-d)\right) \\
&\geq \mathbb{P}(Y \in B, D = d|Z = 1-d) - \omega\mathbb{P}(Y \in B, D = d|Z = d) \\
&\qquad + \omega\mathbb{P}(Y \in B, D = d|Z = d) - \omega\mathbb{P}(Y \in B, D = d|Z = 1-d) \geq 0.
\end{aligned}$$

The same reasoning applies to the compliers.

2.) It clearly holds that the outcome probabilities add up to their respectively population sizes as

$$\begin{aligned}
\mathbb{P}(\tilde{Y}_d \in \mathbb{Y}, T = dT) &= \mathbb{P}(D = d|Z = d) - \tilde{\pi}_{CO} = \tilde{\pi}_{dT} \\
\mathbb{P}(\tilde{Y}_d \in \mathbb{Y}, T = CO) &= \mathbb{P}(D = d|Z = d) - \mathbb{P}(D = d|Z = d) + \tilde{\pi}_{CO} = \tilde{\pi}_{CO} \\
\mathbb{P}(\tilde{Y}_d \in \mathbb{Y}, T = DF) &= \mathbb{P}(D = d|Z = 1-d) - \mathbb{P}(D = d|Z = d) + \tilde{\pi}_{CO} = \tilde{\pi}_{DF}.
\end{aligned}$$

I lastly show that the probability distributions are nondecreasing such that 3.) holds. Let

$B' \subseteq B$. I want to show that $\mathbb{P}(\tilde{Y}_d \in B, T = dT) - \mathbb{P}(\tilde{Y}_d \in B', T = dT) \geq 0$, which follows as

$$
\begin{aligned}
&\mathbb{P}(Y \in B, D = d|Z = d) - \sup_{C \in \mathscr{B}_B} (\mathbb{P}(Y \in C, D = d|Z = d) - \mathbb{P}(Y \in C, D = d|Z = 1 - d)) \\
&\qquad - \mathbb{P}(Y \in B', D = d|Z = d) + \sup_{C \in \mathscr{B}'_B} (\mathbb{P}(Y \in C, D = d|Z = d) - \mathbb{P}(Y \in C, D = d|Z = 1 - d)) \\
&= \mathbb{P}(Y \in B, D = d|Z = d) - \mathbb{P}(Y \in B', D = d|Z = d) \\
&\qquad - \sup_{C \in \mathscr{B}_{B \setminus B'}} (\mathbb{P}(Y \in C, D = d|Z = d) - \mathbb{P}(Y \in C, D = d|Z = 1 - d)) \\
&\geq \mathbb{P}(Y \in B, D = d|Z = d) - \mathbb{P}(Y \in B', D = d|Z = d) - \sup_{C \in \mathscr{B}_{B \setminus B'}} (\mathbb{P}(Y \in C, D = d|Z = d)) \geq 0.
\end{aligned}
$$

I further note that $\mathbb{P}(\tilde{Y}_d \in B, T = CO) \geq \mathbb{P}(\tilde{Y}_d \in B', T = CO)$ as

$$
\begin{aligned}
&\mathbb{P}(\tilde{Y}_d \in B, T = dT) - \mathbb{P}(\tilde{Y}_d \in B, T = dT) = \mathbb{P}(Y \in B, D = d|Z = d) - \mathbb{P}(Y \in B', D = d|Z = d) \\
&\qquad - \sup_{C \in \mathscr{B}_{B \setminus B'}} (\mathbb{P}(Y \in C, D = d|Z = d) - \mathbb{P}(Y \in C, D = d|Z = 1 - d)) \\
&\leq \mathbb{P}(Y \in B, D = d|Z = d) - \mathbb{P}(Y \in B', D = d|Z = d).
\end{aligned}
$$

Moreover, $\mathbb{P}(\tilde{Y}_d \in B, T = DF) \geq \mathbb{P}(\tilde{Y}_d \in B', T = DF)$ as

$$
\begin{aligned}
&\mathbb{P}(\tilde{Y}_d \in B, T = dT) - \mathbb{P}(\tilde{Y}_d \in B, T = dT) = \mathbb{P}(Y \in B, D = d|Z = d) - \mathbb{P}(Y \in B', D = d|Z = d) \\
&\qquad - \sup_{C \in \mathscr{B}_{B \setminus B'}} (\mathbb{P}(Y \in C, D = d|Z = d) - \mathbb{P}(Y \in C, D = d|Z = 1 - d)) \\
&\leq \mathbb{P}(Y \in B, D = d|Z = 1 - d) - \mathbb{P}(Y \in B', D = d|Z = 1 - d).
\end{aligned}
$$

This reasoning completes this proof.

□

B.3. **Proof of Proposition 5.1.** In the absence of treatment the data generating process does not reveal anything about the distribution of the always takers, and neither in the presence of treatment of the never takers. The outcome is therefore just bounded by its support. Sharpness of the other six outcome distribution follows from Theorem 4.1. It follows from the definition of the outcome distribution that in the presence and absence of treatment that

$$F_{Y_d}(y) = \pi_{1-d} F_{Y_d^{(1-d)T}} + \pi_{\mathrm{CO}} F_{Y_d^{CO}} + \pi_{\mathrm{DF}} F_{Y_d^{DF}} + \pi_d F_{Y_d^{dT}},$$

where π_{1-d} is the population size of always takers if $d=0$ and otherwise of the never takers and $F_{Y_d^{(1-d)T}}$ respectively. It follows moreover that

$$\begin{aligned}
F_{Y_d}(y) &= \pi_{1-d}F_{Y_d^{(1-d)T}} + \pi_{\mathrm{CO}}F_{Y_d^{CO}} - \pi_{\Delta}G_d(y) + \pi_{\mathrm{CO}}F_{Y_d^{CO}} + Q_{dd}(y) - \pi_{\mathrm{CO}}F_{Y_d^{CO}}. \\
&= \pi_{1-d}F_{Y_d^{(1-d)T}} + \pi_{\mathrm{CO}}F_{Y_d^{CO}} - \pi_{\Delta}G_d(y) + Q_{dd}(y) \\
&= \pi_{1-d}F_{Y_d^{(1-d)T}} + \pi_{\mathrm{CO}}F_{Y_d^{CO}} + Q_{d(1-d)}(y)
\end{aligned}$$

Sharp bounds in a first order stochastic dominacen sense of $F_{Y_d}(y)$ are therefore obtained by taking the compliers sharp lower and upper bounds and setting $Y_d^{(1-d)T}$ to one of its boundaries of its support. Moreover, it is clear that an upper bound on the compliers' distribution (in a first order stochastic dominance sense) implies an bound on the defiers' distribution and a lower bound on the never and respectively always takers outcome distribution. The Proposition 5.1 follows from this reasoning. □

B.4. **Proof of Proposition 5.2.** By the same arguments of the Theorem 4.1, one can show that these bounds are sharp conditionally on x, if x fulfills the conditions. □

B.5. **Proof of Corollary 5.3.** By the reasoning of the main text, the construction of the bounds implies that they are sharp implying that any values of those bounds cannot be proper bounds. It is left to show that they are bounds. As each of those equation only satisfy one side, it is left to show that the left hand side is always smaller than the right one.

I first show that the RHS is always greater than zero. It is obvious for (ii). It also applies to (iii) as

$$\frac{P_{dd} - P_{d(1-d)} + \pi_{\mathrm{DF}}}{\pi_{\mathrm{CO}}} = \frac{\pi_{\mathrm{CO}}P_d^{CO} - \pi_{\mathrm{DF}}P_d^{DF} + \pi_{\mathrm{DF}}}{\pi_{\mathrm{CO}}} \geq \frac{\pi_{\mathrm{CO}}P_d^{CO}}{\pi_{\mathrm{CO}}} \geq 0.$$

By the same reasoning for (iv), it holds that

$$P_{dd} - P_{d(1-d)} + \pi_{\mathrm{DF}}\delta \geq P_{dd} - P_{d(1-d)} + \pi_{\mathrm{DF}}\left(\frac{P_{d(1-d)} - P_{dd}}{\pi_{\mathrm{DF}}}\right) = 0.$$

I show that the RHS is always greater than $(P_{dd} - \pi_{\mathrm{d}})/\pi_{\mathrm{CO}}$. I therefore note that

$$\frac{P_{dd} - \pi_{\mathrm{d}}}{\pi_{\mathrm{CO}}} = \frac{\pi_{\mathrm{CO}}P_d^{CO} + \pi_{\mathrm{d}}P_d^{DT} - \pi_{\mathrm{d}}}{\pi_{\mathrm{CO}}} \leq \frac{\pi_{\mathrm{CO}}P_d^{CO}}{\pi_{\mathrm{CO}}} \leq 1.$$

The third term follows if $P_{d(1-d)} - \pi_{\mathrm{DF}} \leq \pi_{\mathrm{d}}$, which holds as

$$P_{d(1-d)} - \pi_{\mathrm{DF}} = \pi_{\mathrm{DF}}P_d^{DF} + \pi_{\mathrm{d}}P_d^{DT} - \pi_{\mathrm{DF}} \leq \pi_{\mathrm{d}}P_d^{DT} \geq -\pi_{\mathrm{d}}.$$

Condition (iv) is left to show

$$\begin{aligned}
&\frac{P_{dd}-\pi_{\mathrm{d}}}{\pi_{\mathrm{CO}}}-\frac{1}{\pi_{\mathrm{CO}}-\pi_{\mathrm{DF}}}(P_{dd}-P_{d(1-d)}+\pi_{\mathrm{DF}}\delta)\\
\leq\ &\frac{\pi_{\mathrm{CO}}P_d^{CO}+\pi_{\mathrm{d}}P_d^{DT}-\pi_{\mathrm{d}}}{\pi_{\mathrm{CO}}}-\frac{1}{\pi_{\mathrm{CO}}-\pi_{\mathrm{DF}}}(\pi_{\mathrm{CO}}-\pi_{\mathrm{DF}})P_d^{CO}\leq 0
\end{aligned}$$

I show that the RHS is always greater than $(P_{dd}-P_{d(1-d)})/\pi_{\mathrm{CO}}$. I note that

$$P_{dd}-P_{d(1-d)}=\pi_{\mathrm{CO}}P_d^{CO}-\pi_{\mathrm{DF}}P_d^{DF}\leq\pi_{\mathrm{CO}}P_d^{CO}\leq\pi_{\mathrm{CO}}$$

Condition (iii) follows immediately. Condition (iv) follows immediately, if $P_{dd}-P_{d(1-d)}\geq 0$ as $\pi_{\mathrm{CO}}\leq\pi_{\mathrm{CO}}-\pi_{\mathrm{DF}}$. If $P_{dd}-P_{d(1-d)}<0$

$$P_{dd}-P_{d(1-d)}+\pi_{\mathrm{DF}}\delta=(\pi_{\mathrm{CO}}-\pi_{\mathrm{DF}})P_d^{CO}+\pi_{\mathrm{DF}}(P_d^{CO}-P_d^{DF})+\pi_{\mathrm{DF}}\delta\geq P_d^{CO}\geq 0$$

I show next that the RHS is always greater than $(P_{dd}-P_{d(1-d)}-\pi_{\mathrm{DF}}\delta)/(\pi_{\mathrm{CO}}-\pi_{\mathrm{DF}})$. (i) holds as

$$P_{dd}-P_{d(1-d)}-\pi_{\mathrm{DF}}\delta=\pi_{\mathrm{CO}}P_d^{CO}-\pi_{\mathrm{DF}}P_d^{DF}-\pi_{\mathrm{DF}}\delta\leq(\pi_{\mathrm{CO}}-\pi_{\mathrm{DF}})P_d^{CO}\leq\pi_{\mathrm{CO}}-\pi_{\mathrm{DF}}$$

Condition (ii) holds as

$$\begin{aligned}
&\frac{1}{\pi_{\mathrm{CO}}-\pi_{\mathrm{DF}}}\Big(P_{dd}-P_{d(1-d)}-\pi_{\mathrm{DF}}\delta\Big)\leq\frac{1}{\pi_{\mathrm{CO}}-\pi_{\mathrm{DF}}}\Big(\pi_{\mathrm{CO}}P_d^{CO}-\pi_{\mathrm{DF}}P_d^{DF}-\pi_{\mathrm{DF}}\delta\Big)\\
&\leq\max\{0,\frac{1}{\pi_{\mathrm{CO}}-\pi_{\mathrm{DF}}}(\pi_{\mathrm{CO}}-\pi_{\mathrm{DF}})P_d^{CO}\}\leq\frac{\pi_{\mathrm{CO}}P_d^{CO}+\pi_{\mathrm{d}}P_d^{DT}}{\pi_{\mathrm{CO}}}\leq\frac{P_{dd}}{\pi_{\mathrm{CO}}}
\end{aligned}$$

and condition (iii) follows as the difference of the RHS and the LHS is negative and equals

$$\begin{aligned}
&\frac{\pi_{\mathrm{DF}}}{\pi_{\mathrm{CO}}(\pi_{\mathrm{CO}}-\pi_{\mathrm{DF}})}(P_{dd}-P_{d(1-d)})-\pi_{\mathrm{DF}}(\frac{1}{\pi_{\mathrm{CO}}}+\frac{1}{\pi_{\mathrm{CO}}-\pi_{\mathrm{DF}}}\delta)\\
&=\frac{\pi_{\mathrm{DF}}}{(\pi_{\mathrm{CO}}-\pi_{\mathrm{DF}})}(P_d^{CO}-P_d^{DF})+\frac{\pi_{\mathrm{DF}}}{\pi_{\mathrm{CO}}}P_d^{DF}-\pi_{\mathrm{DF}}(\frac{1}{\pi_{\mathrm{CO}}}+\frac{1}{\pi_{\mathrm{CO}}-\pi_{\mathrm{DF}}}\delta)\leq 0
\end{aligned}$$

which proves the claim.

It follows that the breakdown frontier for the claim that $\Delta_{CO}\geq\mu$ is given by the following expression, if $(\pi_{\mathrm{DF}},BP(\pi_{\mathrm{DF}}))$ lies in the sensitivity region and otherwise by $(\pi_{\mathrm{DF}},\underline{\delta}(\pi_{\mathrm{DF}}))$.

$BP(\pi_{\text{DF}})$ equals

$$\frac{1}{\pi_{\text{DF}}}\max\Big\{-(\mu\cdot\pi_\Delta)+P_{00}-P_{01}),-((\mu-\frac{P_{11}-\pi_{\text{AT}}}{\pi_{\text{CO}}})\pi_\Delta+P_{00}-P_{01}),$$
$$-((\mu-\frac{P_{11}-P_{01}}{\pi_{\text{CO}}})\pi_\Delta+P_{00}-P_{01}),P_{11}-P_{10}-(\mu+1)\pi_\Delta,P_{11}-P_{10}-(\mu+\frac{P_{11}}{\pi_{\text{CO}}})\pi_\Delta,$$
$$P_{11}-P_{10}-(\mu+\frac{P_{11}-P_{10}+\pi_{\text{DF}}}{\pi_{\text{CO}}})\pi_\Delta,\frac{1}{2}(P_{11}-P_{01}-P_{00}+P_{10}-\mu\cdot\pi_\Delta)\Big\}.$$

□

B.6. **Proof of Theorem 6.1.** To simplify the exposition, I present and proof the following Lemma.

Lemma B.1. *Let $A_y=[\underline{y},y]$. It holds that $\sqrt{n}(\widehat{\widetilde{\theta}}(y,s)-\widetilde{\theta}(y,s))$ is given by*

$$\sqrt{n}\begin{pmatrix}\int_{A_y}\min\{\hat{q}_{ss}(z),\hat{q}_{s(1-s)}(z)\}dz-\int_{A_y}\min\{q_{ss}(z),q_{s(1-s)}(z)\}dz\\ \int_{A_y}\hat{q}_{ss}(z)dz-\int_{A_y}q_{ss}(z)dz\\ \int_{A_y}\hat{q}_{s(1-s)}(z)dz-\int_{A_y}q_{s(1-s)}(z)dz\end{pmatrix}\to\mathcal{Z}_6(y,s),$$

where $\mathcal{Z}_6(y,s)$ is a tight Gaussian process with continuous path in $\ell^\infty(\mathbb{R}\times\{0,1\},\mathbb{R}^3)$ with zero mean and variance given by $\Sigma_6(y,s,\widetilde{y},\widetilde{s})=0_{3\times3}$ zero if $\widetilde{s}\neq s$ and otherwise by[33]

$$\Sigma_6(y,s,\widetilde{y},s)=\begin{pmatrix}\sigma_{11}(y,s,\widetilde{y},s)&\sigma_{12}(y,s,\widetilde{y},s)&\sigma_{13}(y,s,\widetilde{y},s)\\ \sigma_{12}(y,s,\widetilde{y},s)&\sigma_{22}(y,s,\widetilde{y},s)&0\\ \sigma_{13}(y,s,\widetilde{y},s)&0&\sigma_{33}(y,s,\widetilde{y},s)\end{pmatrix}\qquad \textit{with}$$

$$\sigma_{11}(y,s,\widetilde{y},s)=(p_{ss}(y\wedge\widetilde{y})-p_{ss}(y)p_{ss}(\widetilde{y})+p_{s(1-s)}(y\wedge\widetilde{y})-p_{s(1-s)}(y)p_{s(1-s)}(\widetilde{y}))$$
$$\sigma_{22}(y,s,\widetilde{y},s)=(p_s(y\wedge\widetilde{y})-p_s(y)p_s(\widetilde{y}))\qquad\sigma_{33}(y,s,\widetilde{y},s)=(p_{1-s}(y\wedge\widetilde{y})-p_{1-s}(y)p_{1-s}(\widetilde{y}))$$
$$\sigma_{12}(y,s,\widetilde{y},s)=(p_{ss}(y\wedge\widetilde{y})-p_{ss}(y)p_{ss}(\widetilde{y}))\quad\sigma_{13}(y,s,\widetilde{y},s)=(p_{s(1-s)}(y\wedge\widetilde{y})-p_{s(1-s)}(y)p_{s(1-s)}(\widetilde{y}))$$

where $C_{ss}(y)=\{z\in A_y:q_{ss}(z)<q_{s(1-s)}(z)\}$, $C_{s(1-s)}(y)=\{z\in A_y:q_{ss}(z)>q_{s(1-s)}(z)\}$ and

$$p_s(y)=\mathbb{P}(Y_s\in A_y,D_s=s)\quad p_{1-s}(y)=\mathbb{P}(Y_{1-s}\in A_y,D_{1-s}=s)$$
$$p_{ss}(y)=\mathbb{P}(Y_s\in C_{ss}(y),D_s=s)\quad p_{s(1-s)}(y)=\mathbb{P}(Y_{1-s}\in C_{s(1-s)}(y),D_s=s).$$

[33] I denote $\min(a,b)=a\wedge b$.

B.6.1. *Proof of Lemma B.1.* This proof uses arguments of Anderson et al. (2012). By Assumption 3 (iv), it holds that

$$\begin{aligned}
&\sqrt{n}\int_{A_y} \min\{\hat{q}_{ss}(z), \hat{q}_{s(1-s)}(z)\} - \min\{q_{ss}(z), q_{s(1-s)}(z)\}dz \\
&= \sqrt{n}\int_{C_{ss}(y)} \min\{\hat{q}_{ss}(z), \hat{q}_{s(1-s)}(z)\} - q_{ss}(z)dz + \\
&\quad \sqrt{n}\int_{C_{s(1-s)}(y)} \min\{\hat{q}_{ss}(z), \hat{q}_{s(1-s)}(z)\} - q_{s(1-s)}(z)\}dz
\end{aligned}$$

I first consider the first summand, where it holds that

$$\begin{aligned}
&\int_{C_{ss}(y)} \min\{\hat{q}_{ss}(z), \hat{q}_{s(1-s)}(z)\} - q_{ss}(z)dz = \int_{C_{ss}(y)} \hat{q}_{ss}(z) - \mathbb{E}q_{ss}(z)dz - R_n^s(y) \\
&\text{with} \qquad R_n^s(y) = \int_{C_{ss}(y)} \min\{\hat{q}_{ss}(z), \hat{q}_{s(1-s)}(z)\} - q_{ss}(z) - \hat{q}_{ss}(z) + \mathbb{E}q_{ss}(z)dz.
\end{aligned}$$

I note that it further follows that $R_n^s(y)$ equals

$$\sqrt{n}\int_{C_{ss}(y)} \min\{\hat{q}_{ss}(z) - q_{ss}(z), \hat{q}_{s(1-s)}(z) - \mathbb{E}\hat{q}_{s(1-s)}(z) + q_{s(1-s)}(z) - q_{ss}(z)\}dz + O_p(n^{1/2}h^2)$$

I further note that

$$\begin{aligned}
&|\sqrt{n}\int_{C_{ss}(y)} \min\{\hat{q}_{ss}(z) - q_{ss}(z), \hat{q}_{s(1-s)}(z) - \mathbb{E}\hat{q}_{s(1-s)}(z) + q_{s(1-s)}(z) - q_{ss}(z)\}dz \\
&\quad - \int_{C_{ss}(y)} \hat{q}_{ss}(z) - q_{ss}(z)dz| \\
&= \sqrt{n}\int_{C_{ss}(y)} \max\{0, \hat{q}_{s(1-s)}(z) - \hat{q}_{ss}(z) - \mathbb{E}\hat{q}_{s(1-s)}(z) + \mathbb{E}q_{ss}(z) - q_{s(1-s)}(z) - q_{ss}(z)dz\}dz
\end{aligned}$$

It is clear that for any $\varepsilon > 0$

$$\begin{aligned}
&\mathbb{P}(\sqrt{n}\int_{C_{ss}(y)} \hat{q}_{s(1-s)}(z) - \hat{q}_{ss}(z) - \mathbb{E}\hat{q}_{s(1-s)}(z) + \mathbb{E}q_{ss}(z) > \varepsilon) \\
&\leq \mathbb{P}(|\hat{q}_{s(1-s)}(z) - \hat{q}_{ss}(z)| + |\mathbb{E}\hat{q}_{s(1-s)}(z) - \mathbb{E}q_{ss}(z)| > \varepsilon) = 0
\end{aligned}$$

and as $\mu(q_{s(1-s)}(z) - q_{ss}(z) > \varepsilon) = 0$ for ε sufficiently small the claim follows.

It therefore holds that

$$\sup_{y\in\mathbb{Y}} ||(\widehat{\widetilde{\theta}}(y,s) - \widetilde{\theta}(y,s)) - \begin{pmatrix} \int_{C_{ss}(y)} \hat{q}_{ss}(z) - q_{ss}(z)dz + \int_{C_{s(1-s)}(y)} \hat{q}_{s(1-s)}(z) - q_{s(1-s)}(z)dz \\ \int_{A_y} \hat{q}_{ss}(z)dz - \int_{A_y} q_{ss}(z)dz \\ \int_{A_y} \hat{q}_{s(1-s)}(z)dz - \int_{A_y} q_{s(1-s)}(z)dz \end{pmatrix}||$$

is $o_P(\sqrt{1/n})$ It then follows from Theorem 4 in Giné and Nickl (2008) that

$$(\widehat{\widetilde{\theta}}(y,s) - \widetilde{\theta}(y,s)) \to \mathcal{Z}_6(y,s),$$

because of the structure of the set C_{ss}.[34] $\square$

Theorem 6.1 directly follows from Lemma B.1 by observing that $\theta(y,s)$ is a a linear transformation of $\tilde{\theta}$. $\square$

B.7. **Proof of Lemma A.2.** I first show that ψ^U_{av} satisfies the above criterion

1. Follows immediately, as $|f(y)| = \sqrt{f(y)}^2 \leq \sqrt{f(y) + 1/\kappa}^2$ for all $y \in \mathbb{Y}$ and any $f \in l^\infty(\mathbb{Y})$.
2. It is also clear that $\sqrt{f(y) + 1/\kappa}^2 \to |f(y)|$ uniformly for all $x \in \mathbb{R}$ as $\kappa \to \infty$.
3. It is also clear that $(\psi^U_{av}(f;\kappa))'(y) = (f(y)^2 + \frac{1}{\kappa})^{-1/2} \cdot f'(y)$ and $(f(y)^2 + \frac{1}{\kappa}) \geq \frac{1}{\kappa}$ and therefore the denominator is bounded away from zero. ψ^U_{av} is Hadamard-differentiable.

Second, I show that ψ^L_{av} satisfies the above criterion

1. Follows immediately, as $|f(y)| = f(y)\frac{f(y)}{|f(y)|} \geq \frac{f(y)^2}{\psi^U_{av}(f)}$ for all $y \in \mathbb{Y}$.
2. It is also clear that as $\sqrt{f(x) + 1/\kappa}^2 \to |f(x)|$ uniformly for all $x \in \mathbb{R}$ as $\kappa \to \infty$, $\frac{f(x)^2}{\psi^U_{av}(f)} \to f(x)\frac{f(x)}{|f(x)|} = |f(x)|$.
3. It is also clear that

$$(\psi^{av}_\kappa)' = \frac{2f(y)f'(y)}{\sqrt{f(y)^2 + \frac{1}{\kappa}}} - (f(x)^2 + \frac{1}{\kappa})^{-1/2} \cdot f'(y) \cdot \frac{f(y)^2}{f(y)^2 + \frac{1}{\kappa}}$$

and $(f(x)^2 + \frac{1}{\kappa}) \geq \frac{1}{\kappa}$ and therefore the denominator is bounded away from zero, that this function is Hadamard-differentiable.

$\square$

B.8. **Proof of Lemma A.3.** I first show that $\psi^L_{\sup,\leq}$ satisfies the above criterion

1. Follows immediately, as for any $y \in \mathbb{Y}$ and any $\kappa \in \mathbb{N}$

$$\begin{aligned}\psi^L_{\sup,\leq}(f,g;\kappa)(y) &= \psi^L_{\max}(\{g(k_j) - f(k_j)\}_{j:k_j\leq y};\kappa) \leq \max(\{g(k_j) - f(k_j)\}_{j:k_j\leq y}) \\ &\leq \sup_{\tilde{y}\leq y}(g(\tilde{y}) - f(\tilde{y})) = \psi_{\sup,\leq}(f,g).\end{aligned}$$

[34] The results of Giné and Nickl (2008) show that under the assumptions of this Theorem $\int_{\underline{y}}^{y} f(y)\,(\hat{q}_{ss}(z)dz - q_{ss}(z))\,dz$ converges to a Gaussian-Bridge as long as $f(y)$ is Donsker. As by Assumption 3(iv), $f(y)$ can be represented as a finite collection of indicator functions, $f(y)$ is obviously Donsker.

The first inequality follows from the definition of $\psi^L_{\max}$ and the second inequality as g and f are nondecreasing.

2. It follows that for $\kappa \to \infty$, $\psi^L_{\sup,\leq}(f, g; \kappa) \to \psi_{\sup,\leq}(f, g)$

3. Hadamard-differentiability follows from the chain rule of Hadamard-differentiable functions Theorem 20.9 van der Vaart (1998) and as the difference is a linear operator.

I second consider $\psi^U_{\sup,\leq}$.

1. Follows immediately, as for any $y \in \mathbb{Y}$ and any $\kappa \in \mathbb{N}$

$$\begin{aligned}\psi^U_{\sup,\leq}(f, g; \kappa)(y) &= \psi^U_{\max}(\{g(k_j) - f(\min(y, k_{j+1}))\}_{j:k_j<y}\ ; \kappa)\\ &\geq \max(\{g(k_j) - f(\min(y, k_{j+1}))\}_{j:k_j<y})\\ &\geq \sup_{\tilde{y}\leq y}(g(\tilde{y}) - f(\tilde{y})) = \psi_{\sup,\leq}(f, g).\end{aligned}$$

The first inequality follows from the definition of $\psi^U_{\max}$ and the second inequality as g and f are nondecreasing.

2. It follows that for $\kappa \to \infty$, $\psi^L_{\sup}(f, g; \kappa) \to \psi_{\sup,\leq}(f, g)$ by continuity of f and g.

3. Hadamard-differentiability follows from the chain rule of Hadamard-differentiable functions Theorem 20.9 van der Vaart (1998) and as the difference is a linear operator.

The same arguments apply to the other mappings $\psi^U_{\sup,\geq}(f, g; \kappa)$, $\psi^U_{\sup,\geq}(f, g; \kappa)$ $\psi^L_{\inf,\leq}(f, g; \kappa)$, $\psi^U_{\inf,\leq}(f, g; \kappa)$, $\psi^L_{\inf,\geq}(f, g; \kappa)$, $\psi^U_{\inf,\geq}(f, g; \kappa)$.

□

B.9. **Proof of Proposition .** This mapping is by Lemma 20.10 by van der Vaart (1998) Hadamard differentiable. By the chain rule of Hamdard differentiable functions Lemma 20.9 by van der Vaart (1998) and Lemma **??** it follows that ϕ_{TE} is Hadamard differentiable. The Theorem then follows by the Delta-method Theorem 20.8 van der Vaart (1998)) and by Theorem 6.1.

B.10. **Proof of Theorem B.9.** The Theorem follows from Theorem 6.1 and the Delta method (see for example Theorem 20.8 van der Vaart (1998)), as all underlying functions are uniformly Hadamard-differentiable by construction and by Lemma A.2 and Lemma A.3. □

B.11. **Proof of Proposition 6.3.** The underlying parameters converge to a Gaussian distribution by Theorem 6.1. As the mapping is Hadamard differentiable by construction, the nonparametric bootstrap is consistent by Theorem 3.1 in Fang and Santos (2018). It therefore holds that

$$\lim_{n\to\infty} \mathbb{P}(\max_{k\leq 5} e'_k\ (\phi_{TE,\kappa}(\widehat{\theta}, \pi_{\text{DF}}) + cv_{1-\alpha}(\pi_{\text{DF}}) - \phi_{TE,\kappa}(\widehat{\theta}, \pi_{\text{DF}}))) \geq 1 - \alpha.$$

By definition of $\phi_{TE,\kappa}$ it further holds that

$$\lim_{n\to\infty} \inf_{\pi_{\rm DF}} \mathbb{P}(\max_{k\le 5} e_k' \ (\phi_{TE,\kappa}(\widehat{\theta},\pi_{\rm DF}) + cv_{1-\alpha}(\pi_{\rm DF}) - \phi_{TE}(\widehat{\theta},\pi_{\rm DF}))) \geq 1-\alpha,$$

which in turn implies the result that

$$\lim_{n\to\infty} \mathbb{P}(\widehat{RR}_{\kappa,B} \subseteq RR,\ \mathrm{SR} \subseteq \widehat{\mathrm{SR}}_{\kappa,B}) \geq 1-\alpha.$$

□

B.12. **Proof of Proposition A.5.** This Proposition is an application of Theorem 3.2 Fang and Santos (2018). It first follows from standard central limit arguments as θ^b is the arithmetic mean of binary outcome variables and by the assumed sampling process that

$$\sqrt{n}(\widehat{\theta}^b - \theta^b) \to \mathcal{N}(0_{1\times 5}, \Sigma),$$

where $\Sigma ii = \theta_{b,i}(1-\theta_{b,i})$. It is clear that $\sqrt{n}(\widehat{\theta}_b^* - \widehat{\theta}_b) \to \mathcal{N}(0,\Sigma)$. The proposition follows by showing that

$$\lim \inf_{n\to\infty} \mathbb{P}(\widehat{\phi}_n'(h) = \phi_\theta'(h) \forall h \in \mathbb{R}^6) = 1,$$

which holds trivially here as

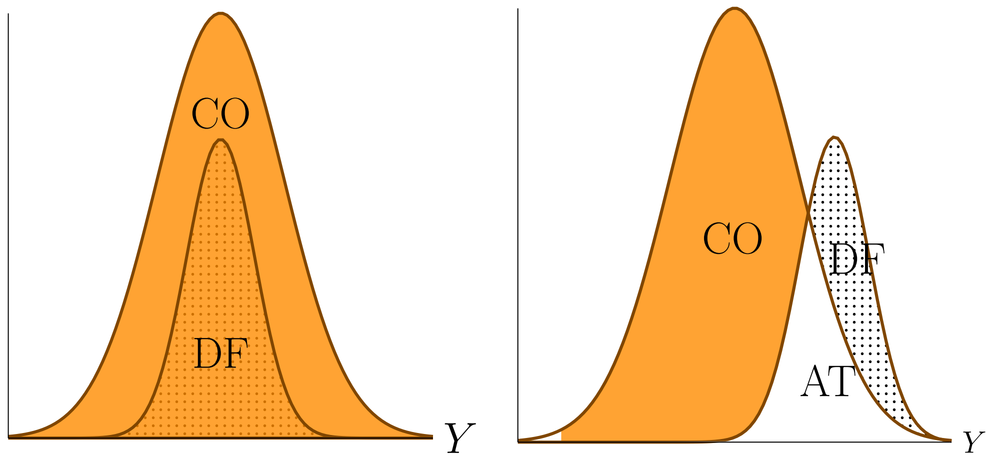
$$||\widehat{\phi}_n'(h_1) - \widehat{\phi}_n'(h_2)|| \leq C||h_1 - h_2||$$

and therefore for any $h \in \mathbb{R}^6$ it follows that $||\hat{\phi}_n'(h) - \phi_\theta(h)|| = o_P(1)$. It therefore follows from Theorem 3.2 Fang and Santos (2018) that the proposed bootstrap procedure is consistent.

□

C. ADDITIONAL RESULTS FOR THE SENSITIVITY REGION

In the section, I give some intuition on how the bounds on the sensitivity parameters δ are derived. Let us first consider the largest value $\overline{\pi}_{DF}$. This value implies that both the compliers and defiers' outcome distributions are point identified, as the population size of always takers would be zero. Thus the defiers' outcome distribution function equals $Q_{10}(y)$, and the compliers' one equals $Q_{11}(y)$ up to normalization. In this specific example, the outcome heterogeneity would be point identified but especially bounded from above by 0.5. Moreover, if I now consider the right panel and the smallest possible value of outcome heterogeneity $\pi_{\text{DF}} = \underline{\pi}_{DF}$, then the two outcome distributions are again point identified, and the outcome heterogeneity would be close to one, but especially it would be bounded from below.



(a) above for large values of π_{DF} (b) below for small values of π_{DF}.

Figure C.1: Breakdown Frontiers and Robust Regions for different Claims of δ

D. COMPARISON TO THE LITERATURE

In this section I show that the derived bounds are equivalent to the bounds of Balke and Pearl (1997) in the special case if no assumptions about the sensitivity parameters are imposed.

I first consider just the lower bound. Following the notation of this paper, Balke and Pearl (1997) show that $\mathbb{P}(Y = 1|D = 1)$ is bounded from below by

$$\begin{aligned}\max \{&P_{10}, P_{11}, P_{11} + \mathbb{P}(Y = 0, D = 0|Z = 1) - \mathbb{P}(Y = 0, D = 0|Z = 0) - \mathbb{P}(Y = 0, D = 1|Z = 0) \\ &P_{11} + \mathbb{P}(Y = 1, D = 0|Z = 1) - \mathbb{P}(Y = 0, D = 1|Z = 0) - \mathbb{P}(Y = 1, D = 0|Z = 0)\},\end{aligned}$$

In Section 5.4, I have shown that the compliers' distribution is sharply bounded by

$$\underline{P}_d^{CO}(\delta, \pi_{\text{DF}}) = \max\{0, \frac{P_{dd} - \pi_{\text{d}}}{\pi_{\text{CO}}}, \frac{P_{dd} - P_{d(1-d)}}{\pi_{\text{CO}}}, \frac{P_{dd} - P_{d(1-d)} - \pi_{\text{DF}}\delta}{\pi_{\text{CO}} - \pi_{\text{DF}}}\}$$

It follows from the reasoning of Section 5.1 that

$$\underline{P}_d(\delta, \pi_{\mathrm{DF}}) = \pi_{\mathrm{CO}} \underline{P}_d^{CO}(\delta, \pi_{\mathrm{DF}}) + P_{d(1-d)}$$

These bounds therefore simplify to

$$\begin{aligned}\underline{P}_d(\delta, \pi_{\mathrm{DF}}) &= \max\{P_{d(1-d)}, P_{dd} + P_{d(1-d)} - \pi_{\mathrm{d}}, P_{dd}\} \\ &= \max\{P_{d(1-d)}, P_{dd} + P_{d(1-d)} - \mathbb{P}(D = 1||Z = 0) + \pi_{\mathrm{DF}}, P_{dd}\}\end{aligned}$$

The second line follows as the population size of always takers is given by $\pi_{\mathrm{AT}} = \mathbb{P}(D = 1|Z = 0) - \pi_{\mathrm{DF}}$. I first note that the minimal population size of defiers is given by

$$\begin{aligned}\underline{\pi}_{DF} = \max\{0, &\mathbb{P}(Y = 1, D = 0|Z = 1) - \mathbb{P}(Y = 1, D = 0|Z = 0), \\ &\mathbb{P}(Y = 0, D = 0|Z = 1) - \mathbb{P}(Y = 0, D = 0|Z = 0), \\ &\mathbb{P}(Y = 1, D = 1|Z = 0) - \mathbb{P}(Y = 1, D = 1|Z = 1), \\ &\mathbb{P}(Y = 0, D = 1|Z = 0) - \mathbb{P}(Y = 0, D = 1|Z = 0)\}.\end{aligned}$$

This bound then obviously equals the bounds of Balke and Pearl (1997) for $\pi_{\mathrm{DF}} = \underline{\pi}_{DF}$.

Consider similarly the upper bound. Balke and Pearl (1997) show that $\mathbb{P}(Y = 1|D = 1)$ is bounded from above by

$$\begin{aligned}\min\{&1 - \mathbb{P}(Y = 0, D = 1|Z = 1), 1 - \mathbb{P}(Y = 0, D = 1|Z = 0) \\ &P_{11} + P_{10} + \mathbb{P}(Y = 0, D = 0|Z = 0) + \mathbb{P}(Y = 1D = 0|Z = 1) \\ &P_{11} + P_{10} + \mathbb{P}(Y = 1, D = 0|Z = 0) + \mathbb{P}(Y = 0, D = 0|Z = 1)\}.\end{aligned} \tag{16}$$

Following Section 5.4, the sharp bound on the population size of compliers is given by

$$\overline{P}_d^{CO}(\delta, \pi_{\mathrm{DF}}) = \min\{1, \frac{P_{dd}}{\pi_{\mathrm{CO}}}, \frac{P_{dd} - P_{d(1-d)} + \pi_{\mathrm{DF}}}{\pi_{\mathrm{CO}}}, \frac{P_{dd} - P_{d(1-d)} + \pi_{\mathrm{DF}}\delta}{\pi_{\mathrm{CO}} - \pi_{\mathrm{DF}}}\},$$

where the distribution of the population potential outcome variable is bounded by

$$\begin{aligned}\underline{P}_d(\delta, \pi_{\mathrm{DF}}) &= \pi_{\mathrm{CO}} \underline{P}_d^{CO}(\delta, \pi_{\mathrm{DF}}) + P_{d(1-d)} + \pi_{\mathrm{d}} \\ &= \min\{\pi_{\mathrm{CO}} + P_{d(1-d)} + \pi_{\mathrm{d}}, P_{dd} + P_{d(1-d)} + \pi_{\mathrm{d}}, P_{dd} + \pi_{\mathrm{DF}} + \pi_{\mathrm{d}}\} \\ &= \min\Big\{P_{d(1-d)} + \mathbb{P}(D = d|Z = d), P_{dd} + P_{d(1-d)} + \mathbb{P}(D = d|Z = 1 - s) - \pi_{\mathrm{DF}}, \\ &\qquad P_{dd} + \mathbb{P}(D = d|Z = 1 - d)\}\end{aligned}$$

By plugging-in $\underline{\pi}_{DF}$, the bound of Balke and Pearl (1997) follows.

Considering a continuous outcome model, a similar reasoning applies to the bounds of Kitagawa (2020).