Sensitivity Analysis for Linear Estimators

Jacob Dorn and Luther Yap*

November 3, 2025

Abstract

We propose a novel unified sensitivity analysis framework for linear estimators. Our approach measures the degree of identification failure through the change in measure between the observed distribution and a hypothetical target distribution that would identify the structural parameter of interest. The target distribution is synthetic, but formalizes the known correspondence between tools for distributionally robust optimization and sensitivity analysis of structural parameters. We illustrate the framework by generalizing existing bounds for Average Potential Outcome (APO) designs to allow unbounded likelihood ratios, and providing new bounds for regression discontinuity (RD), difference in differences (DD) and instrumental variables (IV) designs. We apply our framework to two empirical applications in RD to show how sensitivity analysis allows us to obtain bounds that are far more informative than worst-case bounds.

1 Introduction

Many important research designs in economics employ estimators that are linear in observed outcomes. These designs include regression discontinuity, differences in differences, and instrumental variables. These linear estimators leverage powerful identifying restrictions to target meaningful estimands. In observational settings, these identifying restrictions may

*This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE-2039656 and by grant T32 HS026116 from the Agency for Healthcare Research and Quality. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation or the Agency for Healthcare Research and Quality.

not be satisfied: many treatment choices can be selected on unobservables, manipulators can strategically sort across treatment cutoffs, and instruments can directly affect outcomes.

In light of how identifying restrictions for linear estimators may not hold, there is a large but disparate literature that places bounds on the object of interest when these restrictions fail. Many of these proposals are specific to the design: Gerard et al. (2020) is specific to regression discontinuity; Masten and Poirier (2018) is specific to selected treatments; Ramsahai (2012) is specific to instrumental variables; and Rambachan and Roth (2023) is specific to differences in differences. The core econometric problem in all of these designs is linear in outcomes. As a result, there is an open question of whether there is a unified framework for sensitivity analysis across these linear estimators. Having a unified framework removes the need to develop case-specific methods, allows researchers to compare sensitivity across different identification strategies, and allows the technology developed in one case to be more easily transferable to other linear estimators. We provide such a unified framework.

While constructing the unified framework is the primary goal, it turns out that our framework also solves a few more problems in the literature. First, there is a large literature for bounding average potential outcomes (APO) when there is treatment selection under bounds on selection in terms of odds (Tan, 2006) or propensity scores (Masten and Poirier, 2018). We provide a simple unification of these frameworks, including a characterization that is smooth when there is a region in which the bounds barely allow a propensity of zero. Second, in the regression discontinuity (RD) design with manipulation, there is currently no method available to bound the average treatment effect at the cutoff. A suite of bounds for conditional treatment effects after manipulation follow from our analysis as a corollary.

In this paper, we propose a framework for the sensitivity analysis of identification failures for linear estimators. The framework proceeds as follows. First, we define a target distribution: a synthetic distribution over observed variables that would enable a standard estimator to point-identify the structural estimand of interest. Second, we consider a structural model that implies restrictions on the divergence between the target population and the population the practitioner observes. Third, we leverage work from the literature in statistics, economics, and distributionally robust optimization (DRO) to estimate bounds implied by the restrictions from the second step. The framework is especially powerful when the implied restrictions correspond to a family of restrictions on the Radon-Nikodym derivative (the generalization of likelihood ratio to allow point masses) between the observed distribution of observables and the distribution of observables under the target distribution.

Our proposed framework has several empirical advantages. The target distribution only has to be defined once per application. Once the target distribution is defined, structural restrictions on the true distribution imply bounds on the target distribution. The approach typically yields a sensitivity analysis: under the strongest restriction, the bounds reduce to the standard point estimate; under the weakest restriction, the bounds reduce to a Manski-type worst-case exercise. Between the two extremes, restrictions on the structural model often correspond to new restrictions on the target distribution that facilitate computation based on the emerging average potential outcome literature.

Unifying the existing approaches requires a novel framework. There is a burgeoning literature on partial identification of the APO under bounds on the strength of confounding of treatment selection. In other settings like RD, there is no immediate correspondence from the object of interest to propensity scores. Our proposed synthetic-population framework defines a synthetic target population that immediately extends APO partial identification results to accommodate sensitivity analysis in these other settings.

We illustrate the framework by generalizing existing APO results to allow propensity score bounds that include zero. Without care, a needlessly conservative characterization of the partially identified set may be unbounded when Manski (1990) worst-case bounds show that the identified set is finite. We study an extension of the APO literature where distributional restrictions correspond to limits on treatment selection through L_∞ bounds on the Radon–Nikodym derivative (Tan, 2006; Masten and Poirier, 2018; Zhao et al., 2019; Dorn and Guo, 2023). By rewriting and extending Tan (2024a)'s bounds, we obtain an expression that prevents incorrectly invoking infinite bounds, thereby enabling us to provide unified bounds for families of assumptions between unconfoundedness and Manski-type bounds. The bounds depend on nuisance functions that are typically estimated in the primary analysis and a certain conditional outcome quantile function; plugging in a misspecified quantile estimate yields bounds that are too wide rather than too narrow. As a corollary, we obtain a simple characterization of bounds under Masten and Poirier (2018)'s conditional c-dependence model that bypasses the previous need to integrate the conditional quantile function with respect to the quantile index (Masten et al., 2024).

We apply our framework to yield novel results for three more applications. In our first application, we study bounds on causal effects from sharp RD with one-sided manipulation.

McCrary (2008) proposes a test for the RD assumption of no manipulation. When McCrary's test fails, Gerard et al. (2020) propose a worst-case bound on the conditional average treatment effect for non-manipulators: a Conditional Local Average Treatment Effect (CLATE). We show that a stronger restriction on manipulation choice allows us to obtain meaningful bounds on the more standard conditional average treatment effect (CATE), which to the best of our knowledge has been an open issue in the literature. Our framework provides a sensitivity analysis by nesting an unconfoundedness-type assumption and the Gerard et al. (2020) assumption as extreme cases.

In our second application, we study the method of differences in differences, which identifies the average treatment effect on the treated either through randomization (Athey and Imbens, 2022) or parallel trends (Callaway and Sant'Anna, 2021). Our framework allows us to bound the structural estimand with limited violations of randomization or parallel trends. In contrast to existing sensitivity analyses in DD designs (e.g., Rambachan and Roth (2023)), our sensitivity parameters are invariant to the functional form.

In our third application, we study treatment effects with an instrument that fails the exclusion or exogeneity restriction. When the instrument is allowed to affect the outcome directly, we consider estimation of a generic weighted average of local average treatment effects (LATEs) across instrument values. In the continuous outcome case, we contribute a sensitivity analysis for a measure of exclusion failure that is unit-free, and separates the potential outcomes for always-takers, never-takers, and compliers, which is more transparent than the convolution used in Ramsahai (2012). To our knowledge, our proposed sensitivity analysis includes the first unit-free measure of exclusion failure with continuous outcomes. The proposed bounds are simple and tractable. An important caveat is that the implied LATE bounds can be conservative because of the generality of our framework.

Our empirical applications quantify the value of our procedure for RD with manipulation. We first apply our procedure to study the incumbency advantage, in which previous work has evaluated the contribution of incumbent victory to the incumbent party winning in the following election. While the CATE and CLATE point estimates are identical, we find that the CLATE estimate is more robust to a given level of unobserved confounding. In our second application, we study the effects of council size on expenditure. The worst-case bounds are highly uninformative in this context, while our sensitivity analysis allows us to obtain meaningfully tighter bounds.

1.1 Related Work

Our framework unifies ideas from the operations research and economics literature. Two closely related frameworks are Bertsimas et al. (2022) and Christensen and Connault (2023), both of which are limited to discrete covariates. In the discrete covariate case, relative to Bertsimas et al., we propose justifying distributional distances in terms of underlying structural models of economic failure and propose constructing target distributions that apply in settings beyond average treatment effects, and offer a specific analysis of L_∞ bounds rather than other distributional distances; and relative to Christensen and Connault, we analyze a nominally different class of identification failures and provide closed-form bounds for specifically linear estimators.

Our work is related to the recent literature on sensitivity analysis for IPW estimators, which relates to our first application. A sensitivity analysis is an approach to partial identification that begins from assumptions that point-identify the structural estimand of interest and then considers increasing relaxations of those assumptions (Molinari, 2020). Our analysis extends Dorn and Guo (2023)'s sharp characterization of bounds beyond Tan (2006)'s marginal sensitivity model. Tan (2024a) and Frauen et al. (2023) previously extended this characterization to families that bound the Radon-Nikodym derivative of interest. We generalize these results to also include unbounded Radon-Nikodym derivative, so that we can include a compact characterization of bounds under Masten and Poirier (2018)'s conditional c-dependence model as a special case. Our paper uses the same DRO ideas as Dorn et al. (2024) and Tan (2024a), but we generalize the class of assumptions that can be considered in the unconfoundedness case and show how to use these ideas in settings besides unconfoundedness. There is rich work in this literature under other sensitivity assumptions like f-divergences and Total Variation distance (e.g. Huang and Pimentel, 2024; Freidling and Zhao, 2022; Chernozhukov et al., 2022; Christensen and Connault, 2023; Ishikawa et al., 2023)). These other assumptions also fit within our framework, because our target distribution constructions are independent of the L_∞ sensitivity assumptions that we analyze.

Our framework relates to existing work on sensitivity analysis for other applications. Our proposed sensitivity analysis for sharp RD applies when data on the running variable fails tests for manipulation (McCrary, 2008; Otsu et al., 2013; Bugni and Canay, 2021). Our proposal nests both an exogeneity-type assumption and Gerard et al. (2020)'s bounds as special cases. There is other work in the RD context on partial identification bounds under manipulation (Rosenman et al., 2019; Ishihara and Sawada, 2020) but to our knowledge, our proposal is the first sensitivity analysis for manipulation. There are sensitivity analysis for exclusion failure with instrumental variables (Ramsahai, 2012; Van Kippersluis and Rietveld, 2018; Masten and Poirier, 2021; Freidling and Zhao, 2022), but to our knowledge our proposal is the first sensitivity analysis whose underlying assumptions are invariant to invertible transformations of variables.

Notation. We use $\bar{\mathbb{R}}$ to refer to the extended real number line $\mathbb{R} \cup \{-\infty, \infty\}$. For a real-valued random variable Z and a distribution $\mathbb{Q}$, we use the notation $E_{\mathbb{Q}}[Z] = \int z d\mathbb{Q}(z)$ and we write that the expected value of Z exists under $\mathbb{Q}$ if $E_{\mathbb{Q}}[|Z|]$ is finite or $E_{\mathbb{Q}}[Z]$ is well-defined as exactly one of positive or negative infinity. We write that the expected value of Z exists (without specifying the distribution) if the expected value exists under the observed distribution $\mathbb{P}^{\text{Obs}}$, where $\mathbb{P}^{\text{Obs}}$ is defined below. Similarly, we sometimes suppress the dependence of expectations when referring to the expectation under the observed distribution $\mathbb{P}^{\text{Obs}}$. We write $\{a, b\}_+ = \max\{a, b\}$ and $\{a, b\}_- = \min\{a, b\}$, and abuse notation by writing $\{a\}_+ = \max\{a, 0\}$. For random variables $Y \in \mathbb{R}^1$ and $R \in \mathbb{R}^d$ and a function $t : \mathbb{R}^d \to [0, 1]$, we refer to "the" conditional quantile function $Q_{t(R)}(Y \mid R)$, which is any minimizer of $E_{\mathbb{P}^{\text{Obs}}}[t(R)\{Y - Q(R), 0\}_+ - (1 - t(R))\{Y - Q(R), 0\}_- \mid R]$ in functions Q from the domain of R to the extended real line $\mathbb{R} \cup \{-\infty, \infty\}$; when $t(R) \in \{0, 1\}$, we take the infimum and supremum of the distribution's support respectively. To consolidate notation, when b is infinite, we evaluate the interval $[a, b]$ as $[a, \infty)$. If our estimation procedure calls for an estimate of a nuisance function f that depends on another nuisance function g, we use the notation $\hat{f}$ to denote the full estimated nuisance function, which may include a composition of nuisance estimates. We use $1\{.\}$ to refer to the indicator function that takes value 1 if true and 0 otherwise. We use $\mathbb{P}$ to denote both the probability mass function for a discrete variable and the cumulative distribution function for a continuous (or mixed) variable, and use $d\mathbb{P}$ to denote the probability density of the a continuous (or mixed) variable.

2 Framework for Sensitivity Analysis

In the following section, we illustrate our generic framework for sensitivity analysis: we propose linking a structural model of identification failures to an implied statistical distribution bound and then conduct partial identification under the distributional bound. We characterize the identified set under a family of L_∞ distributional distances that is especially tractable for estimation. We leave specific structural models to Section 3.

2.1 General Setting

We begin with our general partial identification setting. We use the APO application as a running example.

Let $Y \in \mathbb{R}^1$ denote the outcome of interest, $R \in \mathcal{R}$ denote observable quantities that do not include Y, and $Y(r)$ denote potential outcomes. We call R the regressors, and the observed outcome is $Y = Y(R)$.

We consider three different but related probability measures. There is a true distribution $\mathbb{P}^{\text{True}}$ over $(\{Y(r)\}_{r\in\mathcal{R}}, R, \xi)$, where ξ are unobserved variables that do not include potential outcomes. This ξ will not feature in our running APO example, but it will become relevant in other applications in Section 3. While ξ itself is unobserved, some features of ξ may be point- or partially-identified. For instance, in the instrumental variables framework with potential treatments $T(z)$ where z is the instrument, ξ could include $T(z)$, which is itself unobserved, but the proportion of compliers, $\Pr(T(1) = 1, T(0) = 0)$, is identified in the population under monotonicity.

The object of interest is a linear reweighting over potential outcomes with treatment status r, $E_{\mathbb{P}^{\text{True}}}\left[\lambda(R)\int_r Y(r)dG(r \mid R)\right]$, where λ is a real-valued function and G is some known distribution weighting over values of r. We observe data (Y, R) from the observable distribution $\mathbb{P}^{\text{Obs}}$, which is a coarsening of $\mathbb{P}^{\text{True}}$. We face the fundamental problem of casual inference and do not observe $Y(r)$.

In our applications, for every distribution $\mathbb{P}^{\text{True}}$, there exists a target distribution $\mathbb{P}^{\text{Target}}$ that point identifies the structural estimand through the moment

$$E_{\mathbb{P}^{\text{Target}}}[\lambda(R)Y] = E_{\mathbb{P}^{\text{True}}}\left[\lambda(R)\int_r Y(r)dG(r \mid R)\right].$$

$\mathbb{P}^{\text{Target}}$ is not necessarily unique, and it measures the same space (i.e., of (R, Y)) as $\mathbb{P}^{\text{Obs}}$. The observed distribution $\mathbb{P}^{\text{Obs}}$ may not be able to point-identify the structural estimand under the researcher's preferred estimator, for example if R is a selected treatment or if R includes an instrument that fails the exclusion restriction.

Assumption Support. The marginal distribution of R is the same under $\mathbb{P}^{\text{Target}}$ and $\mathbb{P}^{\text{Obs}}$. The distribution $\mathbb{P}^{\text{Target}}$ is absolutely continuous with respect to $\mathbb{P}^{\text{Obs}}$.

Assumption Support has two components. The first component is that $\mathbb{P}^{\text{Obs}}$ identifies the target distribution of regressors R, but may have a different conditional distribution of outcomes Y. The second component is an absolute continuity assumption that ensures the existence of a Radon-Nikodym derivative $\frac{d\mathbb{P}^{\text{Target}}}{d\mathbb{P}^{\text{Obs}}}$. The assumption accommodates unbounded Radon-Nikodym derivatives, for example $Y \mid R \sim Unif(0,1)$ under $\mathbb{P}^{\text{Obs}}$ and $f_Y(y \mid R) = 1\{y \in (0,1]\}y^{-1/2}/2$ under $\mathbb{P}^{\text{Target}}$. The absolute continuity assumption is stronger than necessary for our results, and could be reduced to assuming that the support of (Y, R) under $\mathbb{P}^{\text{Target}}$ is a subset of the support under $\mathbb{P}^{\text{Obs}}$. Absolute continuity rules out $\mathbb{P}^{\text{Target}}$ possessing mass points where $\mathbb{P}^{\text{Obs}}$ lacks mass points, except where such distributions can be achieved as limits of distributions within Assumption Support.

Our framework is inspired by the literature on sensitivity analysis for the APO $E_{\mathbb{P}^{\text{True}}}[E_{\mathbb{P}^{\text{True}}}[Y(1) \mid R]]$. In this case, there is a distribution $\mathbb{P}^{\text{True}}$ over $(X, T, Y(1), Y(0))$, but we only observe the distribution $\mathbb{P}^{\text{Obs}}$ over $(X, T, Y = Y(T))$, so that $R = (X, T)$. The researcher can accurately estimate $\mathbb{P}^{\text{True}}(T \mid X)$, but cannot observe $\mathbb{P}^{\text{True}}(Y(1) \mid X, T = 0)$. We propose the target distribution $\mathbb{P}^{\text{Target}}$ that first samples $(X, T) \sim \mathbb{P}^{\text{True}}$ and then samples $Y \mid X, T$ from the distribution of $Y(T) \mid X$ under $\mathbb{P}^{\text{True}}$. To be precise, the target distribution $\mathbb{P}^{\text{Target}}$ is defined as $\mathbb{P}^{\text{Target}}(X, T, Y) = \mathbb{P}^{\text{Obs}}(X, T)d\mathbb{P}^{\text{True}}(Y(T) \mid X, T)$. The target distribution $\mathbb{P}^{\text{Target}}$ does not correspond to $\mathbb{P}^{\text{True}}$, but if a causal estimator like an inverse propensity weighted (IPW) estimator were applied to $\mathbb{P}^{\text{Target}}$ instead of $\mathbb{P}^{\text{Obs}}$, the researcher would identify the APO under $\mathbb{P}^{\text{True}}$. Our later examples illustrate the applicability of this framework for other structural estimands.

We refer to our target as the structural estimand, and the target value of the primary observational analysis as the primary estimand. We restrict ourselves to estimands that correspond to linear estimators.

Definition 1. The *structural estimand* is $\psi_0 = E_{\mathbb{P}^{\text{Target}}}[\lambda(R)Y]$ for some real-valued function λ. The *primary estimand* is $E_{\mathbb{P}^{\text{Obs}}}[\lambda(R)Y]$.

We generally assume that λ is identified and the researcher has conducted a primary analysis that consistently estimates λ. For example, in the APO case, the λ function is the inverse propensity $\frac{T}{\mathbb{P}^{\text{Obs}}(T=1|X)}$. When the researcher is able to make strong enough assumptions that $\mathbb{P}^{\text{Target}}$ is equivalent to $\mathbb{P}^{\text{Obs}}$, then the primary estimand will be equal to the structural estimand. When the researcher is only able to bound the difference between $\mathbb{P}^{\text{Target}}$ and $\mathbb{P}^{\text{Obs}}$, then the researcher can partially identify the structural estimand through a restriction on the Radon-Nikodym derivative between the two distributions.

Lemma 1. *Suppose Assumption Support holds and $\lambda(R)Y$ is integrable under $\mathbb{P}^{\text{Target}}$. Then*

$$\psi_0 = E_{\mathbb{P}^{\text{Obs}}}\left[\lambda(R)Y\frac{d\mathbb{P}^{\text{Target}}(Y \mid R)}{d\mathbb{P}^{\text{Obs}}(Y \mid R)}\right].$$

Proof. Lemma 1 is a standard result for Importance Sampling and an immediate property of the Radon-Nikodym derivative. □

This reweighing characterization is useful for partial identification because any non-negative putative outcome reweighing $\bar{W} = \lambda(R)\frac{d\mathbb{P}^{\text{Target}}(Y|R)}{d\mathbb{P}^{\text{Obs}}(Y|R)}$ that satisfies $E[\bar{W} \mid R] = 1$ almost surely will correspond to a well-defined distribution $\mathbb{Q}$ over (R, Y).

The Radon-Nikodym derivative characterization in Lemma 1 often maps to interpretable quantities. For example, in the APO case, the Radon-Nikodym derivative maps to odds ratios as follows[1]:

$$\begin{aligned}\lambda(R)\frac{d\mathbb{P}^{\text{Target}}(Y \mid R)}{d\mathbb{P}^{\text{Obs}}(Y \mid R)} &= \lambda(R)\mathbb{P}^{\text{Obs}}(T = 1 \mid X) \\ &\quad + \lambda(R)\mathbb{P}^{\text{Obs}}(T = 0 \mid X)\frac{\mathbb{P}^{\text{True}}(T = 0 \mid X, Y(1))}{\mathbb{P}^{\text{True}}(T = 1 \mid X, Y(1))}\frac{\mathbb{P}^{\text{Obs}}(T = 1 \mid X)}{\mathbb{P}^{\text{Obs}}(T = 0 \mid X)}.\end{aligned} \tag{1}$$

As a result, there is a bijection between structural models of treatment selection and causal models of treatment effects. More generally, principled restrictions on an underlying treatment selection model may imply, or be equivalent to, restrictions on the Radon-Nikodym derivative. The mapping from some structural model on objects such as $\frac{\mathbb{P}^{\text{Target}}(T=0|X,Y(1))}{\mathbb{P}^{\text{Target}}(T=1|X,Y(1))}$ to the Radon-Nikodym derivative as in Equation (1) is application-dependent, but holds across sensitivity models.

Our general framework defines $\mathbb{P}^{\text{Target}}$ in this way and derives a correspondence between

[1] To see why the above result is true, observe that, by applying Bayes' Rule and dropping conditioning on X, $d\mathbb{P}^{\text{Obs}}(Y \mid T = 1) = d\mathbb{P}^{\text{True}}(Y(1) \mid T = 1) = \frac{\mathbb{P}^{\text{True}}(T=1|Y(1))d\mathbb{P}^{\text{True}}(Y(1))}{\mathbb{P}^{\text{Obs}}(T=1)}$ and $d\mathbb{P}^{\text{Target}}(Y \mid T = 1) = \frac{\mathbb{P}^{\text{Target}}(T=1|Y(1))d\mathbb{P}^{\text{True}}(Y(1))}{\mathbb{P}^{\text{Obs}}(T=1)} = d\mathbb{P}^{\text{True}}(Y(1))$ as $\mathbb{P}^{\text{Target}}(T = 1 \mid Y(1)) = \mathbb{P}^{\text{Obs}}(T = 1)$. Hence, $\frac{d\mathbb{P}^{\text{Target}}(Y|R)}{d\mathbb{P}^{\text{Obs}}(Y|R)} = \frac{\mathbb{P}^{\text{Obs}}(T=1)}{\mathbb{P}^{\text{True}}(T=1|Y(1))}$, which is easily verified to be the same as the RHS by factoring out $\mathbb{P}^{\text{Obs}}(T = 1)$, then observing that $\mathbb{P}^{\text{True}}(T = 0 \mid X, Y(1)) + \mathbb{P}^{\text{True}}(T = 1 \mid X, Y(1)) = 1$.

bounds on $\mathbb{P}^{\text{True}}$ and bounds on the divergence between $\mathbb{P}^{\text{Target}}$ and $\mathbb{P}^{\text{Obs}}$. We provide formal results for this approach applied to a class of L_∞ sensitivity models.

2.2 Partial Identification Assumption and Result

In the this subsection, we characterize the sharp bounds on the identified set under an L_∞ restriction on the largest and smallest Radon-Nikodym derivative.

Definition 2 (Sensitivity assumption)**.** For any pair of functions $\underline{w}, \bar{w} : \mathbb{R}^d \to \bar{\mathbb{R}}$ satisfying $0 \le \underline{w}(R) \le 1 \le \bar{w}(R)$ almost surely, we define $\mathcal{M}(\underline{w}, \bar{w})$ as the set of distributions $\mathbb{Q}$ over $(\{Y(r)\}_{r\in\mathcal{R}}, R, Y = Y(R), \xi)$ satisfying $\lambda(R)\frac{d\mathbb{Q}(R,Y)}{d\mathbb{P}^{\text{Obs}}(R,Y)} \in [\lambda(R)\underline{w}(R), \lambda(R)\bar{w}(R)]$ almost surely.

The space that the distribution $\mathbb{Q}$ measures is the union of the space measured by $\mathbb{P}^{\text{True}}$ and $\mathbb{P}^{\text{Obs}}$. This family has several advantages. The restrictions on $\frac{d\mathbb{Q}}{d\mathbb{P}^{\text{Obs}}}$ decouple across values of R, enabling tractable characterizations of the identified set. The family nests extreme assumptions as special cases: point identification corresponds to the case $\underline{w}(R) = \bar{w}(R) = 1$ almost surely, while Manski-type bounds that only restrict the support of Y correspond to $\bar{w}(R) = \infty$ with domain-appropriate $\underline{w}(R)$. In between, the structural estimand is only partially identified. When the outcome Y is binary, the restriction can equivalently be viewed as a restriction on the conditional mean of $Y \mid R$. We show below that, as in the Tan (2006) model that inspired this generalization, the resulting bounds are highly tractable for estimating sharp and valid bounds.

We adapt standard notation from Ho and Rosen (2017).

Definition 3. The *identified set* is $\mathcal{I}(\underline{w}, \bar{w}) = \{E_{\mathbb{Q}}[\lambda(R)Y] \mid \mathbb{Q} \in \mathcal{M}(\underline{w}, \bar{w})\}$. The *sharp bounds* on the identified set are $\psi^-(\underline{w}, \bar{w}) = \inf_{\psi\in\mathcal{I}(\underline{w},\bar{w})} \psi$ and $\psi^+(\underline{w}, \bar{w}) = \sup_{\psi\in\mathcal{I}(\underline{w},\bar{w})} \psi$. A *valid* identified set is a superset of $\mathcal{I}(\underline{w}, \bar{w})$, and valid bounds is an interval that contains $[\psi^-(\underline{w}, \bar{w}), \psi^+(\underline{w}, \bar{w})]$. We call bounds *conservative* if they are valid but not sharp.

We abuse notation and simply write $\mathcal{I}$, ψ^-, and ψ^+ as shorthand for the identified set and bounds under a generic model family. We work under convex restrictions for which, as we verify, the distinction between bounds and the identified set largely collapses. Definition 3 corresponds to the statistical partial identification bounds implied by an underlying structural model. As we illustrate in our applications, a structural model underlying Definition 2 can sometimes, but not always, imply a narrower bound.

We will require some nuisance functions to characterize the identified set under the family of L_∞ restrictions we allow. In particular:

Definition 4. The threshold probability is $\tau(R) \equiv \frac{\bar{w}(R)-1}{\bar{w}(R)-\underline{w}(R)}$. The threshold quantiles are $Q^+(R) \equiv Q_{\tau(R)}(\lambda(R)Y \mid R)$ and $Q^-(R) \equiv Q_{1-\tau(R)}(\lambda(R)Y \mid R)$. The likelihood shifting term is

$$a(\bar{w}, \underline{w}, s) = (\bar{w} - \underline{w})1\{s > 0\} - (1 - \underline{w}).$$

For a function $\bar{Q} : \mathbb{R}^d \to \bar{\mathbb{R}}$, define the pseudo-outcome

$$\phi^+(\underline{w}, \bar{w}, r, y \mid \bar{Q}) \equiv \lambda(r)y + (\lambda(r)y - \bar{Q}(r))a(\bar{w}(r), \underline{w}(r), \lambda(r)y - \bar{Q}(r)),$$

and define $\phi^-(\underline{w}, \bar{w}, r, y \mid \bar{Q})$ analogously:

$$\phi^-(\underline{w}, \bar{w}, r, y \mid \bar{Q}) \equiv \lambda(r)y + (\lambda(r)y - \bar{Q}(r))a(\bar{w}(r), \underline{w}(r), \bar{Q}(r) - \lambda(r)y).$$

Define the indicator variables $F \equiv 1\{\bar{w}(R) \text{ finite}\}$, $G^+ \equiv 1\{\lambda(R)Y - Q^+(R) > 0\}$, $G^- \equiv 1\{\lambda(R)Y - Q^-(R) < 0\}$, and $H \equiv 1\{\tau(R) < 1\}$.

When $\bar{w}(R)$ is finite, the formula for $\tau(R)$ can be found in Tan (2024a) and Frauen et al. (2023).

We will make additional regularity assumptions.

Assumption Moments. The expected values of $F|Q^+(R)|$, $F|Q^-(R)|$, $|\phi^+(\underline{w}, \bar{w}, R, Y \mid Q^+)|$ and $|\phi^-(\underline{w}, \bar{w}, R, Y \mid Q^-)|$ are well-defined.

We allow infinite $E_{\mathbb{P}^{\text{Obs}}}[Q^+(R)]$ for completeness with unbounded outcomes. While $\bar{w}$ can be infinite, we rule out heavy tails that would yield infinite $E_{\mathbb{P}^{\text{Obs}}}[1\{\bar{w}(R) \text{ finite}\}Q^+(R)]$.

Our work focuses on the upper bound of the identified set.

Theorem 1. *Suppose Assumptions Support and Moments hold. Then the sharp upper bound on the identified set is:*

$$\psi^+(\underline{w}, \bar{w}) = E_{\mathbb{P}^{\text{Obs}}}\left[\lambda(R)Y + (\lambda(R)Y - Q^+(R))a(\bar{w}(R), \underline{w}(R), \lambda(R)Y - Q^+(R))\right]. \quad (2)$$

The sharp lower bound $\psi^-(\underline{w}, \bar{w})$ follows symmetrically, i.e.,

$$\psi^-(\underline{w}, \bar{w}) = E_{\mathbb{P}^{\text{Obs}}}\left[\lambda(R)Y + (\lambda(R)Y - Q^-(R))a(\bar{w}(R), \underline{w}(R), Q^-(R) - \lambda(R)Y)\right]. \quad (3)$$

Further, the identified set is convex.

The theorem states that the sharp upper bound can be written as a closed-form moment: after the constituent components such as $\lambda(.), Q^+(.)$, and $a(.)$ have been estimated, no further optimization problem needs to be solved. Proofs are in the appendix.

A proof sketch is in order. We abuse notation to write $\psi^+ = E_{\mathbb{P}^{\text{Obs}}}[\psi^+(R)]$, where $\psi^+(r)$ is a pointwise upper bound function satisfying:

$$\psi^+(R) = \sup_W E_{\mathbb{P}^{\text{Obs}}}[W\lambda(R)Y \mid R] \quad \text{s.t.} \quad W \in [\underline{w}(R), \bar{w}(R)] \quad \& \quad E_{\mathbb{P}^{\text{Obs}}}[W \mid R] = 1,$$

and $\phi^+(R) \equiv E_{\mathbb{P}^{\text{Obs}}}[\phi^+(\underline{w}, \bar{w}, R, Y \mid Q^+) \mid R]$. The claim is $E_{\mathbb{P}^{\text{Obs}}}[\psi^+(R)] = E_{\mathbb{P}^{\text{Obs}}}[\phi^+(R)]$. We prove the stronger claim that $\psi^+(R) = \phi^+(R)$ almost surely. For simplicity, this sketch will ignore the possibility that $\lambda(R)Y = Q^+(R)$ with positive probability and drop "almost sure" caveats. As Dorn et al. (2024) note, $\psi^+(R)$ can be mapped to a simple DRO problem over $(W - \underline{w}(R))/(1 - \underline{w}(R)) \in [0, (\bar{w}(R) - \underline{w}(R))/(1 - \underline{w}(R))]$. The solution is $\psi^+(R) = E_{\mathbb{P}^{\text{Obs}}}[\underline{w}(R)\lambda(R)Y + (1 - \underline{w}(R))CVaR^+_{\tau(R)}(R) \mid R]$, where $CVaR^+_{\tau(R)} = E\left[Q^+ + \frac{\{Y - Q^+\}_+}{1 - \tau(R)} \mid R\right]$ is the level-$\tau(R)$ conditional value at risk of $\lambda(R)Y$.

We split on the event $\tau(R) < 1$: by previous work (Tan, 2024a; Frauen et al., 2023), $\tau(R) < 1$ implies $\psi^+(R) = \phi^+(R)$. Theorem 1 is novel for the case $\tau(R) = 1$, which corresponds to either $\underline{w}(R) = 1$ or $\bar{w}(R) = \infty$ ($e(X) \notin (c, 1 - c)$ in the case of conditional c-dependence). On that event, $\phi^+(R)$ evaluates to $E[\underline{w}(R)\lambda(R)Y + (1 - \underline{w}(R))Q^+(R) \mid R]$ and $Q^+(R) = CVaR^+_1(R)$. As a result, $\psi^+(R) = \phi^+(R)$ even if $\bar{w}(R)$ is infinite with positive probability. Intuitively, at extreme probabilities, the conditional value at risk and quantile either converge (with bounded outcomes) or tend to the same infinite limit (with unbounded outcomes). In either case, replacing the conditional value at risk with the quantile has no effect. In the other extreme case when $\underline{w}(R) = \bar{w}(R) = 1$, we have $a(\bar{w}, \underline{w}, s) = 0$ so we have point identification regardless of the assignment of $\tau(R) \in (0, 1)$; in our implementation, we take $\tau(R) = 0.5$.

One could characterize the partial identification bounds using the conditional value at risk directly (Dorn et al., 2024), but the reweighing characterization Equation (2) is valuable for sensitivity analysis. We present the worst-case Radon-Nikodym derivative here for convenience.

Lemma 2. *Suppose $\bar{w}(R)$ is bounded. Then one can construct a distribution $\mathbb{Q}^+ \in \mathcal{M}(\underline{w}, \bar{w})$*

and a random variable satisfying $\gamma(R) \in [\underline{w}(R), \bar{w}(R)]$ *almost surely such that* $\psi^+ = E_{\mathbb{Q}^+}[\lambda(R)Y]$ *and:*

$$\frac{d\mathbb{Q}^+(R,Y)}{d\mathbb{P}^{\mathrm{Obs}}(R,Y)} = W^*_{sup} = \begin{cases} \bar{w}(R) & \text{if} \quad \lambda(R)Y > Q^+(R) \\ \underline{w}(R) & \text{if} \quad \lambda(R)Y < Q^+(R) \\ \gamma(R) & \text{if} \quad \lambda(R)Y = Q^+(R). \end{cases} \tag{4}$$

The construction in Equation (2) sidesteps the need to specify $\gamma(R)$.

A researcher with a primary estimate of $\lambda(R)$ and who is capable of quantile regression can easily construct a plug-in estimate of ψ^+. Even if the researcher fails to consistently estimate the additional nuisance parameter Q^+, they can still easily estimate valid bounds:

Lemma 3. *Suppose Assumptions Support and Moments hold and there is a putative quantile function* $\bar{Q}$ *such that the expectation of* $|\phi^+(\underline{w}, \bar{w}, r, y \mid \bar{Q})|$ *exists. Then replacing* Q^+ *with the potentially incorrect quantile function* $\bar{Q}$ *yields valid bounds:*

$$\psi^+(\underline{w}, \bar{w}) \leq E_{\mathbb{P}^{\mathrm{Obs}}}\left[\lambda(R)Y + (\lambda(R)Y - \bar{Q}(R))a(\bar{w}(R), \underline{w}(R), \lambda(R)Y - \bar{Q}(R))\right].$$

An analogous result holds for the lower bound ψ^-*.*

The argument generally follows by Tan (2024a)'s decomposition of $(\lambda(R)Y - \bar{Q}(R))a(\ldots)$ in terms of the Quantile Regression check function, although some care is needed to handle the case $\tau(R) = 1$. $Q^+(R)$ is the minimizer of the associated check function by classic arguments.

Note that a similar recipe as in this section extends to other sensitivity assumptions. With other restrictions on $\mathbb{Q}$, the model family in Definition 2 would change, the sharp upper bound in Theorem 1 would change and may lack a closed form, and the validity result Lemma 3 may or may not hold, but the fundamental logic would carry through.

2.3 Illustration with Average Potential Outcomes

To complete the exposition of our procedure, we show how the framework applies to our APO running example. Many of our results are restatements of recent work; the extension of partial identification bounds to unbounded Radon-Nikodym derivatives is new.

We assume there is a distribution $\mathbb{P}^{\mathrm{True}}$ over $(X, T, \{Y(t)\}_{t\in\{0,1\}}, U)$, where X are controls, T is a discrete treatment, $Y(t)$ is the potential outcome corresponding to treatment level t, and U are potential unobserved confounders. We only observe the coarsened distribution $\mathbb{P}^{\text{Obs}}$ over $(X, T, Y = Y(T))$, where Y is the observed outcome. We write $I_t = 1\{T = t\}$.

We tailor our application to IPW estimation of the APO, and consider average treatment effects at the end of the section. We write the observable propensity $e(X) = \mathbb{P}^{\text{Obs}}(T = 1 \mid X)$, where we assume $e(X) \in (0, 1)$ almost surely. The primary estimate is the IPW estimator that first estimates $e(X)$ and then estimates the APO as $E_{\mathbb{P}^{\text{Obs}}}[\lambda(R)Y]$, where $\lambda(R) = I_1/e(X)$. When unconfoundedness holds given the observed covariates so that $I_t \perp\!\!\!\perp Y(t) \mid X$, the IPW strategy consistently estimates the structural estimand $E_{\mathbb{P}^{\text{True}}}[Y(1)]$. However, we will only assume unconfoundedness holds if the researcher had access to both the observed and unobserved covariates: $I_t \perp\!\!\!\perp Y(t) \mid X, U$.[2] We write $e(X, U) := \mathbb{P}^{\text{True}}(T = 1 \mid X, U)$ for the (partially identified) propensity score that conditions on U.

Recall the definition of $\mathbb{P}^{\text{Target}}$ for the APO: draw R from the same distribution as $\mathbb{P}^{\text{Obs}}$, and draw $Y \mid R$ from $\mathbb{P}^{\text{Obs}}(Y \mid R)$ with probability $P(T \mid X)$ and from a certain reweighted distribution (Equation 1) with probability $1 - P(T \mid X)$. The distribution of R is the same under $\mathbb{P}^{\text{Target}}$ and $\mathbb{P}^{\text{Obs}}$ and the distribution $\mathbb{P}^{\text{Target}}$ satisfies $E_{\mathbb{P}^{\text{True}}}[Y(1)] = E_{\mathbb{P}^{\text{Target}}}[\lambda(R)Y]$, so that the structural estimand is $E_{\mathbb{P}^{\text{Target}}}[\lambda(R)Y]$. Notice that $\mathbb{P}^{\text{Target}}$ and $\mathbb{P}^{\text{Obs}}$ may have different distributions of $Y \mid X$: $\mathbb{P}^{\text{Obs}}(X, T, Y) = \mathbb{P}^{\text{Obs}}(X, T) d\mathbb{P}^{\text{Obs}}(Y \mid X, T)$ which differs from $\mathbb{P}^{\text{Target}}(X, T, Y)$ under unobserved confounding.

Many researchers study models that imply there are functions $\ell(X), u(X)$ such that

$$\ell(X) \leq \frac{e(X, U)/(1 - e(X, U))}{e(X)/(1 - e(X))} \leq u(X)$$

almost surely (Manski, 1990; Tan, 2006; Aronow and Lee, 2013; Masten and Poirier, 2018; Zhao et al., 2019; Dorn et al., 2024; Tan, 2024b; Frauen et al., 2023). These models imply pointwise restrictions on $\mathbb{P}^{\text{True}}(T = 1 \mid X, Y(1))$ and the Radon-Nikodym derivative $\frac{d\mathbb{P}^{\text{Target}}(Y|X,T=1)}{d\mathbb{P}^{\text{Obs}}(Y|X,T=1)}$. In particular, the selection assumption would imply the following almost sure bound:

$$\lambda(R) \frac{d\mathbb{P}^{\text{Target}}(Y \mid R)}{d\mathbb{P}^{\text{Obs}}(Y \mid R)} \in \left[\lambda(R)(e(X) + (1 - e(X))u(X)^{-1}), \quad \lambda(R)(e(X) + (1 - e(X))\ell(X)^{-1})\right].$$

Unconfoundedness corresponds to the special case $\ell(X) = u(X) = 1$. Manski bounds correspond to the special case $\ell(X) = 0$, $u(X) = \infty$. Tan (2006)'s Marginal Sensitivity Model

[2] This is without loss of generality by setting $U = \{Y(t)\}$.

(called such for $e(X)$ implicitly marginalizing over U) corresponds to $\underline{w}(X,t) = \mathbb{P}^{\text{Obs}}(T = t \mid X) + \mathbb{P}^{\text{Obs}}(T \neq t \mid X)\Lambda^{-1}$ and $\bar{w}(X,t) = \mathbb{P}^{\text{Obs}}(T = t \mid X) + \mathbb{P}^{\text{Obs}}(T \neq t \mid X)\Lambda$ with their sensitivity parameter Λ. Basit et al. (2023)'s risk ratio marginal sensitivity model, that restricts $e(X,U) \in \left[\Gamma_0^{-1}e(X), \Gamma_1^{-1}e(X)\right]$, corresponds to $\underline{w}(R) = \Gamma_1$ and $\bar{w}(R) = \Gamma_0$ when $e(X) \leq \Gamma_1$, where Γ_1 and Γ_0 are their sensitivity parameters. Masten and Poirier's conditional c-dependence model, that restricts $e(X,U) \in [e(X) - c, e(X) + c] \cap [0,1]$, corresponds to $\underline{w}(X,t) = \mathbb{P}^{\text{Obs}}(T = t \mid X) + \mathbb{P}^{\text{Obs}}(T \neq t \mid X)\frac{\max\{0,1-(e(X)+c)\}/\min\{1,e(X)+c\}}{(1-e(X))/e(X)}$ and $\bar{w}(X,t) = \mathbb{P}^{\text{Obs}}(T = t \mid X) + \mathbb{P}^{\text{Obs}}(T \neq t \mid X)\frac{\min\{1,1-(e(X)-c)\}/\max\{0,e(X)-c\}}{(1-e(X))/e(X)}$, where c is their sensitivity parameter.

Theorem 1 yields valid bounds on the identified set with the following quantities:

Proposition 1. *In the APO case, our method can be implemented on $E_{\mathbb{P}^{\text{True}}}[Y(1)]$ with:*

$$\lambda(R) = \frac{I_1}{e(X)}, \quad \underline{w}(R) = e(X) + (1 - e(X))u(X)^{-1}, \text{ and } \bar{w}(R) = e(X) + (1 - e(X))\ell(X)^{-1}.$$

The result is intuitive. The w bounds would be 1 under the unconfoundedness assumption $\ell(X) = u(X) = 1$. As $e(X)$ gets closer to one, we see a greater share of the treated potential outcomes and the w bounds grow closer to one. An analogous approach can be used to bound $E[Y(0)] = E[Y(1 - Z)/(1 - e(X, Y(0)))]$.

The result in Proposition 1 extends the analysis of Tan (2024a); Frauen et al. (2023) to allow $\ell(X)$ to be arbitrarily small or equal to zero and $u(X)$ to be equal to one. As a result, it includes Masten and Poirier (2018)'s conditional c-dependence assumption so long as $\lambda(R)Y$ is integrable under the target distribution. This characterization of bounds under conditional c-dependence is also simpler than Masten and Poirier (2018)'s characterization, which involves an integral over a transformation of the full quantile regression function, and by Lemma 3, our characterization enables estimation of valid bounds under appropriate regularity conditions and consistency of the propensity score alone. An interesting question for future work is whether Masten et al. (2024)'s proposed estimator for conditional c-dependence possesses similar validity guarantees.

The resulting bounds are also sharp for the average treatment effect (ATE). It is not difficult to verify that the bounds are sharp for each APO separately, because the bounds $\underline{w}$ and $\bar{w}$ correspond to propensity scores between 0 and 1. One may also ask whether the separate APO bounds can be achieved by the same unobserved confounder, for example the potential outcomes. An extension of Dorn and Guo (2023)'s sharpness argument shows that in fact the separate bounds can be achieved at the same time, so that these bounds also yield sharp bounds for the ATE.

3 Novel Applications of Framework

In this section, we illustrate the wide applicability of our framework by applying it to several potential failures of identifying assumptions: RD with manipulation, DD without randomization or parallel trends, and IV without exogeneity or exclusion. We show that the sharp statistical bounds under our approach contain all restrictions implied by the underlying structural model for APO selection and RD manipulation, but not IV exclusion failure.

3.1 Regression Discontinuity

In this section, we introduce a novel sensitivity analysis for sharp RD designs. We explain this design in more detail: we bound standard structural estimands that previous work could not meaningfully quantify, and the extension to other linear estimators is relatively straightforward. We explain the setup, describe the estimands, then propose our sensitivity assumption and show how our framework applies. Finally, we show that our bounds are sharp for our structural model.

We work under a slight modification of Gerard et al. (2020)'s model. There is a full distribution $\mathbb{P}^{\text{True}}$ over $(M, X(1), X(0), X, Y(1), Y(0), T, T(0))$, where M is a variable corresponding to manipulation status, $X(m)$ is a potential running variable corresponding to manipulation status $M = m$, $Y(t)$ is a potential outcome corresponding to treatment status $T = t$, $T \in \{0, 1\}$ is the treatment status, and $T(0)$ is the potential treatment status for $M = 0$. We face the fundamental problem of causal inference and only observe the coarsening $\mathbb{P}^{\text{Obs}}$ over $(X = X(M), Y = Y(T), T)$, so $R = (X, T)$ and $\xi = (M, X(1), X(0), T(0))$ in this context.

We study sharp RD designs. We assume there is a cutoff c such that $\mathbb{P}^{\text{True}}(T = 1 \mid X > c) = 1$, $\mathbb{P}^{\text{True}}(T = 1 \mid X < c) = 0$, which mimics assumption (RD) of Hahn et al. (2001).[3] It is observationally testable and often obvious in applications: if X is the net reported vote share of an election candidate, then the candidate wins if and only if $X > 0$. We use the

[3] Gerard et al. (2020) extend their approach to fuzzy RD, where there are complex shape restrictions.

notation "$X = c$" to mean that we take the limit as $\varepsilon \to 0$ of $X \in [c-\varepsilon, c+\varepsilon]$ for bandwidth ε. Similarly, we let $X = c^+$ denote the limit corresponding to $X \in (c, c+\varepsilon]$ and let $X = c^-$ correspond to $X \in [c-\varepsilon, c)$. We use this notation to mimic the notation in Gerard et al. (2020) so that the results are immediately comparable. We operate in the bandwidth near the threshold, so indicators such as $1\{T = 0\}$ implicitly takes value zero for observations outside of the bandwidth.

We assume that observations can be partitioned into manipulators and non-manipulators. We assume manipulation is one-sided, and without loss of generality assume manipulators choose treatment ($F_{X|M=1}(c) = 0$). As in Gerard et al., the probability of treated observations being manipulated,

$$\eta = \mathbb{P}^{\text{True}}(M = 1 \mid X = c^+),$$

is identified.[4] We assume appropriate continuity of potential outcomes and running variables given the manipulation status, and discuss other regularity conditions in Assumption RD. The assumptions imply that non-manipulator average treatment effects would be identified by the change in non-manipulator outcomes across the cutoff c. However, when $\eta > 0$ so that there is manipulation, the distribution of $Y \mid X = c^+$ includes manipulators' treated potential outcomes, so that standard treatment effects are not point-identified.

We may be interested in the conditional average treatment effect (CATE), the conditional local average treatment effect (CLATE), and the conditional average treatment effect on the treated (CATT):

$$\begin{aligned}
\psi^{CATE} &\equiv E\left[Y(1) - Y(0) \mid X = c\right] \\
\psi^{CLATE} &\equiv E\left[Y(1) - Y(0) \mid X = c, M = 0\right] \\
\psi^{CATT} &\equiv E\left[Y(1) - Y(0) \mid X = c^+\right].
\end{aligned} \tag{5}$$

$E\left[Y(1) - Y(0) \mid X = c, M = 0\right]$ is called the CLATE because it is an average treatment effect for the subgroup of potentially-assigned "compliers" who receive treatment if and only if their X_i is above the cutoff (Gerard et al., 2020).

We write the structural estimands of interest in terms of conditional expectations of observed variables and potential outcomes as follows:

[4] Gerard et al. call this quantity τ. Without manipulation, the density of X should be smooth at the cutoff, so this object is identified from the jump in the density of X at the cutoff.

Lemma 4. *Our main estimands of interest have the following expressions:*

$$\begin{aligned}
\psi_{CATE} &= \frac{1}{2-\eta} E\left[Y \mid X = c^+\right] + \frac{1-\eta}{2-\eta} E_{\mathbb{P}^{\mathrm{True}}}\left[Y(1) \mid X = c, M = 0\right] \\
&\quad - \frac{\eta}{2-\eta} E_{\mathbb{P}^{\mathrm{True}}}\left[Y(0) \mid X = c, M = 1\right] - \left(1 - \frac{\eta}{2-\eta}\right) E\left[Y \mid X = c^-\right] \\
\psi_{CLATE} &= E_{\mathbb{P}^{\mathrm{True}}}[Y(1) \mid X = c, M = 0] - E[Y \mid X = c^-] \\
\psi_{CATT} &= \frac{E\left[(2T-1)Y \mid X = c\right] - \frac{\eta}{2-\eta} E_{\mathbb{P}^{\mathrm{True}}}[Y(0) \mid X = c, M = 1]}{\mathbb{P}^{\mathrm{Obs}}(X = c^+ \mid X = c)}.
\end{aligned}$$

Most quantities in the expressions above are point-identified. The only unidentified objects are the conditional expectations $E_{\mathbb{P}^{\mathrm{True}}}\left[Y(1) \mid X = c, M = 0\right]$ and $E_{\mathbb{P}^{\mathrm{True}}}\left[Y(0) \mid X = c, M = 1\right]$.

The specific target distribution $\mathbb{P}^{\mathrm{Target}}$ will depend on the estimand of interest as follows:[5]

$$\begin{aligned}
d\mathbb{P}^{\mathrm{Target}}_{CATE}(Y \mid X = c, T = t) &= d\mathbb{P}^{\mathrm{True}}(Y(t) \mid X = c) \\
d\mathbb{P}^{\mathrm{Target}}_{CLATE}(Y \mid X = c, T = t) &= t d\mathbb{P}^{\mathrm{True}}(Y(1) \mid X = c, M = 0) + (1-t) d\mathbb{P}^{\mathrm{Obs}}(Y \mid X = c^-) \\
d\mathbb{P}^{\mathrm{Target}}_{CATT}(Y \mid X = c, T = t) &= t d\mathbb{P}^{\mathrm{Obs}}(Y \mid X = c^+) + (1-t) d\mathbb{P}^{\mathrm{True}}(Y(0) \mid X = c^-).
\end{aligned}$$

Then the target distribution for estimated $E \in \{CATE, CLATE, CATT\}$ is defined as $d\mathbb{P}^{\mathrm{Target}}(X, T, Y) \equiv \mathbb{P}^{\mathrm{Obs}}(X, T) d\mathbb{P}^{\mathrm{Target}}_{E}(Y \mid X, T)$. The distribution of R is the same under $\mathbb{P}^{\mathrm{Target}}$ and $\mathbb{P}^{\mathrm{Obs}}$ and $E_{\mathbb{P}^{\mathrm{Target}}}[\lambda(R)Y]$ is equal to the target estimand for $\lambda(R)$ we define below.

We work under a relatively simple restriction on manipulation. Suppose we have a structural model that implies there are values $\Lambda_0, \Lambda_1 \geq 1$ such that for $t = 1, 0$:

$$\frac{\mathbb{P}^{\mathrm{True}}(M = 1 \mid Y(t), X = c)}{\mathbb{P}^{\mathrm{True}}(M = 0 \mid Y(t), X = c)} \bigg/ \frac{\mathbb{P}^{\mathrm{True}}(M = 1 \mid X = c)}{\mathbb{P}^{\mathrm{True}}(M = 0 \mid X = c)} \in [\Lambda_t^{-1}, \Lambda_t]. \tag{6}$$

Define $\mathcal{M}(\Lambda_0, \Lambda_1)$ as the set of distributions $\mathbb{Q}$ over $(M, X(1), X(0), Y(1), Y(0), T)$ that marginalize to the distribution of $(X(M), Y(T), T = 1\{X(M) > c\})$ under $\mathbb{P}^{\mathrm{True}}$, satisfy the restrictions of this section, and satisfy $\frac{\mathbb{Q}(M=1|Y(t),X=c)}{\mathbb{Q}(M=0|Y(t),X=c)} / \frac{\mathbb{P}^{\mathrm{True}}(M=1|X=c)}{\mathbb{P}^{\mathrm{True}}(M=0|X=c)} \in [\Lambda_t^{-1}, \Lambda_t]$ almost surely.[6]

This functional form (and hence its interpretation) is inspired by Tan (2006)'s Marginal

[5] A formal construction would define the appropriate distributions at $X = x$ based on the distribution conditional on $|x - c|$ and then take the limit as $x \to c$ from either side, which would yield well-defined distributions by the maintained continuity assumptions.

[6] To be precise, $\mathbb{P}^{\mathrm{True}}(M = 1 \mid X = c) = \mathbb{P}^{\mathrm{True}}(M = 1 \mid X = c^+) \mathbb{P}^{\mathrm{Obs}}(X = c^+ \mid X = c) = \mathbb{P}^{\mathrm{True}}(M = 1 \mid X = c^+) \lim_{\varepsilon \to 0^+} \mathbb{P}^{\mathrm{Obs}}(T = 1 \mid |X - c| \leq \varepsilon)$, which is well-defined by Assumption RD.

Sensitivity Model, so the sensitivity parameter can be interpreted as the largest variation in the odds ratio of manipulation given any potential outcome compared to the average. $\Lambda_t = 1$ corresponds to an unconfounded case in which the difference in regression values yields the CATE. $\Lambda_t = \infty$ corresponds to Gerard et al. (2020)'s assumption, in which the distribution of $Y(0) \mid M = 1, X = c$ is only constrained by the support of the potential outcome. In between, larger values of Λ accommodate larger degrees of manipulation on potential outcomes.

Proposition 2. *In the RD case, our method can be implemented for the CATE $E_{\mathbb{P}^{\text{True}}}[Y(1) - Y(0) \mid X = c]$ with:*

$$\begin{aligned}
\lambda(R) &= \frac{1\{T=1\}}{\mathbb{P}^{\text{Obs}}(X = c^+ \mid X = c)} - \frac{1\{T=0\}}{\mathbb{P}^{\text{Obs}}(X = c^- \mid X = c)}\\
\underline{w}(R) &= 1\{T=1\}\left[\frac{1}{2-\eta} + \frac{1-\eta}{2-\eta}\frac{1}{1-\eta+\eta\Lambda_1}\right] + 1\{T=0\}\left[\left(1 - \frac{\eta}{2-\eta}\right) + \frac{\eta}{2-\eta}\Lambda_0^{-1}\right]\\
\bar{w}(R) &= 1\{T=1\}\left[\frac{1}{2-\eta} + \frac{1-\eta}{2-\eta}\frac{1}{1-\eta+\eta\Lambda_1^{-1}}\right] + 1\{T=0\}\left[\left(1 - \frac{\eta}{2-\eta}\right) + \frac{\eta}{2-\eta}\Lambda_0\right];
\end{aligned}$$

our method can be implemented for the CLATE $E_{\mathbb{P}^{\text{True}}}[Y(1) \mid X = c, M = 0] - E_{\mathbb{P}^{\text{Obs}}}[Y \mid X = c^-]$ with the same $\lambda(R)$ and:

$$\underline{w}(R) = \frac{1\{T=1\}}{1-\eta+\eta\Lambda_1} + 1\{T=0\}, \qquad \bar{w}(R) = \frac{1\{T=1\}}{1-\eta+\eta\Lambda_1^{-1}} + 1\{T=0\};$$

our method can be implemented for the CATT $E_{\mathbb{P}^{\text{True}}}[Y(1) - Y(0) \mid X = c^+]$ with the same $\lambda(R)$ and:

$$\begin{aligned}
\underline{w}(R) &= 1\{T=1\} + 1\{T=0\}\left((1-\eta) + \eta\Lambda_0^{-1}\right),\\
\bar{w}(R) &= 1\{T=1\} + 1\{T=0\}\left((1-\eta) + \eta\Lambda_0\right).
\end{aligned}$$

The connection between Proposition 2 and the characterization from Lemma 4 comes from correspondences that we derive in Appendix Lemma 8. Our CLATE bounds for $\Lambda_1 = \infty$ are finite and identical to the bounds in Gerard et al. (2020) for sharp RD.

When Λ_1 is finite, we are also able to obtain meaningful CATE bounds.

Remark 1. CATE bounds require some restriction on the untreated potential outcomes among the always-treated ($M = 1$) population. As a result, we need some assumption to narrow these bounds beyond range restrictions. Equation (6) narrows these bounds by bounding the probability of manipulation conditional on potential outcomes, allowing us to learn about the potential outcomes of the $M = 1$ observations from the untreated potential outcomes of the $M = 0$ observations.

Our analysis so far bounds manipulation on each potential outcome separately. There turns out to be no additional information available from a structural model that bounds manipulation on both potential outcomes simultaneously.

Proposition 3. *Suppose there is a finite $\Lambda \geq 1$ such that $\Lambda_1 = \Lambda_0 = \Lambda$. Let $\mathcal{M}'(\Lambda)$ be the set of distributions $\mathbb{Q} \in \mathcal{M}(\infty, \infty)$ satisfying:*

$$\frac{\mathbb{Q}(M = 1 \mid Y(1), Y(0), X = c)}{\mathbb{Q}(M = 0 \mid Y(1), Y(0), X = c)} \Big/ \frac{\mathbb{P}^{\text{True}}(M = 1 \mid X = c)}{\mathbb{P}^{\text{True}}(M = 0 \mid X = c)} \in [\Lambda^{-1}, \Lambda]. \tag{7}$$

Then $\mathcal{M}'(\Lambda) \subseteq \mathcal{M}(\Lambda, \Lambda)$. Further, take any distributions $\mathbb{Q}_1, \mathbb{Q}_0 \in \mathcal{M}(\Lambda, \Lambda)$ and write $\psi_t = E_{\mathbb{Q}_t}[Y(t) \mid X = c, M = 1 - t]$. Then there is a distribution $\mathbb{Q}' \in \mathcal{M}'(\Lambda)$ satisfying $E_{\mathbb{Q}'}[Y(t) \mid X = c, M = 1 - t] = \psi_t$ for $t = 1, 0$.

Proposition 3 shows that the bounds are sharp, i.e., separate bounds on both potential outcomes suffice to bound our structural estimands of interest. The proposition states that a distribution that satisfies the conditions of the model can be constructed to achieve the bounds. The quantities $E_{\mathbb{P}^{\text{True}}}[Y(t) \mid X = c, M = 1-t]$ still identify our structural estimands of interest.

The high-level logic of Proposition 3 is fundamentally similar to Section 2.3. Any given pair of Radon-Nikodym derivatives may not be simultaneously achievable for both potential outcomes. However, the pair corresponding to the worst-case bounds are simultaneously achievable. As a result, the identified set is the same whether we bound manipulation on one potential outcome or both potential outcomes simultaneously.

3.2 Differences in Differences

We explain how our framework applies to difference in differences (DD) designs with selected treatments, focusing on a 2-group-2-period structure for simplicity. In this context, the estimand is the average treatment effect on the treated (ATT). With $Y_t(d)$ denoting the potential outcome at time $t = 0, 1$ with treatment status $d = 0, 1$, D the treatment status, T the time period, and X covariates, there is a full distribution $\mathbb{P}^{\text{True}}$ over $(D, T, Y_0(0), Y_0(1), Y_1(0), Y_1(1), X)$, but we only observe the coarsening $\mathbb{P}^{\text{Obs}}$ over $R = (D, T, Y_0, Y_1, X)$, and ξ is empty.

The object of interest is $E_{\mathbb{P}^{\text{True}}}[Y_1(1) - Y_1(0) \mid D = 1]$, and the primary estimand is $E_{\mathbb{P}^{\text{Obs}}}[Y_1 - Y_0 \mid D = 1] - E_{\mathbb{P}^{\text{Obs}}}[Y_1 - Y_0 \mid D = 0]$. Two prominent identification assumptions have been used in the literature: randomization and parallel trends. Our approach is more directly applicable to sensitivity analysis to violations of randomization. We discuss our method's application to parallel trends later.

When randomization (Athey and Imbens, 2022) holds, both the structural estimand and the primary estimand are also the ATE. Randomization is the assumption that $D \perp\!\!\!\perp Y_t(d) \mid X, T$. This problem is analogous to the IPW environment. If randomization holds, $E_{\mathbb{P}^{\text{True}}}[Y_1(1) - Y_1(0) \mid D = 1] = E_{\mathbb{P}^{\text{True}}}[Y_1(1) - Y_1(0)]$, so our structural estimand is the ATE in the post period:

$$ATE = E_{\mathbb{P}^{\text{True}}}[Y_1(1) - Y_1(0)].$$

We let $Y = Y_1 - Y_0$ denote the difference in observed outcomes in the two periods. The DD estimand is $E_{\mathbb{P}^{\text{Obs}}}[\lambda(R)Y]$, for $\lambda(R) = \frac{D}{\mathbb{P}^{\text{Obs}}(D=1|X)} - \frac{1-D}{\mathbb{P}^{\text{Obs}}(D=0|X)}$. If randomization holds, then $E_{\mathbb{P}^{\text{Obs}}}[Y_1(0) - Y_0(0) \mid D = 0] = E_{\mathbb{P}^{\text{Obs}}}[Y_1(0) - Y_0(0) \mid D = 1]$, which can be interpreted as parallel trends. We allow for violations of randomization in the form of:

$$\ell_d(X) \leq \frac{e(X,U)/(1 - e(X,U))}{e(X)/(1 - e(X))} \leq u_d(X)$$

where $e(X) = \mathbb{P}^{\text{Obs}}(D = 1 \mid X)$ and $e(X,U) = \mathbb{P}^{\text{True}}(D = 1 \mid X, U)$, where $U = (Y_1(1) - Y_0(0), Y_1(0) - Y_0(0))$. Then, our method can be implemented with:

$$\begin{aligned}
\lambda(R) &= \frac{D}{e(X)} - \frac{1 - D}{1 - e(X)}, \\
\underline{w}(R) &= D\left(e(X) + \frac{1 - e(X)}{u_1(X)}\right) + (1 - D)\left(1 - e(X) + e(X) l_0(X)\right) \\
\bar{w}(R) &= D\left(e(X) + \frac{1 - e(X)}{l_1(X)}\right) + (1 - D)\left(1 - e(X) + e(X) u_0(X)\right) \\
\frac{d\mathbb{P}^{\text{Target}}(Y \mid R)}{d\mathbb{P}^{\text{Obs}}(Y \mid R)} &= \frac{De(X)}{e(X,U)} + \frac{(1 - D)(1 - e(X))}{1 - e(X,U)}
\end{aligned}$$

With this reweighting, the DD estimand recovers $E_{\mathbb{P}^{\text{True}}}[Y_1(1) - Y_0(0)] - E_{\mathbb{P}^{\text{True}}}[Y_1(0) - Y_0(0)] = E_{\mathbb{P}^{\text{True}}}[Y_1(1) - Y_1(0)]$. Due to their analogy to IPW, these bounds are sharp due to the argument in Dorn and Guo (2023).

Remark 2. The analysis here is applicable to selection on parallel trends. The typical parallel trends assumption is some statement like, $E_{\mathbb{P}^{\text{Obs}}}[Y_1(0) - Y_0(0) \mid D = 1] = E_{\mathbb{P}^{\text{Obs}}}[Y_1(0) - Y_0(0) \mid D = 0]$, i.e., the difference in the untreated potential outcome is the same for both treated and untreated groups (e.g., Callaway and Sant'Anna (2021) Assumption 4). Roth and Sant'Anna (2023) show that this statement is functional form-independent only under strong randomization-type assumption. Our approach can therefore be interpreted as a sensitivity analysis when parallel trends is interpreted as a functional form-independent restriction. Alternatively, when the representation of the outcome variable is important to identification, our analysis also carries through for the ATT under a bound on the degree of selection on the trends $Y_1(0) - Y_0(0)$. Other approaches to sensitivity analysis for difference-in-difference designs include bounding the difference in trends (Rambachan and Roth, 2023) or distributional distances (Bertsimas et al., 2022) that can be interpreted as f-divergence restriction on selection in our framework (Jin et al., 2022). Parallel trends can also be microfounded (Marx et al., 2024) and the bias of the DD estimand can also be characterized when parallel trends fails (Ghanem et al., 2022). These constructions can inform the sensitivity parameter used, but the exact mapping is beyond the scope of this paper.

3.3 Instrumental Variables

In this subsection, we consider the canonical context of instrumental variables to qualify the advantages and limitations of our framework. When considering sensitivity analysis to failure of exogeneity, our framework can be applied in a straightforward manner. However, there are also some limitations of our framework when considering other IV assumptions: our framework can be applied to sensitivity analysis to exclusion failure, but our bounds are not sharp; and our framework is not applicable to monotonicity failure.

We assume there is a distribution $\mathbb{P}^{\text{True}}$ over $\left(Z, \{T(z)\}_{z\in\{0,1\}}, T, \{Y(t,z)\}_{t,z\in\{0,1\}}, X\right)$, where Z is a binary instrument, $Y(t,z)$ is the potential outcome with binary treatment status t and instrument status z, and $T(z)$ is the potential treatment given the instrument z is $T(z)$. However, we only observe the coarsening $\mathbb{P}^{\text{Obs}}$ over $(Z, T = T(z), Y = Y(T, Z), X)$.

We will maintain that monotonicity holds ($T(1) \geq T(0)$). Under monotonicity, there are three treatment response groups: always-takers (At, $T(1) = T(0) = 1$), never-takers (Nt, $T(1) = T(0) = 0$), and compliers (Co, $1 = T(1) > T(0) = 0$). The regressors are $R = (T, Z, X)$ and $\xi = \left(\{T(z)\}_{z\in\{0,1\}}\right)$.

We first consider sensitivity analysis to failure of exogeneity. Suppose exclusion holds, so $Y(T, z) = Y(T)$ for all z, and we target the treatment effect on compliers. Then, $\psi = E_{\mathbb{P}^{\text{True}}}\left[\frac{1\{Co\}}{\mathbb{P}^{\text{Obs}}(Co)}(Y(1) - Y(0))\right]$. With slight abuse of notation, we let $e(X) = \mathbb{P}^{\text{Obs}}(Z = 1 \mid X), e(X, U) = \mathbb{P}^{\text{True}}(Z = 1 \mid X, U)$, where $U = (Y(0), Y(1), T(0), T(1))$ denotes the individual unobservable as before. Suppose we allow exogeneity to fail for compliers, but require exogeneity to nonetheless hold for always takers and never takers, i.e., $(Y(1), Y(0)) \perp\!\!\!\perp Z \mid X, T(1), T(0) = T(1)$.[7] Then, it can be shown that:[8]

$$
\begin{aligned}
\psi &= E_{\mathbb{P}^{\text{True}}}\left[\frac{Z1\{Co\}Y}{e(X,U)} - \frac{(1-Z)1\{Co\}Y}{1-e(X,U)}\right] \\
&= E_{\mathbb{P}^{\text{True}}}\left[\frac{ZY}{e(X,U)} - \frac{(1-Z)Y}{1-e(X,U)}\right] = E_{\mathbb{P}^{\text{Target}}}\left[\lambda(R)Y\right]
\end{aligned}
$$

for $\lambda(R) = \frac{Z - \mathbb{P}^{\text{Obs}}(Z=1|X)}{\mathbb{P}^{\text{Obs}}(Z=1|X)\mathbb{P}^{\text{Obs}}(Z=0|X)}$. Then, the problem becomes entirely analogous to IPW, so the form of $\underline{w}, \bar{w}$ is identical, and we have sharp bounds.

Turning to sensitivity analysis to failure of exclusion, now suppose that the instrument is randomly assigned ($Y(t, z) \perp\!\!\!\perp Z \mid X$). The exclusion restriction is$Y(t, 1) = Y(t, 0)$. If exclusion fails, then the standard conditional IV estimand, $(E[Y \mid Z = 1, X] - E[Y \mid Z = 0, X])/(E[T \mid Z = 1, X] - E[T \mid Z = 0, X])$, will incorrectly assign any direct effect of the instrument on outcomes to treatment effects. We target an instrument-weighted average

[7] This assumption is reasonable in most empirical settings. Suppose we are interested in the impact of participation in 401(k) tax-deferred savings plans (T) on additional savings (Y), and we instrument for T with eligibility for 401(k) (Z) based on employment. Never-takers have no incentive to manipulate their eligibility because they would not take the plan up even if they were offered. However, compliers who really want the plan may manipulate their eligibility by choosing an occupation appropriately.

[8] We use the results that:

$$
\begin{aligned}
E_{\mathbb{P}^{\text{True}}}\left[\frac{ZY}{e(X,U)}\right] &= E_{\mathbb{P}^{\text{True}}}\left[\frac{Z1\{Co\}Y}{e(X,U)}\right] + E_{\mathbb{P}^{\text{True}}}\left[\frac{(1-Z)TY}{e(X)}\right] + E_{\mathbb{P}^{\text{True}}}\left[\frac{Z(1-T)Y}{e(X)}\right] \\
E_{\mathbb{P}^{\text{True}}}\left[\frac{(1-Z)Y}{1-e(X,U)}\right] &= E_{\mathbb{P}^{\text{True}}}\left[\frac{(1-Z)1\{Co\}Y}{1-e(X,U)}\right] + E_{\mathbb{P}^{\text{True}}}\left[\frac{(1-Z)TY}{e(X)}\right] + E_{\mathbb{P}^{\text{True}}}\left[\frac{Z(1-T)Y}{e(X)}\right]
\end{aligned}
$$

LATE given by:

$$\psi := E_{\mathbb{P}^{\text{True}}}\left[\eta(X)1\{Co\}\sum_z \omega(z \mid X)\left(Y(1,z) - Y(0,z)\right)\right]$$

where $\eta(X)$ reflects covariate weighting and $\omega(1 \mid X) = 1 - \omega(0 \mid X)$ is a researcher-estimated instrument weighting function in $[0,1]$. When exclusion holds, this class includes the average complier treatment effect for $\eta(X) = 1/\mathbb{P}^{\text{True}}(Co)$. When exclusion fails, this specification also allows the researcher to target a particular weighted average of treatment effects across instrument statuses. For example, $\omega(Z \mid X) = 1$ targets the treatment effect at $Z = 1$ and $\omega(Z \mid X) = E[Z \mid X]$ targets the treatment effect at the average observed instrument status.

We rewrite the structural estimand ψ in terms of conditional expected outcomes as follows:

$$\psi = E_{\mathbb{P}^{\text{True}}}\left[\lambda(R)E_{\mathbb{P}^{\text{True}}}\left[\sum_z \omega(z \mid X)Y(T,z) \mid X, T(1), T(0)\right]\right],$$

where $\lambda(Z,X,T) = \eta(X)\left(\frac{Z - \mathbb{P}^{\text{Obs}}(Z=1|X)}{\mathbb{P}^{\text{Obs}}(Z=1|X)\mathbb{P}^{\text{Obs}}(Z=0|X)}\right)$. Always-takers and never-takers are included in this statement of ψ for convenience, but their potential outcomes cancel out for $\lambda(R)$ because $E_{\mathbb{P}^{\text{True}}}[\lambda(R) \mid T(1), T(0), X] = 0$.

The target distribution $\mathbb{P}^{\text{Target}}$ is constructed as a marginal distribution. For $v \in \{0,1\}$, define the distribution $\mathbb{P}^{\text{Target}}$, which re-draws a value v from a Bernoulli($\omega(1 \mid X)$) distribution in order to achieve appropriate outcome weighting, as follows:

$$\mathbb{P}^{\text{Target}}(X,T,Z,Y) \equiv \sum_{v,t_1,t_0 \in \{0,1\}} \mathbb{P}^{\text{True}}(X,t,Z,T(1)=t_1,T(0)=t_0)\omega(v \mid X)\mathbb{P}^{\text{True}}(Y(T,v) \mid X, T(1), T(0)).$$

The distribution of R is the same under $\mathbb{P}^{\text{Target}}$ and $\mathbb{P}^{\text{Obs}}$. Further, by Abadie (2003)'s argument, the distribution satisfies $E_{\mathbb{P}^{\text{Target}}}[\lambda(R)Y] = E_{\mathbb{P}^{\text{True}}}[\lambda(R)E_{\mathbb{P}^{\text{True}}}[Y(T,Z) \mid X, Co]] = \psi$, so that $E_{\mathbb{P}^{\text{Target}}}[\lambda(R)Y]$ is the structural estimand.

Suppose we have a structural model, where, for $z \in \{0,1\}$,

$$\ell(X) \leq \frac{d\mathbb{P}^{\text{True}}(Y(T(z), 1-z) \mid \{T(z)\}, X)}{d\mathbb{P}^{\text{True}}(Y(T(z), z) \mid \{T(z)\}, X)} \leq u(X), \tag{8}$$

and suppose further that all associated Radon-Nikodym derivatives are finite and strictly positive. The exclusion restriction $Y(t,1) = Y(t,0)$ implies $\ell(x) = u(x) = 1$. Worst-case bounds that only restrict the support of the potential outcomes correspond to $\ell(x) = 0$, $u(x) = \infty$. When Y is binary, Ramsahai (2012) proposes a sensitivity analysis that places a bound on $\mathbb{P}^{\text{True}}(Y = 1 \mid X, T, Z = 1, U) - \mathbb{P}^{\text{True}}(Y = 1 \mid X, T, Z = 0, U)$ for some unobserved U, which immediately translates to bounds on always- and never-taker likelihood ratios by taking $U = \{T(1), T(0)\}$ and implies bounds on complier likelihood ratios. This object can be decomposed as a convolution of several of the probability objects above, so its interpretation is less transparent in a causal framework with potential treatments. Alternatively, Λ bounds on the effect of Z on the odds of a binary potential outcome given X and potential treatments imply Λ bounds on the odds ratios here, though narrower partially identified regions could be obtained by leveraging estimated regression functions.

As we show in Appendix Lemma 5, these objects further imply bounds on $\frac{d\mathbb{P}^{\text{Target}}(Y|X,T=t,Z=z)}{d\mathbb{P}^{\text{Obs}}(Y|X,T=t,Z=z)}$ for each value of T and Z.

Proposition 4. *Suppose we write the implied bounds from the worst-case ℓ and u applied to the formulas from Appendix Lemma 5 as*

$$\underline{w}(t,z|X) \leq \frac{d\mathbb{P}^{\text{Target}}(Y \mid X, T = t, Z = z)}{d\mathbb{P}^{\text{Obs}}(Y \mid X, T = t, Z = z)} \leq \bar{w}(t,z|X),$$

where we write $\underline{w}(t,z|X) = \bar{w}(t,z|X) = 1$ for values such that $\mathbb{P}^{\text{Obs}}(T = t, Z = z \mid X) = 0$. Then our method can be implemented on $E_{\mathbb{P}^{\text{True}}}[\eta(X)1\{Co\}\sum_z \omega(z \mid X)(Y(1,z) - Y(0,z))]$ with the following values:

$$\lambda(R) = \eta(X)\frac{Z - \mathbb{P}^{\text{Obs}}(Z = 1 \mid X)}{\mathbb{P}^{\text{Obs}}(Z = 1 \mid X)\mathbb{P}^{\text{Obs}}(Z = 0 \mid X)}, \quad \underline{w}(R) = \underline{w}(T, Z \mid X), \quad \bar{w}(R) = \bar{w}(T, Z \mid X).$$

Unlike the previous examples, the characterization in Proposition 4 is conservative.

Proposition 5. *Suppose the observed distribution follows $X = 1$; $Z \mid X \sim Bern(0.5)$; $T \mid Z, X \sim Bern(Z/2)$; and $Y \mid X, Z, T \sim Unif(-1,1)$. Suppose we are interested in the average complier treatment effect at $z = 1$, i.e. $\eta(x) = 2$ and $\omega(z \mid x) = z$. Suppose a structural model implies lower bounds of $\ell(x) = 1$ and $u(x) = \infty$. Then the structural model implies the sharp bounds are the singleton $\{0\}$, but the bounds from Proposition 4 are $[-1,1]$.*

Intuitively, the structural model implicitly includes cross-restrictions between the Co, At, and Nt Radon-Nikodym derivatives. In this case, the $\frac{d\mathbb{P}^{\text{Target}}(Y|T=0,Z=0)}{d\mathbb{P}^{\text{Obs}}(Y|T=0,Z=0)}$ includes a product of Radon-Nikodym derivatives for compliers and never-takers, which must satisfy further constraints that our approach omits for the purpose of ease of use.[9]

Remark 3. Our procedure can be applied somewhat trivially to linear projections using Ordinary Least Squares (OLS). Consider a hypothetical linear model $Y = X\beta + u$. Suppose we are interested in a linear combination of coefficients $\delta'\beta$, where X includes an intercept but δ puts no weight on the intercept term, so that without loss of generality we can assume $E_{\mathbb{P}^{\text{True}}}[u] = 0$. However, suppose there may be endogeneity in the sense that $E[Xu] \neq 0$. Such problems are considered in Cinelli and Hazlett (2020). If we targeted a distribution $\mathbb{P}^{\text{Target}}$ that first sampled $X \sim \mathbb{P}^{\text{Obs}}$, then drew $u \mid X$ from the distribution of u under $\mathbb{P}^{\text{True}}$, and then returned $Y = X\beta + u$, then the target distribution would obtain the correct coefficients. The sensitivity assumption is then on $\frac{d\mathbb{P}^{\text{Target}}}{d\mathbb{P}^{\text{Obs}}}$. For instance, we may have $\underline{w} \leq \frac{d\mathbb{P}^{\text{Target}}}{d\mathbb{P}^{\text{Obs}}} \leq \bar{w}$. Using the notation of our general framework, $\lambda(R) = \delta' E[X'X]^{-1}X$. An open question is whether there is a microfounded interpretation of such an assumption.

Remark 4. Due to our interpretation of the sensitivity parameter as the Radon-Nikodym derivative of the conditional distribution of Y, one case that is not easily covered by our framework is when the sensitivity assumption does not concern Y directly. In the context of instrumental variables, some sensitivity analysis to violations of the monotonicity assumption that $T(1) \geq T(0)$ makes assumptions on the mass of defiers with $T(1) < T(0)$ that do not condition on potential outcomes (e.g., Noack (2021); Yap (2025)), which do not map to Y directly.

4 Empirical Applications

We apply plug-in estimates of the partially identified bounds in two empirical settings that use sharp regression discontinuity designs. Since the focus of this paper is on identification, we relegate our estimation and inference strategy to Appendix B; the approach is largely similar to that of Gerard et al. (2020). In the first application on incumbency advantage, our exercise shows how CLATE and CATE can deliver different bounds once we allow for

[9] Non-sharpness is attributed to how the sensitivity parameter maps to $\frac{d\mathbb{P}^{\text{Target}}(Y|X,T,Z)}{d\mathbb{P}^{\text{Obs}}(Y|X,T,Z)}$. If we use $\frac{d\mathbb{P}^{\text{Target}}(Y|X,T,Z)}{d\mathbb{P}^{\text{Obs}}(Y|X,T,Z)}$ as the sensitivity parameter, for instance, we will automatically get sharp bounds, though $\frac{d\mathbb{P}^{\text{Target}}(Y|X,T,Z)}{d\mathbb{P}^{\text{Obs}}(Y|X,T,Z)}$ is not interpretable.

manipulators. In the second application on the effects of council size, we show how, even if we target the CLATE, sensitivity analysis allows us to obtain bounds that are far more informative than worst-case bounds.

4.1 Incumbency Advantage

We illustrate our procedure using the dataset from Caughey and Sekhon (2011). We study incumbency advantage in Congressional elections by using an RD design to estimate the impact of the Democratic party winning in the current period (T) on the probability of the the democratic party winning in the following election (Y). The running variable is the marginal share of votes in the current period (X), normalized such that crossing zero results in a win.

In this dataset, a Democratic win in marginal elections (defined as within a margin of 5 percentage points, which is the smallest window reported in Lee (2008)) is associated with a 48 percentage point increase in the probability of a democratic win in the next election. This estimate is somewhat larger than the 35 percentage point estimate in Lee (2008) that uses a different dataset and a larger bandwidth, but in line with his 45 percentage point estimate for the same candidate winning in the next election.[10]

In this sharp RD design, manipulation occurs only in one direction as candidates aim to be elected. As documented by Caughey and Sekhon (2011), there is a discontinuity in the density of the running variable at the cutoff (see their Figure 1). We estimate the density on either side of the cutoff to obtain $\hat{\eta} = 0.11$ in our specification of interest. This sorting is attributed to activities before election day rather than post election, as Caughey and Sekhon (2011) document that incumbents raise more funds for their campaigns, which is positively correlated with election outcomes. Caughey and Sekhon (2011) estimate Rosenbaum bounds in terms of treatment probabilities that fail to exploit the nature of identification failing via manipulation. In light of manipulation, we implement our procedure to calculate bounds on the CATE, CLATE, and CATT, which are reported in Figure 1. In contrast, the bounds of Gerard et al. (2020) do not characterize the CATE. We use $\Lambda = \Lambda_1 = \Lambda_0$, and a linear quantile regression to obtain the quantiles.

While the results are largely similar across the three estimands, it is notable that the estimated bounds for CLATE is higher than that of CATE and CATT. This observation

[10]See Section 3.4 in Lee (2008).

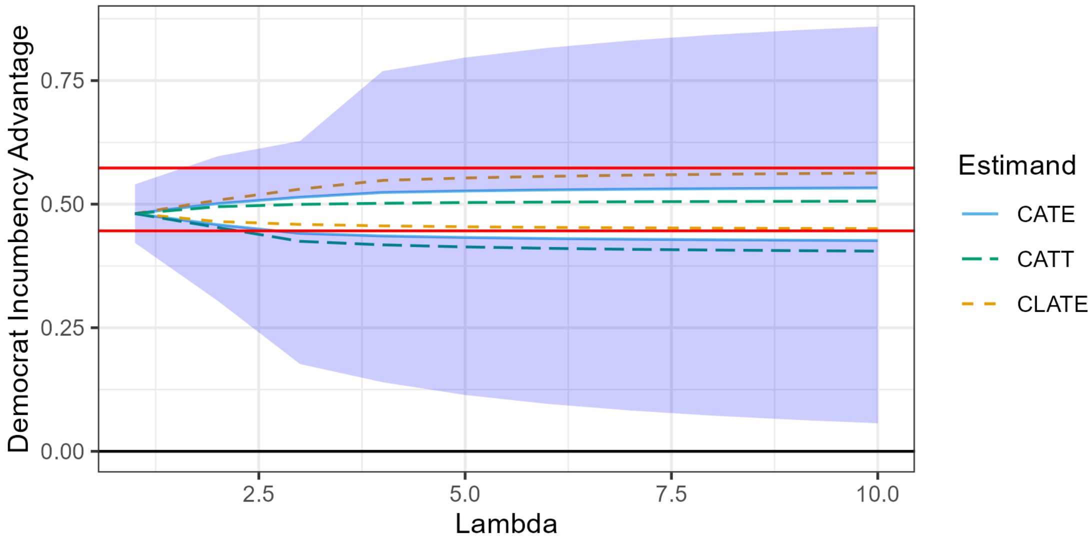

Figure 1: Bounds on CATE, CLATE, and CATT all exclude zero (solid line). The upper and lower bounds of the shaded region above and below the CATE estimate indicates the bounds of the estimated 95% confidence set. The horizontal red lines denote the worst-case GRR bounds for CLATE. $n = 856$.

suggests that non-manipulators are more likely to win in the next election than manipulators, which would go against the intuition that manipulators in the current election are also able to manipulate in the following election. Accounting for the difference in the treatment effects of various subgroups is beyond the scope of this paper, but this result shows that distinguishing CLATE from CATE sheds light on heterogeneity in treatment effects across different subgroups of the population.

We find that the incumbency advantage is not very sensitive to Λ. Our estimated confidence sets easily exclude treatment effects of zero for even large Λ. Our results that the conclusion is robust are simple to implement and align with the setting: our limiting bound estimates when $\Lambda \to \infty$ are just slightly wider than the worst-case Manski bounds.[11]

[11] Using a back-of-the-envelope calculation, exploiting the fact that $\eta = 0.11$, the proportion of manipulators is about 0.06. Further, $E[Y \mid X = c^-] = E[Y \mid M = 0, X = c^-] = 1/4$ and $E[Y \mid X = c^+] = P(M = 0 \mid X = c^+)E[Y \mid M = 0, X = c^+] + P(M = 1 \mid X = c^+)E[Y \mid M = 1, X = c^+] = .73$ so that $E[Y \mid M = 0, X = c^+] = \frac{E[Y|X=c^+]-P(M=1|X=c^+)E[Y|M=1,X=c^+]}{P(M=0|X=c^+)} \geq \frac{.73-0.11}{0.89}$. Using the expression for CATE and the worst-case bounds with $E[Y \mid M = 1, X = c^-] = E[Y \mid M = 1, X = c^+] = 1$,

$$\begin{aligned} CATE = {} & P(M = 0)\left(E[Y \mid M = 0, X = c^+] - E[Y \mid M = 0, X = c^-]\right) \\ & + P(M = 1)(E[Y \mid M = 1, X = c^+] - E[Y \mid M = 1, X = c^-]) \geq 0.42. \end{aligned}$$

As we take Λ large, our estimated lower bound converges to 0.42, even though we do not explicitly impose bounded outcomes. This finding suggests that even without exploiting all available information, our

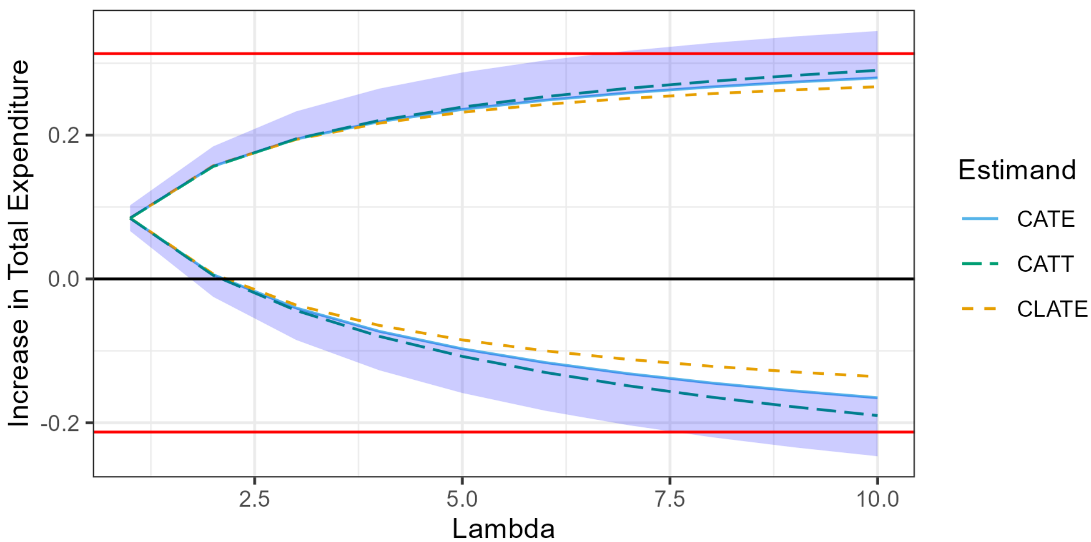

Figure 2: The upper and lower bounds of the shaded region above and below the CATE estimate indicates the bounds of the estimated 95% confidence set. The horizontal red lines denote the worst-case GRR bounds for CLATE. $n = 44,276$.

4.2 Effects of Council Size

Egger and Koethenbuerger (2010) were interested in the effect of council size on total expenditure in Bavaria. They apply a sharp RD design as the council size is a deterministic (step-wise) function of the population size of the municipality (see their Table 1). In their main specification (their Table 3), they studied the effect of crossing the population threshold (T) on the log of total expenditure (Y), using the log population (X) as the running variable. Using their replication data, we estimate that crossing the threshold increases expenditure increases by 8% for the 30% window, which is smaller than the authors' estimates as we do not rely on a polynomial approximation.[12]

Manipulation has been documented in this context by Eggers et al. (2018), and it can occur for several reasons in this context, including selective precision (e.g., ordering extra checks to ensure new arrivals are properly processed), and strategic recruitment (e.g., encouraging friends to change their address to the municipality). The manipulation also occurs in one direction, as mayors have little incentive to have smaller council sizes. With our given bandwidth, we estimate $\hat{\eta} = 0.15$.

procedure yields reasonably tight bounds.

[12] Our method compares the difference in means above and below the threshold so that it is immediately comparable to our benchmark in Gerard et al. (2020).

In this application, the bounds of the three estimands are fairly similar. If we were to use the GRR bounds for the CLATE by default, then we would automatically overturn the results of the study as the bounds are extremely uninformative. However, our sensitivity analysis provides a more nuanced perspective on the significance of the effects: at small Λ values, the effect on total expenditure remains significant and positive. In this context, small Λ values may be defensible if the higher expenditures (Y) do not translate directly to, say, the mayors' salaries which would incentivize the manipulation: if we believe that the odds ratio of manipulation for any given expenditure after crossing the threshold is not more than 2, then the estimated lower bounds for CLATE is still above 0. Alternatively, for the estimated CLATE bound to be lower than 0, we would need manipulation odds ratio of at least 2, which may be perceived as rather extreme. Our confidence intervals are also shorter there than in the previous application as we have many more observations.

5 Conclusion

This paper proposes a novel sensitivity analysis framework for identification failures for linear estimators. By placing bounds on the distributional distance between the observed distribution and a target distribution that identifies the structural parameter of interest, we obtain sharp and tractable analytic bounds. This framework generalizes existing sensitivity models in RD and IPW and motivates a new sensitivity model for IV exclusion failures. We provide new results on sharp and valid sensitivity analysis that allow even unbounded likelihood ratios. We illustrate how our framework and partial identification results contribute to three important applications, including new procedures for sensitivity analysis for the CATE under RD with manipulation and for instrumental variables with exclusion.

A Additional Assumptions and Results

Assumption RD. We extend Gerard et al.'s continuity condition and assume that there is a well-defined conditional distribution function $\mathbb{P}^{\text{True}}(Y(1), Y(0) \mid X, M)$ such that for each y_1, y_0, $\mathbb{P}^{\text{True}}(Y(1) \leq y_1, Y(0) \leq y_0 \mid X = x, M = 0)$ can be defined as continuous in x and that the derivative of $\mathbb{P}^{\text{True}}(X \leq x \mid M = 0)$ is continuous at $x = c$. We further assume that $\mathbb{P}^{\text{True}}(Y(1) \leq y_1, Y(0) \leq y_0 \mid X = x, M = 1)$ is right-continuous x, assume that

$\mathbb{P}^{\text{True}}(M = 1 \mid X = x)$ is right-continuous at $x = c$, assume that marginals $\mathbb{P}^{\text{True}}(Y(t) \leq y_t \mid X = x, M = m)$ are right-continuous for $m = 1$ (continuous for $m = 0$) at $X = c$ and define conditional distributions at $X = c$, $X = c^+$, and $X = c^-$ using the appropriate limits, such as $\mathbb{P}^{\text{Obs}}(X = c^+ \mid X = c) = \lim_{\varepsilon \to 0^+} \mathbb{P}^{\text{Obs}}(T = 1 \mid |X - c| \leq \varepsilon)$.

The following claim shows that the conditional c-dependence identified set in our implementation example is finite and bounded for all $c \leq 0.1$ but is infinite for all $c > 0.1$.

Proposition 6. *Suppose $Y \mid X, Z \sim \mathcal{N}(\mu(X,Z), \sigma(X,Z)^2)$, the support of the observed propensity function $e(X)$ is the closed interval $[\eta_1, 1-\eta_2] \subset (0,1)$, and the conditional outcome variance $\sigma(X,Z)$ is positive and bounded. Then there is a finite $B > 0$ such that the ATE identified set is a subset of $[E[\mu(X,1) - \mu(X,0)] - B, E[\mu(X,1) - \mu(X,0)] + B]$ for all $c < \min\{\eta_1, \eta_2\}$ but is $(-\infty, \infty)$ for all $c > \min\{\eta_1, \eta_2\}$.*

For exclusion failure, we consider a more general model than the one stated in the main text. Suppose we have a structural model that implies there are functions $\ell_{Nt}(X), u_{Nt}(X)$, $\ell_{At}(X), u_{At}(X)$, $\ell^1_{Co}(X), u^1_{Co}(X)$, $\ell^0_{Co}(X), u^0_{Co}(X)$ such that:

$$\ell_{Nt}(X) \leq \frac{d\mathbb{P}^{\text{True}}(Y(0,0) \mid Nt, X)}{d\mathbb{P}^{\text{True}}(Y(0,1) \mid Nt, X)} \leq u_{Nt}(X), \qquad \ell_{At}(X) \leq \frac{d\mathbb{P}^{\text{True}}(Y(1,1) \mid At, X)}{d\mathbb{P}^{\text{True}}(Y(1,0) \mid At, X)} \leq u_{At}(X),$$
$$\ell^1_{Co}(X) \leq \frac{d\mathbb{P}^{\text{True}}(Y(1,0) \mid Co, X)}{d\mathbb{P}^{\text{True}}(Y(1,1) \mid Co, X)} \leq u^1_{Co}(X), \qquad \ell^0_{Co}(X) \leq \frac{d\mathbb{P}^{\text{True}}(Y(0,1) \mid Co, X)}{d\mathbb{P}^{\text{True}}(Y(0,0) \mid Co, X)} \leq u^0_{Co}(X).$$

The assumption stated in the main text is a special case of this setting where $\ell_{Nt}(X) = \ell_{At}(X) = \ell^1_{Co}(X) = \ell^0_{Co}(X) = \ell(X)$ and $u_{Nt}(X) = u_{At}(X) = u^1_{Co}(X) = u^0_{Co}(X) = e(X)$. The following lemma shows how $\frac{d\mathbb{P}^{\text{Target}}(Y|X,T=t,Z=z)}{d\mathbb{P}^{\text{Obs}}(Y|X,T=t,Z=z)}$ can be written as objects that we bound above.

Lemma 5. *In the Instrumental Variables application in Section 3.3, the following decompositions hold:*

$$\frac{d\mathbb{P}^{\text{Target}}(Y|X, T=0, Z=0)}{d\mathbb{P}^{\text{Obs}}(Y|X, T=0, Z=0)} = \omega(0 \mid X) + \omega(1 \mid X) \frac{\mathbb{P}^{\text{Obs}}(Nt \mid X)}{\mathbb{P}^{\text{Obs}}(Co \mid X) + \mathbb{P}^{\text{Obs}}(Nt \mid X)} \frac{d\mathbb{P}^{\text{Obs}}(Y \mid X, T=0, Z=1)}{d\mathbb{P}^{\text{Obs}}(Y \mid X, T=0, Z=0)}$$
$$+ \left\{ \omega(1 \mid X) \frac{d\mathbb{P}^{\text{True}}(Y(0,1)|X, Co)}{d\mathbb{P}^{\text{True}}(Y(0,0)|X, Co)} \right.$$
$$\left. \times \left(1 - \frac{\mathbb{P}^{\text{Obs}}(Nt \mid X)}{\mathbb{P}^{\text{Obs}}(Co \mid X) + \mathbb{P}^{\text{Obs}}(Nt \mid X)} \frac{d\mathbb{P}^{\text{True}}(Y(0,0)|X, Nt)}{d\mathbb{P}^{\text{True}}(Y(0,1)|X, Nt)} \frac{d\mathbb{P}^{\text{Obs}}(Y \mid X, T=0, Z=1)}{d\mathbb{P}^{\text{Obs}}(Y \mid X, T=0, Z=0)}\right) \right\}$$
$$\frac{d\mathbb{P}^{\text{Target}}(Y|X, T=1, Z=1)}{d\mathbb{P}^{\text{Obs}}(Y|X, T=1, Z=1)} = \omega(1 \mid X) + \omega(0 \mid X) \frac{\mathbb{P}^{\text{Obs}}(At \mid X)}{\mathbb{P}^{\text{Obs}}(Co \mid X) + \mathbb{P}^{\text{Obs}}(At \mid X)} \frac{d\mathbb{P}^{\text{Obs}}(Y \mid X, T=1, Z=0)}{d\mathbb{P}^{\text{Obs}}(Y \mid X, T=1, Z=1)}$$

$$+\left\{\omega(0 \mid X)\frac{d\mathbb{P}^{\text{True}}(Y(1,0)|X,Co)}{d\mathbb{P}^{\text{True}}(Y(1,1)|X,Co)}\right.$$
$$\left.\times\left(1-\frac{\mathbb{P}^{\text{Obs}}(At \mid X)}{\mathbb{P}^{\text{Obs}}(Co \mid X)+\mathbb{P}^{\text{Obs}}(At \mid X)}\frac{d\mathbb{P}^{\text{True}}(Y(1,1) \mid X,At)}{d\mathbb{P}^{\text{True}}(Y(1,0) \mid X,At)}\frac{d\mathbb{P}^{\text{Obs}}(Y \mid X,T=1,Z=0)}{d\mathbb{P}^{\text{Obs}}(Y \mid X,T=1,Z=1)}\right)\right\}$$
$$\frac{d\mathbb{P}^{\text{Target}}(Y \mid X,T=0,Z=1)}{d\mathbb{P}^{\text{Obs}}(Y \mid X,T=0,Z=1)}=\omega(1 \mid X)+\omega(0 \mid X)\frac{d\mathbb{P}^{\text{True}}(Y(0,0) \mid X,Nt)}{d\mathbb{P}^{\text{True}}(Y(0,1) \mid X,Nt)}$$
$$\frac{d\mathbb{P}^{\text{Target}}(Y \mid X,T=1,Z=0)}{d\mathbb{P}^{\text{Obs}}(Y \mid X,T=1,Z=0)}=\omega(0 \mid X)+\omega(1 \mid X)\frac{d\mathbb{P}^{\text{True}}(Y(1,1) \mid X,At)}{d\mathbb{P}^{\text{True}}(Y(1,0) \mid X,At)},$$

where for values of X with $\mathbb{P}^{\text{Obs}}(T,Z \mid X)=0$, we write $\frac{d\mathbb{P}^{\text{Target}}(Y|X,T,Z)}{d\mathbb{P}^{\text{Obs}}(Y|X,T,Z)}=1$.

B Estimation, Robustness, and Inference

This subsection shows that natural plug-in estimators are robust under reasonable conditions. We focus on the upper bound for exposition. The target object is:

$$T \equiv E_{\mathbb{P}^{\text{Obs}}}\left[\lambda(R)Y+\left(\lambda(R)Y-Q_{\tau(R)}\left(\lambda(R)Y|R\right)\right)a(\underline{w}(R),\bar{w}(R),\lambda(R)Y,Q_{\tau(R)}(\lambda(R)Y \mid R))\right],$$

where

$$a(\underline{w},\bar{w},\lambda y,q) \equiv (\bar{w}-\underline{w})\,1\{\lambda y>q\}-(1-\underline{w}).$$

This is a statistical quantity that depends on the observed distribution $\mathbb{P}^{\text{Obs}}$ alone, so we now suppress the dependence on $\mathbb{P}^{\text{Obs}}$ for concision.

We use hatted objects to denote the estimated objects and $\hat{E}$ to denote the sample mean.

$$\hat{T} \equiv \hat{E}\left[\hat{\lambda}(R)Y+\left(\hat{\lambda}(R)Y-\hat{Q}_{\hat{\tau}(R)}\left(\hat{\lambda}(R)Y|R\right)\right)a(\hat{\underline{w}}(R),\hat{\bar{w}}(R),\hat{\lambda}(R)Y,\hat{Q}_{\hat{\tau}(R)}(\hat{\lambda}(R)Y \mid R))\right]$$

Consistency follows under natural conditions. Further, we have a form of one-sided robustness.

Proposition 7. *If we have iid sampling, finite second moments, and $\hat{\lambda} \xrightarrow{p} \lambda, \hat{Q} \xrightarrow{p} Q, \hat{\bar{w}} \xrightarrow{p} \bar{w}, \hat{\underline{w}} \xrightarrow{p} \underline{w}$ uniformly, then $\hat{T} \xrightarrow{p} T$. Further, even if $\hat{Q} \xrightarrow{p} \bar{Q} \neq Q$, there exists some $\hat{T}^*$ such that $\hat{T} \geq \hat{T}^*$ and $\hat{T}^* \xrightarrow{p} T$.*

The first part of the proposition is immediate by applying the continuous mapping theorem and the law of large numbers for iid observations. The $a(\cdot)$ dependence on an indicator function only introduces a kink point rather than a discontinuity, so the function is still continuous. The assumptions of the proposition are made at a high level, so that we can accommodate various forms of consistent estimators for functions such as $\hat{\hat{w}}$. In particular, we can accommodate machine learning nuisance estimators, as we do in our simulation. Note that our consistency assumption may not be achievable when the quantiles are finite but unbounded, as at one point in our simulation.

The second part of the proposition states a useful robustness guarantee. Even if the quantile is not estimated correctly, the resulting estimated bounds will be valid: too wide for the estimated sensitivity model rather than too narrow. This validity property corresponds to Lemma 3's one-sided validity guarantee.

For inference in the general model, we use a standard percentile bootstrap. Namely,

1. For every $b = 1, \cdots, B$,
   (a) Sample n observations from the data iid with replacement to get the bootstrap data.
   (b) Calculate $t^{(b)} \equiv \sqrt{n}(\hat{T}^{(b)} - \hat{T})$, where $\hat{T}^{(b)}$ is the estimator that uses the bootstrapped data.
2. Let G_N denote the CDF of $t^{(b)}$ and q_α denote that α quantile of G_N. For a size $1-\alpha$ confidence interval, use $[\hat{T} - \frac{1}{\sqrt{n}} q_{1-\alpha/2}, \hat{T} - \frac{1}{\sqrt{n}} q_{\alpha/2}] = CI(\alpha)$

The bootstrap will have coverage at least as large as nominal under standard conditions, like smoothness of the $\hat{w}$ estimates, even if the quantile estimator tends to an inconsistent limit. The core argument is that an infeasible bootstrap estimator that replaces the estimated $1\{\hat{\lambda}^{(b)} Y > \hat{Q}^+\}$ with the true $1\{\lambda Y > Q^+\}$ in the construction of a would be valid and have weakly more aggressive confidence intervals. As a result, the quantiles do not even need to be reestimated in the bootstraps (Dorn and Guo, 2023). Further, estimation error in the quantiles exhibit a second-order influence on the estimated bounds, so that the confidence intervals can asymptotically achieve the nominal rate under moderate conditions (Dorn et al., 2024). A generic proof of bootstrap consistency is outside the scope of this paper. The presence of extreme quantiles or kink points, as in Masten and Poirier (2018)'s conditional c-dependence model, may call for more exotic bootstraps and a subtle proof of bootstrap validity.

In our empirical applications to regression discontinuity, we mimic the inference procedure from Gerard et al. (2020) that "tilts" the confidence set. The main complication is that

$\hat{\eta}$ is an estimated object, and there is potential non-normality when the true η is close to zero. Due to the potential for case-by-case nonnormality that would call for a more sophisticated bootstrap (Fang and Santos, 2019), we do not pursue showing validity of any general bootstrap procedure.

For $\kappa_n = \log(n)^{1/2}$, $\tilde{\eta} = 1 - \hat{f}^-/\hat{f}^+$, and $\hat{\eta} = \max\{\tilde{\eta}, 0\}$, where $\hat{f}^-$ and $\hat{f}^+$ are the estimated densities of the running variable just below and above the threshold (which can be calculated using standard nonparametric methods),

1. Generate bootstrap samples $\{Y_{i,b}, T_{i,b}, X_{i,b}\}_{i=1}^n$, $b = 1, \cdots, B$ by sampling with replacement from the original data $\{Y_i, T_i, X_i\}_{i=1}^n$, for some large integer B.

2. Calculate $\tilde{\eta}_b^* = 1 - \hat{f}_b^-/\hat{f}_b^+$, and put $\hat{\sigma}_{\tilde{\eta}}$ as the sample standard deviation of $\{\tilde{\eta}_b^*\}_{b=1}^B$.

3. Calculate $\tilde{\eta}_b = \tilde{\eta}_b^* - \tilde{\eta} + \max\{\hat{\eta}, \kappa_n \hat{\sigma}_{\tilde{\eta}}\}$ and $\hat{\eta}_b = \max\{\tilde{\eta}_b, 0\}$.

4. For $j \in \{+, -\}$, calculate $\hat{\psi}^j(\hat{\eta}_b)$ using the redefined estimate $\hat{\eta}_b$ from the previous step, and put $\hat{\sigma}^j$ as the sample standard deviation of $\{\hat{\psi}^j(\hat{\eta}_b)\}_{b=1}^B$.

We now define $\hat{\psi}^{-*}$ and $\hat{\psi}^{+*}$ in the same way as $\hat{\psi}^-$ and $\hat{\psi}^+$, except that we use $\hat{\eta}^* = \max\{\tilde{\eta}, \kappa_n \hat{\sigma}_{\tilde{\eta}}\}$ instead of $\hat{\eta}$. The confidence set is

$$\mathcal{C}_{1-\alpha}^{SRD} = \left[\hat{\Gamma}^{L*} - r_\alpha \cdot \hat{\sigma}^L, \hat{\Gamma}^{L*} + r_\alpha \cdot \hat{\sigma}^U\right]$$

where r_α is the value that solves the equation:

$$\Phi\left(r_\alpha + \frac{\hat{\Gamma}^{U*} - \hat{\Gamma}^{L*}}{\max\{\hat{\sigma}^L, \hat{\sigma}^U\}}\right) - \Phi\left(-r_\alpha\right) = 1 - \alpha.$$

C Main Proofs

C.1 Proofs for Section 2

The following lemma is used in the proof of Theorem 1 and results from simple algebra.

Lemma 6. *Recall the definitions $\tau(R) = \frac{\bar{w}(R)-1}{\bar{w}(R)-\underline{w}(R)}$ and $a(\underline{w}, \bar{w}, s) = (\bar{w} - \underline{w})1\{s > 0\} - (1 - \underline{w})$. We may instead write:*

$$a(\underline{w}, \bar{w}, s) = (1 - \underline{w})\left(\frac{1\{s > 0\}}{1 - \tau} - 1\right).$$

Proof of Theorem 1. Formally define $\psi^+(R)$ as a random variable satisfying:

$$\psi^+(R) = \sup_W E_{\mathbb{P}^{\text{Obs}}}[\underline{w}(R)\lambda(R)Y + (1-\underline{w}(R))W\lambda(R)Y \mid R]$$
$$\text{s.t. } W \in [0, 1-\tau(R)] \text{ and } E_{\mathbb{P}^{\text{Obs}}}[W \mid R] = 1 \text{ a.s.}$$

for the given R. By the argument in the proof sketch, we can write the upper bound as:

$$\psi^+ = \sup_W E_{\mathbb{P}^{\text{Obs}}}[W\lambda(R)Y] \quad \text{s.t.} \quad W \in [\underline{w}(R), \bar{w}(R)] \text{ and } E_{PObs}[W \mid R] = 1 \text{ a.s.}$$
$$\psi^+ = \sup_W E_{\mathbb{P}^{\text{Obs}}}[\underline{w}(R)\lambda(R)Y + W\lambda(R)Y] \text{ s.t. } W \in [0, \bar{w}(R) - \underline{w}(R)] \text{ and } E_{PObs}[W \mid R] = 1 - \underline{w}(R) \text{ a.s.}$$
$$\psi^+ = \sup_W E_{\mathbb{P}^{\text{Obs}}}[\underline{w}(R)\lambda(R)Y + (1-\underline{w}(R))W\lambda(R)Y] \text{ s.t. } W \in \left[0, \underbrace{\frac{\bar{w}(R) - \underline{w}(R)}{1-\underline{w}(R)}}_{1-\tau(R)}\right] \text{ and } E_{PObs}[W \mid R] = 1 \text{ a.s.}$$
$$= E_{\mathbb{P}^{\text{Obs}}}[\psi^+(R)].$$

As in the proof sketch, we claim that $\psi^+(R) = \phi^+(R)$ almost surely. By arguments from the DRO literature, e.g. Dorn et al. (2024), $\psi^+(R) = \underline{w}(R)E[\lambda(R)Y \mid R] + (1 - \underline{w}(R))CVaR^+_{\tau(R)}$ almost surely. (In the case that $\tau(R) = 1$, $\sup_W E_{\mathbb{P}^{\text{Obs}}}[W\lambda(R)Y]$ s.t. $E_{\mathbb{P}^{\text{Obs}}}[W \mid R] = 1$ simply reduces to setting $\lambda(R)Y$ to the supremum of the support of $\lambda(R)Y$, i.e. the level-1 CVaR.)

Note that the level-$\tau(R)$ CVaR of $\lambda(R)Y \mid R$ can be defined as $E\left[Q^+ + \frac{\{Y-Q^+\}_+}{1-\tau(R)} \mid R\right]$. (In the case $\tau(R) = 1$, we evaluate this term as $Q^+ = F_1(\lambda(R)Y \mid R)$.)

By Lemma 6, we may write:

$$\begin{aligned}
\phi^+(R) &= \lambda(R)Y + (1-\underline{w}(R))(\lambda(R)Y - Q^+(R))\left(\frac{1\{\lambda(R)Y > Q^+(R)\}}{1-\tau(R)} - 1\right) \\
&= \underline{w}(R)\lambda(R)Y + (1-\underline{w}(R))\left(Q^+(R) + \frac{\{\lambda(R)Y - Q^+(R)\}_+}{1-\tau(R)}\right) \\
&= \underline{w}(R)\lambda(R)Y + (1-\underline{w}(R))CVaR^+_{\tau(R)} \text{ a.s.}
\end{aligned}$$

Completing the proof. The derivation for ϕ^- is analogous. □

Proof of Lemma 1. This is a standard result for importance sampling. □

Proof of Lemma 2. Let $\underline{\tau}(R)$ and $\bar{\tau}(R)$ be random variables satisfying $\underline{\tau}(R) = \mathbb{P}^{\text{Obs}}(\lambda(R)Y < Q+(R) \mid R)$ and $\bar{\tau}(R) = \mathbb{P}^{\text{Obs}}(\lambda(R)Y \leq Q+(R) \mid R)$ almost surely. Note that $\mathbb{P}^{\text{Obs}}(\lambda(R)Y =$

$Q^+(R) \mid R) = \bar{\tau}(R) - \underline{\tau}(R)$ almost surely.

Define the function $\gamma : \mathbb{R}^d \to \mathbb{R}^1$ as follows:

$$\gamma(r) \equiv \begin{cases} \frac{1-\underline{\tau}(r)\underline{w}(r)-(1-\bar{\tau}(r))\bar{w}(r)}{\bar{\tau}(r)-\underline{\tau}(r)} & \text{if} \quad \bar{\tau}(r) > \underline{\tau}(r) \\ \underline{w}(r) & \text{else.} \end{cases}$$

It is clear that $\gamma(r)$ is well-defined. The result that $\gamma(R) \in [\underline{w}(R), \bar{w}(R)]$ almost surely and that the resulting W^*_{sup} solves the upper bound optimization problem follows by an adapted version of Dorn and Guo (2023)'s Proposition 2 argument and the observation that $E[W^*_{sup} \mid R] = 1$ almost surely. □

Proof of Lemma 3. Recall from the proof of Theorem 1 that we may write:

$$\psi^+ = \underline{w}(R)\lambda(R)Y + (1 - \underline{w}(R)) \left(Q^+(R) + \frac{\{\lambda(R)Y - Q^+(R)\}_+}{1 - \tau(R)} \right),$$

where $\{x\}_+ = \max\{x, 0\}$ and where Q^+ is the quantile regression function.

We make the stronger claim that:

$$Q^+(R) \in \arg\min_{\bar{Q}} E\left[\underline{w}(R)\lambda(R)Y + (1 - \underline{w}(R)) \left(\bar{Q}(R) + \frac{\{\lambda(R)Y - \bar{Q}(R)\}_+}{1 - \tau(R)} \right) \mid R \right] \text{ a.s.}$$

If $H \equiv 1\{\tau(R) < 1\} = 0$ and hence $\tau(R) = 1$, then the right hand side should be understood to be infinite if $\lambda(R)Y > Q^+(R)$. Therefore $Q^+(R) = F_1(\lambda(R)Y \mid R)$ is the right-hand side minimizer on the event $H = 0$. Therefore:

$$Q^+(R) \in \arg\min_{\bar{Q}} (1 - H)E\left[\underline{w}(R)\lambda(R)Y + (1 - \underline{w}(R)) \left(\bar{Q}(R) + \frac{\{\lambda(R)Y - \bar{Q}(R)\}_+}{1 - \tau(R)} \right) \mid R \right] \text{ a.s.}$$

On the event $H = 1$, we can write the quantile regression function Q^+ as some function satisfying the weighted quantile regression definition:

$$\begin{aligned} Q^+(R) &\in \arg\min_{\bar{Q}} E_{\mathbb{P}^{\text{Obs}}} \left[(1 - \tau(R))^{-1} \left(\tau(R)\{\lambda(R)Y - \bar{Q}(R)\}_+ + (1 - \tau(R))\{\bar{Q}(R) - \lambda(R)Y\}_+ \right) \right] \\ &\Leftrightarrow \in \arg\min_{\bar{Q}} E_{\mathbb{P}^{\text{Obs}}} \left[(1 - \tau(R))^{-1} \left(\{\lambda(R)Y - \bar{Q}(R)\}_+ + (1 - \tau(R))(\bar{Q}(R) - \lambda(R)Y) \right) \right] \\ &= E_{\mathbb{P}^{\text{Obs}}} \left[\bar{Q}(R) - \lambda(R)Y + \frac{\lambda(R)Y - \bar{Q}(R)\}_+}{1 - \tau(R)} \right]. \end{aligned}$$

The $\lambda(R)Y$ term does not affect the minimizer, so that:

$$Q^{+}(R) \in \arg\min_{\bar{Q}} HE\left[\underline{w}(R)\lambda(R)Y + (1-\underline{w}(R))\left(\bar{Q}(R) + \frac{\{\lambda(R)Y - \bar{Q}(R)\}_{+}}{1-\tau(R)}\right) \mid R\right] \text{ a.s.}$$

Combining terms, we obtain:

$$Q^{+}(R) \in \arg\min_{\bar{Q}} E\left[\underline{w}(R)\lambda(R)Y + (1-\underline{w}(R))\left(\bar{Q}(R) + \frac{\{\lambda(R)Y - \bar{Q}(R)\}_{+}}{1-\tau(R)}\right) \mid R\right] \text{ a.s.}$$

Completing the proof. $\square$

Proof of Proposition 1. This is immediate by the discussion in Section 2.3. $\square$

C.2 Proofs for Section 3

The following lemmas are used for the proofs of Lemma 4 and Proposition 2 in Section 3.

Lemma 7. *Consider the Regression Discontinuity application in Section 3.1. Suppose $F_{X|M=0}(x)$ is differentiable in x at c with a positive derivative. Then $\mathbb{P}^{\text{True}}(X = c^{+} \mid X = c, M = 0) = 1/2$.*

Lemma 8. *Consider the Regression Discontinuity application in Section 3.1. Suppose $\mathbb{Q} \in \mathcal{M}(\infty,\infty)$. Further suppose that the distribution of $(Y(1), Y(0), T(0), M) \mid X = c$ under $\mathbb{Q}$ has associated manipulation selection functions $q_1(y_1) \equiv \mathbb{Q}(M = 1|Y(1) = y_1, X = c)$ and $q_0(y_0) \equiv \mathbb{Q}(M = 1|Y(0) = y_0, X = c)$. Then the Radon-Nikodym derivatives are as follows:*

$$\begin{aligned}
\frac{d\mathbb{Q}(Y(1) \mid X = c, M = 0)}{d\mathbb{P}^{\text{Obs}}(Y \mid X = c^{+})} &= \frac{1}{1-\eta}\frac{1-q_1(Y(1))}{1+q_1(Y(1))} \\
\frac{d\mathbb{Q}(Y(0) \mid X = c, M = 1)}{d\mathbb{P}^{\text{Obs}}(Y \mid X = c^{-})} &= \frac{2(1-\eta)}{\eta}\frac{q_0(Y(0))}{1-q_0(Y(0))}.
\end{aligned}$$

As a result:

$$E_{\mathbb{P}^{\text{True}}}\left[Y(1) \mid X = c, M = 0\right] = \frac{1}{1-\eta}E_{\mathbb{P}^{\text{Obs}}}\left[Y\frac{1-q_1(Y)}{1+q_1(Y)}\frac{1\{X = c^{+}\}}{\mathbb{P}^{\text{Obs}}(X = c^{+})}\right] \tag{9}$$

$$E_{\mathbb{P}^{\text{True}}}\left[Y(0)|X = c, M = 1\right] = \frac{2(1-\eta)}{\eta}E_{\mathbb{P}^{\text{Obs}}}\left[Y\frac{q_0(Y)}{1-q_0(Y)}\frac{1\{X = c^{-}\}}{\mathbb{P}^{\text{Obs}}(X = c^{-})}\right]. \tag{10}$$

Proof of Lemma 4. For our proofs for the Regression Discontinuity application of Section 3.1, it is useful to use $\eta_0 \equiv \mathbb{P}^{\text{True}}(M = 1 \mid X = c)$.

For the CATE, first observe that $\mathbb{P}^{\text{True}}(X = c^- \mid X = c) = \mathbb{P}^{\text{Obs}}(X = c^- \mid X = c) = (1 - \eta_0)/2$. To see this, $\mathbb{P}^{\text{True}}(X = c^-|X = c) = \mathbb{P}^{\text{True}}(X = c^-|M = 0, X = c)\mathbb{P}^{\text{True}}(M = 0|X = c) = \mathbb{P}^{\text{True}}(X = c^-|M = 0, X = c)(1 - \eta_0) = (1 - \eta_0)/2$, due to Lemma 7. Consequently, $\mathbb{P}^{\text{True}}(X = c^+|X = c) = (1 + \eta_0)/2$, so

$$\begin{aligned}
\psi_{CATE} \equiv\ & E[Y(1) - Y(0)|X = c] \\
&= E[Y(1)|X = c^+, X = c]P(X = c^+|X = c) + E[Y(1)|X = c^-, X = c]P(X = c^-|X = c) \\
&\quad - E[Y(0)|X = c, M = 1]P(M = 1|X = c) - P(M = 0|X = c)E[Y(0)|X = c, M = 0] \\
&= E[Y(1)|X = c^+, X = c]\frac{1 + \eta_0}{2} + E[Y(1)|X = c^-, X = c]\frac{1 - \eta_0}{2} \\
&\quad - E[Y(0)|X = c, M = 1]\eta_0 - (1 - \eta_0)E[Y(0)|X = c, M = 0].
\end{aligned}$$

The final equality uses the definition that $\eta_0 = \mathbb{P}^{\text{True}}(M = 1|X = c)$ and the result that $\mathbb{P}^{\text{True}}(X = c^-|X = c) = (1 - \eta_0)/2$. Further, $E[Y|X = c^+] = E[Y(1)|X = c^+]$ and $E[Y|X = c^-] = E[Y(0)|X = c^-]$ due to the sharp RD design of Assumption 1. To obtain the expression in the lemma, observe that $\eta_0 = \mathbb{P}^{\text{True}}(M = 1 \mid X = c^+)\mathbb{P}^{\text{True}}(X = c^+ \mid X = c) = \eta(1 + \eta_0)/2$ as $\mathbb{P}^{\text{True}}(M = 1 \mid X = c^-) = 0$. Then, we can substitute $\eta_0 = \eta/(2 - \eta)$.

For the CATT, observe:

$$\begin{aligned}
\psi_{CATT} &= \frac{E[TY \mid X = c] - E[TY(0) \mid X = c]}{E[T \mid X = c]} \\
&= \frac{E[TY \mid X = c]}{E[T \mid X = c]} - \frac{\mathbb{P}^{\text{True}}[T = 1, M = 0 \mid X = c]E[Y \mid X = c^-]}{E[T \mid X = c]} \\
&- \frac{\mathbb{P}^{\text{True}}[T = 1, M = 1 \mid X = c]E[Y \mid X = c, M = 1]}{E[T \mid X = c]} \\
&= \frac{E[TY \mid X = c]}{E[T \mid X = c]} - \frac{\mathbb{P}^{\text{True}}[T = 0, M = 0 \mid X = c]E[Y \mid X = c^-]}{E[T \mid X = c]} \\
&- \frac{\eta_0 E[Y \mid X = c, M = 1]}{E[D \mid X = c]} \\
&= \frac{E[TY \mid X = c] - E[(1 - T)Y \mid X = c] - \eta_0 E[Y \mid X = c, M = 1]}{E[T \mid X = c]} \\
&= \frac{E[(2T - 1)Y \mid X = c] - \eta_0 E[Y \mid X = c, M = 1]}{\eta_0 + (1 - \eta_0)/2}.
\end{aligned}$$

The CLATE is immediate. □

Proof of Proposition 2. Recall by Lemma 8 that selection bounds on $q_t(Y(t)) = \mathbb{Q}(M = 1 \mid$

$X = c, Y(t))$ from Equation (6) correspond to likelihood ratios for $\mathbb{Q} \in \mathcal{M}(\underline{w}, \bar{w})$ as:

$$\begin{aligned}
\frac{d\mathbb{Q}(Y(1) \mid X = c, M = 0)}{d\mathbb{P}^{\text{Obs}}(Y \mid X = c^+)} &= \frac{1}{1-\eta}\left(1 + 2\frac{q_1(Y(1))}{1 - q_1(Y(1))}\right)^{-1} \\
&\in \left[\frac{1}{1-\eta}\left(1 + 2\frac{\eta}{2(1-\eta)}\Lambda_1\right)^{-1}, \frac{1}{1-\eta}\left(1 + 2\frac{\eta}{2(1-\eta)}\Lambda_1^{-1}\right)^{-1}\right] \\
\frac{d\mathbb{Q}(Y(0) \mid X = c, M = 1)}{d\mathbb{P}^{\text{Obs}}(Y \mid X = c^-)} &= \frac{2(1-\eta)}{\eta}\frac{q_0(Y(0))}{1 - q_0(Y(0))}
\end{aligned}$$

As a result, due to Equation (11),

$$\begin{aligned}
\frac{d\mathbb{Q}(Y(1) \mid X = c, M = 0)}{d\mathbb{P}^{\text{Obs}}(Y \mid X = c^+)} &\in \left[\frac{1}{1 - \eta + \eta\Lambda_1}, \frac{1}{1 - \eta + \eta\Lambda_1^{-1}}\right] \\
\frac{d\mathbb{Q}(Y(0) \mid X = c, M = 1)}{d\mathbb{P}^{\text{Obs}}(Y \mid X = c^-)} &\in \left[\Lambda_0^{-1}, \Lambda_0\right].
\end{aligned}$$

Next, we verify that the target distribution's estimand $E_{\mathbb{P}^{\text{Target}}}[\lambda(R)Y]$ achieves the relevant structural estimands. Suppose we call the target estimator $\bar{\psi}$. We proceed as follows:

$$\begin{aligned}
\bar{\psi}_{CLATE} &= E_{\mathbb{P}^{\text{Target}}}\left[\lambda(R)Y \mid X = c\right] \\
&= \mathbb{P}^{\text{Obs}}(X = c^+ \mid X = c)E_{\mathbb{P}^{\text{Target}}}\left[\lambda(R)Y \mid X = c^+\right] \\
&- \mathbb{P}^{\text{Obs}}(X = c^- \mid X = c)E_{\mathbb{P}^{\text{Target}}}\left[\lambda(R)Y \mid X = c^-\right] \\
&= E_{\mathbb{P}^{\text{Target}}}\left[Y \mid X = c^+\right] - E_{\mathbb{P}^{\text{Target}}}\left[Y \mid X = c^-\right] \\
&= E_{\mathbb{P}^{\text{True}}}\left[Y(1) \mid X = c, M = 0\right] - E_{\mathbb{P}^{\text{Obs}}}[Y(0) \mid X = c, M = 0] = \psi_{CLATE} \\
\bar{\psi}_{CATT} &= E_{\mathbb{P}^{\text{Obs}}}[Y \mid X = c^+] - E_{\mathbb{P}^{\text{True}}}[Y(0) \mid X = c^+] = \psi_{CATT} \\
\bar{\psi}_{CATE} &= E_{\mathbb{P}^{\text{True}}}[Y(1) \mid X = c] - E_{\mathbb{P}^{\text{True}}}[Y(0) \mid X = c] = \psi_{CATE}.
\end{aligned}$$

Finally, we verify that the Radon–Nikodym bounds have the appropriate forms by using Equation (5). The formula for the CLATE is immediate. The formula for the CATT follows by observing that:

$$\begin{aligned}
\frac{d\mathbb{P}^{\text{True}}(Y(0) \mid X = c^+)}{d\mathbb{P}^{\text{Obs}}(Y \mid X = c^-)} &= \mathbb{P}^{\text{Obs}}(M = 0 \mid X = c^+) * \frac{d\mathbb{P}^{\text{True}}(Y(0) \mid M = 0, X = c^+)}{d\mathbb{P}^{\text{Obs}}(Y \mid X = c^-)} \\
&+ \mathbb{P}^{\text{Obs}}(M = 1 \mid X = c^+) * \frac{d\mathbb{P}^{\text{True}}(Y(0) \mid M = 1, X = c^+)}{d\mathbb{P}^{\text{Obs}}(Y \mid X = c^-)}
\end{aligned}$$

$$
\begin{aligned}
&= \mathbb{P}^{\text{Obs}}(M = 0 \mid X = c^+) * 1 + \mathbb{P}^{\text{Obs}}(M = 1 \mid X = c^+) * \frac{d\mathbb{P}^{\text{True}}(Y(0) \mid X = c, M = 1)}{d\mathbb{P}^{\text{Obs}}(Y \mid X = c^-)} \\
&= (1 - \eta) + \eta \frac{d\mathbb{P}^{\text{True}}(Y(0) \mid X = c, M = 1)}{d\mathbb{P}^{\text{Obs}}(Y \mid X = c^-)} \\
&\in \left[(1 - \eta) + \eta\Lambda_0^{-1}, (1 - \eta) + \eta\Lambda_0\right],
\end{aligned}
$$

verifying the form.

The formula for the CATE similarly follows by the argument from Lemma 4. In particular:

$$
\begin{aligned}
\frac{d\mathbb{P}^{\text{True}}(Y(1) \mid X = c)}{d\mathbb{P}^{\text{Obs}}(Y \mid X = c^+)} &= \mathbb{P}^{\text{Obs}}(X = c^+ \mid X = c) * \frac{d\mathbb{P}^{\text{True}}(Y(1) \mid X = c^+)}{d\mathbb{P}^{\text{Obs}}(Y \mid X = c^+)} \\
&\quad + \mathbb{P}^{\text{Obs}}(X = c^- \mid X = c) * \frac{d\mathbb{P}^{\text{True}}(Y(1) \mid X = c^-)}{d\mathbb{P}^{\text{Obs}}(Y \mid X = c^+)} \\
&= \mathbb{P}^{\text{Obs}}(X = c^+ \mid X = c) * 1 + \mathbb{P}^{\text{Obs}}(X = c^- \mid X = c) * \frac{d\mathbb{P}^{\text{True}}(Y(1) \mid X = c, M = 0)}{d\mathbb{P}^{\text{Obs}}(Y \mid X = c^+)}.
\end{aligned}
$$

Similarly:

$$
\begin{aligned}
&\frac{d\mathbb{P}^{\text{True}}(Y(0) \mid X = c)}{d\mathbb{P}^{\text{Obs}}(Y \mid X = c^-)} = \mathbb{P}^{\text{Obs}}(X = c^+ \mid X = c) * \frac{d\mathbb{P}^{\text{True}}(Y(0) \mid X = c^+)}{d\mathbb{P}^{\text{Obs}}(Y \mid X = c^-)} + \mathbb{P}^{\text{Obs}}(X = c^- \mid X = c) * 1 \\
&= \eta\mathbb{P}^{\text{Obs}}(X = c^+ \mid X = c) \frac{d\mathbb{P}^{\text{True}}(Y(0) \mid X = c, M = 1)}{d\mathbb{P}^{\text{Obs}}(Y \mid X = c^-)} + (1 - \eta\mathbb{P}^{\text{Obs}}(X = c^+ \mid X = c)).
\end{aligned}
$$

Recall from the proof of Lemma 8 that $\mathbb{P}^{\text{Obs}}(X = c^+ \mid X = c) = 1/(2 - \eta)$. Therefore:

$$
\begin{aligned}
1\{T = 1\}\underline{w}(R) &= 1\{T = 1\} \left(\frac{1}{2 - \eta} + \frac{1 - \eta}{1 - 2\eta} \left(1 - \eta + \eta\Lambda_1\right)^{-1} \right) \\
1\{T = 1\}\bar{w}(R) &= 1\{T = 1\} \left(\frac{1}{2 - \eta} + \frac{1 - \eta}{1 - 2\eta} \left(1 - \eta + \eta\Lambda_1^{-1}\right)^{-1} \right) \\
1\{T = 0\}\underline{w}(R) &= 1\{T = 0\} \left(\frac{2 - 2\eta}{2 - \eta} + \frac{\eta}{2 - \eta}\Lambda_0^{-1} \right) \\
1\{T = 0\}\bar{w}(R) &= 1\{T = 0\} \left(\frac{2 - 2\eta}{2 - \eta} + \frac{\eta}{2 - \eta}\Lambda_0 \right),
\end{aligned}
$$

completing the final proof. □

Proofs for Proposition 3, Proposition 4, and Proposition 5 are in the online appendix.

References

Abadie, A. (2003): "Semiparametric instrumental variable estimation of treatment response models," *Journal of Econometrics*, 113, 231–263.

Aronow, P. M. and D. K. K. Lee (2013): "Interval estimation of population means under unknown but bounded probabilities of sample selection," *Biometrika*, 100, 235–240.

Athey, S. and G. W. Imbens (2022): "Design-based analysis in difference-in-differences settings with staggered adoption," *Journal of Econometrics*, 226, 62–79.

Basit, M. A., M. A. H. M. Latif, and A. S. Wahed (2023): "A Risk-Ratio-Based Marginal Sensitivity Model for Causal Effects in Observational Studies," .

Bertsimas, D., K. Imai, and M. L. Li (2022): "Distributionally robust causal inference with observational data," .

Bugni, F. A. and I. A. Canay (2021): "Testing continuity of a density via g-order statistics in the regression discontinuity design," *Journal of Econometrics*, 221, 138–159.

Callaway, B. and P. H. Sant'Anna (2021): "Difference-in-differences with multiple time periods," *Journal of Econometrics*, 225, 200–230.

Caughey, D. and J. S. Sekhon (2011): "Elections and the regression discontinuity design: Lessons from close US house races, 1942–2008," *Political Analysis*, 19, 385–408.

Chernozhukov, V., C. Cinelli, W. Newey, A. Sharma, and V. Syrgkanis (2022): "Long Story Short: Omitted Variable Bias in Causal Machine Learning," Working Paper 30302, National Bureau of Economic Research.

Christensen, T. and B. Connault (2023): "Counterfactual Sensitivity and Robustness," *Econometrica*, 91, 263–298.

Cinelli, C. and C. Hazlett (2020): "Making sense of sensitivity: Extending omitted variable bias," *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 82, 39–67.

Dorn, J. and K. Guo (2023): "Sharp sensitivity analysis for inverse propensity weighting via quantile balancing," *Journal of the American Statistical Association*, 118, 2645–2657.

Dorn, J., K. Guo, and N. K. and (2024): "Doubly-Valid/Doubly-Sharp Sensitivity Analysis for Causal Inference with Unmeasured Confounding," *Journal of the American Statistical Association*, 0, 1–12.

Egger, P. and M. Koethenbuerger (2010): "Government spending and legislative organization: Quasi-experimental evidence from Germany," *American Economic Journal: Applied Economics*, 2, 200–212.

Eggers, A. C., R. Freier, V. Grembi, and T. Nannicini (2018): "Regression discontinuity designs based on population thresholds: Pitfalls and solutions," *American Journal of Political Science*, 62, 210–229.

Fang, Z. and A. Santos (2019): "Inference on Directionally Differentiable Functions," *Review of Economic Studies*, 86.

Frauen, D., V. Melnychuk, and S. Feuerriegel (2023): "Sharp Bounds for Generalized Causal Sensitivity Analysis," in *Advances in Neural Information Processing Systems*, ed. by A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine, Curran Associates, Inc., vol. 36, 40556–40586.

Freidling, T. and Q. Zhao (2022): "Sensitivity Analysis with the R^2-calculus," *arXiv preprint arXiv:2301.00040.*

Gerard, F., M. Rokkanen, and C. Rothe (2020): "Bounds on treatment effects in regression discontinuity designs with a manipulated running variable," *Quantitative Economics*, 11, 839–870.

Ghanem, D., P. H. Sant'Anna, and K. Wüthrich (2022): "Selection and parallel trends," *arXiv preprint arXiv:2203.09001.*

Hahn, J., P. Todd, and W. Van der Klaauw (2001): "Identification and estimation of treatment effects with a regression-discontinuity design," *Econometrica*, 69, 201–209.

Ho, K. and A. M. Rosen (2017): "Partial Identification in Applied Research: Benefits and Challenges," in *Advances in Economics and Econometrics: Eleventh World Congress*, ed. by B. Honoré, A. Pakes, M. Piazzesi, and L. Samuelson, Cambridge University Press, Econometric Society Monographs, 307–359.

Huang, M. and S. D. Pimentel (2024): "Variance-based sensitivity analysis for weighting estimators results in more informative bounds," *Biometrika*, 112, asae040.

Ishihara, T. and M. Sawada (2020): "Manipulation-Robust Regression Discontinuity Designs," *arXiv preprint arXiv:2009.07551*.

Ishikawa, K., N. He, and T. Kanamori (2023): "A Convex Framework for Confounding Robust Inference," .

Jin, Y., Z. Ren, and Z. Zhou (2022): "Sensitivity analysis under the f-sensitivity models: a distributional robustness perspective," .

Lee, D. S. (2008): "Randomized experiments from non-random selection in US House elections," *Journal of Econometrics*, 142, 675–697.

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

Marx, P., E. Tamer, and X. Tang (2024): "Parallel trends and dynamic choices," *Journal of Political Economy Microeconomics*, 2, 129–171.

Masten, M. A. and A. Poirier (2018): "Identification of treatment effects under conditional partial independence," *Econometrica*, 86, 317–351.

——— (2021): "Salvaging falsified instrumental variable models," *Econometrica*, 89, 1449–1469.

Masten, M. A., A. Poirier, and L. Zhang (2024): "Assessing Sensitivity to Unconfoundedness: Estimation and Inference," *Journal of Business & Economic Statistics*, 42, 1–13.

McCrary, J. (2008): "Manipulation of the running variable in the regression discontinuity design: A density test," *Journal of Econometrics*, 142, 698–714.

Molinari, F. (2020): "Chapter 5 - Microeconometrics with partial identification," in *Handbook of Econometrics, Volume 7A*, ed. by S. N. Durlauf, L. P. Hansen, J. J. Heckman, and R. L. Matzkin, Elsevier, vol. 7 of *Handbook of Econometrics*, 355–486.

Noack, C. (2021): "Sensitivity of LATE estimates to violations of the monotonicity assumption," *arXiv preprint arXiv:2106.06421*.

Otsu, T., K.-L. Xu, and Y. Matsushita (2013): "Estimation and inference of discontinuity in density," *Journal of Business & Economic Statistics*, 31, 507–524.

Pinelis, I. (2019): "Exact bounds on the inverse Mills ratio and its derivatives," *Complex Analysis and Operator Theory*, 13, 1643–1651.

Rambachan, A. and J. Roth (2023): "A more credible approach to parallel trends," *Review of Economic Studies*.

Ramsahai, R. R. (2012): "Causal Bounds and Observable Constraints for Non-deterministic Models." *Journal of Machine Learning Research*, 13.

Rosenman, E., K. Rajkumar, R. Gauriot, and R. Slonim (2019): "Optimized partial identification bounds for regression discontinuity designs with manipulation," *arXiv preprint arXiv:1910.02170*.

Roth, J. and P. H. Sant'Anna (2023): "When is parallel trends sensitive to functional form?" *Econometrica*, 91, 737–747.

Tan, Z. (2006): "A distributional approach for causal inference using propensity scores," *Journal of the American Statistical Association*, 101, 1619–1637.

——— (2024a): "Model-assisted sensitivity analysis for treatment effects under unmeasured confounding via regularized calibrated estimation," *Journal of the Royal Statistical Society Series B: Statistical Methodology*.

——— (2024b): "Sensitivity models and bounds under sequential unmeasured confounding in longitudinal studies," *Biometrika*, asae044.

Van Kippersluis, H. and C. A. Rietveld (2018): "Beyond plausibly exogenous," *The Econometrics Journal*, 21, 316–331.

Yap, L. (2025): "Sensitivity of Policy-Relevant Treatment Parameters to Violations of Monotonicity," *Journal of Applied Econometrics*.

Zhao, Q., D. S. Small, and B. B. Bhattacharya (2019): "Sensitivity analysis for inverse probability weighting estimators via the percentile bootstrap," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 81, 735–761.

D Online Appendix

D.1 Proof of Proposition 3

Proof of Proposition 3. This proof is extensive, so we split up the argument into separate subsubsections. To prove that the bounds are sharp, we adopt the following strategy. First, we show that our bounds are feasible, by showing that the likelihood constraint implies the allowable distribution and that any ψ_1, ψ_0 that solves our problem are jointly feasible. Second and more extensively, we explicitly describe a worst-case distribution, then show that it is in the model family and achieves the extremal estimands, and then show that a convex combination of distributions enables achieving any pair of interior estimands as well.

D.1.1 Likelihood Constraint Implies Allowable Distribution

First, we show that if $\mathbb{Q} \in \mathcal{M}'(\Lambda)$, then $\mathbb{Q} \in \mathcal{M}(\Lambda, \Lambda)$. It is clear that it only remains to show that for $t = 1, 0$, we have $\frac{\mathbb{Q}(M=1|Y(t),X=c)}{\mathbb{Q}(M=0|Y(t),X=c)} / \frac{\mathbb{P}^{\text{True}}(M=1|X=c)}{\mathbb{P}^{\text{True}}(M=0|X=c)} \in [\Lambda^{-1}, \Lambda]$ almost surely. This holds by iterated expectations over $\int \mathbb{Q}(M = 1 \mid Y(1), Y(0), X = c) d\mathbb{Q}(Y(1-t) \mid X = c, Y(1-t))$ and the restriction on the domain of $\mathbb{Q}(M = 1 \mid Y(1), Y(0), X = c)$ under Equation (7). Therefore $\mathbb{Q} \in \mathcal{M}(\Lambda, \Lambda)$. This completes the first direction of the proof.

D.1.2 Any Two Feasible Potential Outcomes are Jointly Feasible

We now proceed to construct a $\mathbb{Q}'$ corresponding to ψ_1 and ψ_0.

For simplicity, we proceed assuming that Y is continuously distributed. (If the conditional distribution of Y contains mass points at the referenced quantiles, add a uniform random variable to provide a strict ordering on observations of $Y \mid X$ and then apply the construction below to the quantiles of the tie-broken Y.)

Recall that all of our structural estimands of interest can be written as observable linear transformations of the partially identified average potential outcomes $E[Y(t) \mid X = c, M = 1 - t]$ for $t = 1, 0$.

We define a distribution $\mathbb{Q}_{+,+}$ which we show is in the model family $\mathcal{M}'(\Lambda) \subset \mathcal{M}(\Lambda, \Lambda) \subset \mathcal{M}'(\infty)$ and achieves the upper bounds on both average potential outcomes within $\mathcal{M}(\Lambda, \Lambda)$. We then argue that any $\mathbb{Q}_{\pm,\pm}$ combination of upper and lower bounds is feasible. Finally, we argue by mixture that any pair of structural estimands of interest between the two lower and upper bounds under $\mathcal{M}(\Lambda, \Lambda)$ are feasible for some distribution $\mathbb{Q} \in \mathcal{M}'(\Lambda)$.

D.1.3 Construction of Worst-Case Distribution

We will define a distribution $\mathbb{Q}_{+,+}$ with an unobserved confounder "V." $V = 1$ will correspond to a high manipulation probability, small treated potential outcomes (to maximize $E[Y(1) \mid X = c^+, M = 0]$), and large untreated outcomes (to maximize $E[Y(0) \mid X = c^+, M = 1]$). At the threshold, a fraction τ_1 of treated and $1 - \tau_0$ of untreated observations will have $V = 1$, where:

$$\tau_1 = \frac{1 + \eta(\Lambda - 1)}{\Lambda + 1}$$
$$\tau_0 = \frac{\Lambda}{\Lambda + 1}.$$

In order to facilitate this drawing, we define the observable V function:

$$V(x, t, y) = \begin{cases} 1\{y \leq Q_{\tau_1}(Y \mid X = c^+)\} & \text{if } t = 1 \\ 1\{y > Q_{\tau_0}(Y \mid X = c^-)\} & \text{if } t = 0. \end{cases}$$

Finally, in order to match worst-case selection probabilities, $M \mid X = c, V = v$ will be drawn iid from a $Bern(Tq_{++}(v))$ distribution, where $q_{++}(v)$ is defined as:

$$q_{++}(v) = \begin{cases} \frac{\eta}{(\Lambda+1)(1-\tau_1)} & \text{if } v = 0 \\ \frac{\eta\Lambda}{(\Lambda+1)\tau_1} & \text{if } v = 1. \end{cases}$$

Finally, we will draw unobserved potential outcomes from the distribution with the same value of M and V.

Formally, the distribution $\mathbb{Q}_{+,+}$ over $(M, X(1), X(0), Y(1), Y(0), T, T(0))$ is defined as follows:

1. Draw $(X, T, Y) \sim \mathbb{P}^{\text{Obs}}$
2. Define the random variable $V = V(X, T, Y)$
3. Draw $M \sim Bern(Tq_{++}(V))$

4. Set the lower-case variable realizations for later use $x = X$, and $y = Y$, $t = T$, $v = V$, and $m = M$
5. Set $X(M) = X$ and $Y(t) = y$
6. Draw $X(1-M)$ from the distribution of $X \mid M = 1-m$ under $\mathbb{Q}_{+,+}$ defined so far
7. If $m = 0$, draw $Y(1-t)$ from the distribution of $Y \mid X = 2c - x, M = 0, V = v$ under the construction of $\mathbb{Q}_{+,+}$ so far
8. If $m = 1$, draw $Y(0)$ from the distribution of $Y \mid X = 2c - x, V = V(t, y)$ under the construction of $\mathbb{Q}_{+,+}$ so far
9. Set $T(0) = 1\{X(0) > c\}$
10. Return data $(M, X(1), X(0), Y(1), Y(0), T, T(0))$

D.1.4 Showing Constructed Distribution is in Model Family

We wish to show that $\mathbb{Q}_{+,+} \in \mathcal{M}'(\Lambda)$, i.e.:

(a) The distribution of $(X = X(M), Y = Y(T), T)$ under $\mathbb{Q}_{+,+}$ marginalizes to the distribution of (X, T, Y) under $\mathbb{P}^{\text{Obs}}$

(b) $\mathbb{Q}_{+,+}(T = 0, M = 1) = 0$

(c) $\mathbb{Q}_{+,+}(M = 1 \mid X = c^+) = \eta$

(d) $\mathbb{Q}_{+,+}(Y(t) \le y \mid X = x, M = 0)$ is continuous at c and $\mathbb{Q}_{+,+}(Y(1) \le y \mid x = c, M = 1)$ is right-continuous at $x = c$.

(e) $\mathbb{Q}_{+,+}(Y(1), Y(0), T(0), M \mid X = c)$ is a conditional distribution that is well-defined as the appropriate continuous limit.

(f) $\mathbb{Q}_{+,+}$ satisfies Equation (7):

$$\frac{\mathbb{Q}_{+,+}(M = 1 \mid Y(1), Y(0), X = c)}{\mathbb{Q}_{+,+}(M = 0 \mid Y(1), Y(0), X = c)} \Bigg/ \frac{\mathbb{P}^{\text{True}}(M = 1 \mid X = c)}{\mathbb{P}^{\text{True}}(M = 0 \mid X = c)} \in [\Lambda^{-1}, \Lambda].$$

Requirements (a) and (b) are immediate.

Requirement (c) follows by inspection:

$$\begin{aligned}\mathbb{Q}_{+,+}(M=1 \mid X=c^+) &= E_{\mathbb{P}^{\text{Obs}}}[q_{++}(V) \mid X=c^+] \\ &= \sum_v \mathbb{P}^{\text{Obs}}(V=v \mid X=c^+)q_{++}(v) \\ &= \frac{\eta}{\Lambda+1} + \frac{\eta\Lambda}{\Lambda+1} = \eta = \mathbb{P}^{\text{True}}(M=1 \mid X=c^+).\end{aligned}$$

The requirement (d) holds as follows. $\mathbb{P}^{\text{True}}(Y \le y \mid X=x)$ is left- and right-continuous at $x=c$ as follows. For left-continuity, for $x<c$, $\mathbb{P}^{\text{True}}(Y \le y \mid X=x) = \mathbb{P}^{\text{True}}(Y \le y \mid M=0, X=x)$ is left-continuous at $x=c$ by assumption. For right-continuity, for $x>c$,

$$\begin{aligned}\mathbb{P}^{\text{True}}(Y \le y \mid X=x) &= \mathbb{P}^{\text{True}}(M=1 \mid X=x)\mathbb{P}^{\text{True}}(Y \le y \mid M=1, X=x) \\ &\quad + \mathbb{P}^{\text{True}}(M=0 \mid X=x)\mathbb{P}^{\text{True}}(Y \le y \mid M=0, X=x),\end{aligned}$$

which by assumption is right-continuous at $x=c$. As a result, $\mathbb{Q}_{+,+}(Y(t) \le y \mid X=x, M=0)$ is continuous in x at $x=c$. Note also that for $X>c$ and $y \le Q_{\tau_1}(Y \mid X=c^+)$, we have:

$$\begin{aligned}&\mathbb{Q}_{+,+}(Y(t) \le y \mid X, M=1) \\ &\quad= \frac{\mathbb{P}^{\text{Obs}}(Y(t) \le y \mid X)q_{++}(0)}{\mathbb{P}^{\text{Obs}}(Y(t) \le Q_{\tau_1}(Y \mid X=c^+) \mid X)q_{++}(0) + \mathbb{P}^{\text{Obs}}(Y(t) > Q_{\tau_1}(Y \mid X=c^+) \mid X)},\end{aligned}$$

and similarly, for $Y > Q_{\tau_1}(Y \mid X=c^+)$:

$$\begin{aligned}&\mathbb{Q}_{+,+}(Y(t) \le y \mid X, M=1) \\ &\quad= \frac{\mathbb{P}^{\text{Obs}}(Y(t) \le Q_{\tau_1}(Y \mid X=c^+) \mid X)q_{++}(0) + \mathbb{P}^{\text{Obs}}(Y(t) \in [Q_{\tau_1}(Y \mid X=c^+), y] \mid X)q_{++}(0)}{\mathbb{P}^{\text{Obs}}(Y(t) \le Q_{\tau_1}(Y \mid X=c^+) \mid X)q_{++}(0) + \mathbb{P}^{\text{Obs}}(Y(t) > Q_{\tau_1}(Y \mid X=c^+) \mid X)},\end{aligned}$$

both of which are continuous in X.

(e) holds as follows, where we write $T(x) = 1\{x > c\}$. For $m=1$ and some $A \subseteq \mathbb{R}^2$:

$$\begin{aligned}&\lim_{\varepsilon \to 0^+} \mathbb{Q}_{+,+}((Y(1), Y(0)) \in A, T(0)=t, M=1, V=v \mid |X-\varepsilon| \le c) \\ &= \int 1\{X>c, (Y(1), Y(0)) \in A\}q_{++}(v)\frac{1}{2}d\mathbb{Q}_{+,+}(V, Y(0) \mid X, Y, M=1)d\mathbb{P}^{\text{Obs}}(Y, X \mid |X-c| \le \varepsilon) \\ &= \int q_{++}(v)\frac{1\{X>c, (Y(1), Y(0)) \in A, V(X,1,Y)=v\}}{2}d\mathbb{Q}_{+,+}(Y \mid |X-c| \le c, T=0, V=v)d\mathbb{P}^{\text{Obs}}(Y, X \mid |X-c| \le \varepsilon),\end{aligned}$$

which has a well-defined limit by one-sided continuity of the CDF of $Y \mid X$ under our assumptions. The remaining claim holds by taking the sum over v. Similarly, for $m = 0$:

$$\begin{aligned}
&\lim_{\varepsilon \to 0^+} \mathbb{Q}_{+,+}((Y(1), Y(0)) \in A, T(0) = t, M = 1, V = v \mid |X - \varepsilon| \leq c) \\
=&1\{X \leq c\}\mathbb{Q}_{+,+}((Y(1), Y(0)) \in A, T(0) = t, M = 0, V = v \mid |X - \varepsilon| \leq c) \\
+&1\{X > c\}\mathbb{Q}_{+,+}((Y(1), Y(0)) \in A, T(0) = t, M = 0, V = v \mid |X - \varepsilon| \leq c),
\end{aligned}$$

which is continuous by analogous arguments.

Requirement (f) involves a longer argument, so we show it in a separate section.

D.1.5 $\mathbb{Q}_{+,+}$ Satisfies the Odds Ratio Bound

We wish to show:

$$\frac{\mathbb{Q}_{+,+}(M = 1 \mid Y(1), Y(0), X = c)}{\mathbb{Q}_{+,+}(M = 0 \mid Y(1), Y(0), X = c)} \Big/ \frac{\mathbb{P}^{\text{True}}(M = 1 \mid X = c)}{\mathbb{P}^{\text{True}}(M = 0 \mid X = c)} \in [\Lambda^{-1}, \Lambda].$$

Recall that $\frac{\mathbb{P}^{\text{True}}(M=1|X=c)}{\mathbb{P}^{\text{True}}(M=0|X=c)} = \frac{\eta}{2(1-\eta)}$.

Notice that $Y(1), Y(0)$ are iid given M, X, V under $\mathbb{Q}_{+,+}$. As a result, we can equivalently show that for $v = 0, 1$:

$$\begin{aligned}
\frac{\mathbb{Q}_{+,+}(M = 1 \mid V = v, X = c)}{\mathbb{Q}_{+,+}(M = 0 \mid V = v, X = c)} &= \frac{\sum_t \mathbb{Q}_{+,+}(T = t, V = v, M = 1 \mid X = c)}{\sum_t \mathbb{Q}_{+,+}(T = t, V = v, M = 0 \mid X = c)} \\
&\in \left[\Lambda^{-1}\frac{2(1-\eta)}{\eta}, \Lambda\frac{2(1-\eta)}{\eta}\right].
\end{aligned}$$

We do so case-by-case for values of v. In the $v = 0$ case:

$$\begin{aligned}
\sum_t \mathbb{Q}_{+,+}(T = t, V = 0, M = 1 \mid X = c) &= \frac{1}{2-\eta}(1-\tau_1)q_{++}(0) = \frac{\eta}{(\Lambda+1)(2-\eta)} \\
\sum_t \mathbb{Q}_{+,+}(T = t, V = 0, M = 0 \mid X = c) &= \frac{1}{2-\eta}(1-\tau_1)(1-q_{++}(0)) + \frac{1-\eta}{2-\eta}\tau_0 \\
&= \frac{1}{2-\eta}\left(1-\tau_1-\frac{\eta}{\Lambda+1}\right) + \frac{\Lambda(1-\eta)}{(2-\eta)(\Lambda+1)} \\
&= \frac{1}{(2-\eta)(\Lambda+1)}\left(\Lambda(1-\eta)\right) + \frac{\Lambda(1-\eta)}{(2-\eta)(\Lambda+1)} \\
&= \frac{2\Lambda(1-\eta)}{(\Lambda+1)(2-\eta)}
\end{aligned}$$

$$\frac{\mathbb{Q}_{+,+}(M=1 \mid V=0, X=c)}{\mathbb{Q}_{+,+}(M=0 \mid V=0, X=c)} = \Lambda^{-1}\frac{\eta}{2(1-\eta)}.$$

Similarly, in the $v=1$ case:

$$\begin{aligned}
\sum_t \mathbb{Q}_{+,+}(T=t, V=1, M=1 \mid X=c) &= \frac{1}{2-\eta}\tau_1 q_{++}(1) = \frac{\Lambda\eta}{(\Lambda+1)(2-\eta)} \\
\sum_t \mathbb{Q}_{+,+}(T=t, V=1, M=0 \mid X=c) &= \frac{1}{2-\eta}\tau_1(1-q_{++}(1)) + \frac{1-\eta}{2-\eta}(1-\tau_0) \\
&= \frac{1}{(\Lambda+1)(2-\eta)}(1-\eta) + \frac{1-\eta}{(\Lambda+1)(2-\eta)} \\
\frac{\mathbb{Q}_{+,+}(M=1 \mid V=1, X=c)}{\mathbb{Q}_{+,+}(M=0 \mid V=1, X=c)} &= \Lambda\frac{\eta}{2(1-\eta)}.
\end{aligned}$$

Therefore, $\mathbb{Q}_{+,+} \in \mathcal{M}'(\Lambda)$.

D.1.6 Showing Constructed Distribution is Worst-Case for Both Estimands

We wish to show that $E_{\mathbb{Q}_{+,+}}[Y(1) \mid X=c, M=0]$ and $E_{\mathbb{Q}_{+,+}}[Y(0) \mid X=c, M=1]$ are maximal within $\mathcal{M}(\Lambda,\Lambda)$. By the proof of Proposition 2, the maximal conditional expectations are achieved if:

$$\frac{d\mathbb{Q}_{+,+}(Y(1) \mid X=c, M=0)}{d\mathbb{P}^{\text{Obs}}(Y(1) \mid X=c^+)} = \begin{cases} \frac{\Lambda}{\Lambda+\eta(1-\Lambda)} & \text{if } Y(1) > Q_{\tau_1}(Y \mid X=c^+) \\ \frac{1}{1+\eta(\Lambda-1)} & \text{if } Y(1) \le Q_{\tau_1}(Y \mid X=c^+) \end{cases}$$

$$\frac{d\mathbb{Q}_{+,+}(Y(0) \mid X=c, M=1)}{d\mathbb{P}^{\text{Obs}}(Y(0) \mid X=c^-)} = \begin{cases} \Lambda & \text{if } Y(0) > Q_{\tau_0}(Y \mid X=c^+) \\ \Lambda^{-1} & \text{if } Y(0) \le Q_{\tau_0}(Y \mid X=c^+). \end{cases}$$

We begin with the $Y(1)$ distribution:

$$\begin{aligned}
\frac{d\mathbb{Q}_{+,+}(Y(1) \mid X=c, M=0)}{d\mathbb{P}^{\text{Obs}}(Y(1) \mid X=c^+)} &= \frac{d\mathbb{Q}_{+,+}(Y(1) \mid X=c^+, M=0)}{d\mathbb{P}^{\text{Obs}}(Y(1) \mid X=c^+)} \\
&= \frac{d\mathbb{Q}_{+,+}(Y(1), M=0 \mid X=c^+)}{(1-\eta)d\mathbb{P}^{\text{Obs}}(Y(1) \mid X=c^+)} \\
&= \frac{\mathbb{Q}_{+,+}(M=0 \mid X=c^+, Y(1))}{1-\eta} \\
&= \frac{1-\mathbb{Q}_{+,+}(M=1 \mid X=c^+, Y(1))}{1-\eta} \\
&= \frac{1-q_{++}(V(X,1,Y(1))}{1-\eta}
\end{aligned}$$

$$= \begin{cases} \frac{1-\frac{\eta}{\Lambda+\eta(1-\Lambda)}}{1-\eta} & \text{if } Y(1) < Q_{\tau_1}(Y \mid X = c^+) \\ \frac{1-\frac{\eta\Lambda}{1+\eta(\Lambda-1)}}{1-\eta} & \text{if } Y(1) \leq Q_{\tau_1}(Y \mid X = c^+) \end{cases}$$
$$= \begin{cases} \frac{\Lambda}{\Lambda+\eta(1-\Lambda)} & \text{if } Y(1) < Q_{\tau_1}(Y \mid X = c^+) \\ \frac{1}{1+\eta(\Lambda-1)} & \text{if } Y(1) \leq Q_{\tau_1}(Y \mid X = c^+) \end{cases},$$

which are the desired likelihood ratios.

We continue with the $Y(0)$ distribution:

$$\begin{aligned} \frac{d\mathbb{Q}_{+,+}(Y(0) \mid X = c, M = 1)}{d\mathbb{P}^{\text{Obs}}(Y(0) \mid X = c^-)} &= \frac{d\mathbb{Q}_{+,+}(Y(0) \mid X = c^+, M = 1)}{d\mathbb{P}^{\text{Obs}}(Y(0) \mid X = c^-)} \\ &= \frac{\mathbb{Q}_{+,+}(V = V(X, 0, Y(0)) \mid X = c^+, M = 1)}{d\mathbb{P}^{\text{Obs}}(V = V(X, 0, Y(0)) \mid X = c^-)} \\ &= \frac{\mathbb{Q}_{+,+}(V = V(X, 0, Y(0)), M = 1 \mid X = c^+)}{\eta d\mathbb{P}^{\text{Obs}}(V = V(X, 0, Y(0)) \mid X = c^-)}. \end{aligned}$$

By inspection of the two cases, this is $\Lambda V(X, 0, Y(0)) + \Lambda^{-1}(1 - V(X, 0, Y(0)))$, i.e. the desired likelihood ratio.

D.1.7 Construction of All Mixtures of Estimands

By symmetric arguments, we can find a $\mathbb{Q}_{+,-}$ that achieves the maximal value of $E_{\mathbb{Q}}[Y(1) \mid X = c, M = 0]$ and minimal value of $E_{\mathbb{Q}}[Y(0) \mid X = c, M = 1]$ over $\mathbb{Q} \in \mathcal{M}(\Lambda, \Lambda)$, as well as analogous $\mathbb{Q}_{-,+}$ and $\mathbb{Q}_{-,-}$ for all of the other extreme combinations.

Now suppose we are achieving some (ψ_1, ψ_0) within both pointwise bounds, i.e. there are some $\alpha_1, \alpha_0 \in [0, 1]$ such that:

$$\begin{aligned} \alpha_1 \left(E_{\mathbb{Q}_{+,+}}[Y(1) \mid X = c, M = 0] - E_{\mathbb{Q}_{-,-}}[Y(1) \mid X = c, M = 0]\right) &= \psi_1 - E_{\mathbb{Q}_{-,-}}[Y(1) \mid X = c, M = 0] \\ \alpha_0 \left(E_{\mathbb{Q}_{+,+}}[Y(0) \mid X = c, M = 1] - E_{\mathbb{Q}_{-,-}}[Y(0) \mid X = c, M = 1]\right) &= \psi_0 - E_{\mathbb{Q}_{-,-}}[Y(0) \mid X = c, M = 1]. \end{aligned}$$

Define $\mathbb{Q}^*$ as follows:

- Draw $V_1 \sim Bern(\alpha_1)$ and $V_0 \sim Bern(\alpha_0)$
- Write S_d to be $+$ if $V_t = 1$ and S_t to be $-$ if $V_t = 0$
- Draw $(X, M, D, Y(1), Y(0)) \sim \mathbb{Q}_{S_1, S_0}$

By inspection, $\mathbb{Q}^* \in \mathcal{M}'(\Lambda)$. It also has:

$$\begin{aligned} E_{\mathbb{Q}^*}[Y(t) \mid X = c, M = 1-t] &= \alpha_t \psi_t^+(\Lambda) + (1-\alpha_t)\psi_t^-(\Lambda) \\ &= \frac{\psi_t^+(\Lambda)\left(\psi_t - \psi_t^-(\Lambda)\right) + \left(\psi_t^+(\Lambda) - \psi_t\right)\psi_t^-(\Lambda)}{\psi_t^+(\Lambda) - \psi_t^-(\Lambda)} \\ &= \psi_t \end{aligned}$$

Demonstrating the claim.

□

D.2 Additional Proofs for Section 3

Proof of Proposition 4. With the $\lambda(R)$ as stated, we have:

$$\begin{aligned} E_{\mathbb{P}^{\text{Target}}}[\lambda(R)Y \mid X] &= \eta(X)\mathbb{P}^{\text{Obs}}(Z=1 \mid X)\frac{1-\mathbb{P}^{\text{Obs}}(Z=1\mid X)}{\mathbb{P}^{\text{Obs}}(Z=1\mid X)\mathbb{P}^{\text{Obs}}(Z=0\mid X)}E_{\mathbb{P}^{\text{Target}}}[Y \mid X, Z=1] \\ &\quad + \eta(X)\mathbb{P}^{\text{Obs}}(Z=0 \mid X)\frac{0-\mathbb{P}^{\text{Obs}}(Z=1\mid X)}{\mathbb{P}^{\text{Obs}}(Z=1\mid X)\mathbb{P}^{\text{Obs}}(Z=0\mid X)}E_{\mathbb{P}^{\text{Target}}}[Y \mid X, Z=0] \\ &= \eta(X)\left(E_{\mathbb{P}^{\text{Target}}}[Y \mid X, Z=1] - E_{\mathbb{P}^{\text{Target}}}[Y \mid X, Z=0]\right) \\ &= E_{\mathbb{P}^{\text{True}}}\left[\eta(X)E_{\mathbb{P}^{\text{True}}}\left[\sum_z \omega(z \mid X)\{Y(T(1),z) - Y(T(0),z)\} \mid X, T(1), T(0)\right] \mid X\right] \\ &= E_{\mathbb{P}^{\text{True}}}\left[\eta(X)1\{Co\}\sum_z \omega(z\mid X)(Y(1,z)-Y(0,z)) \mid X\right]. \end{aligned}$$

By iterated expectations, $E_{\mathbb{P}^{\text{Target}}}[\lambda(R)Y] = \psi$.

The constraints on $\frac{d\mathbb{P}^{\text{Target}}(Y|X,T,Z)}{d\mathbb{P}^{\text{Obs}}(Y|X,T,Z)}$ follow by any pointwise bounds on those likelihood ratios that are derived as implications of Lemma 5. □

Proof of Proposition 5. We suppress the dependence on X because it is constant.

Notice that in this example, $\mathbb{P}^{\text{Obs}}(Co) = \mathbb{P}^{\text{Obs}}(T=1 \mid Z=1) - \mathbb{P}^{\text{Obs}}(T=1\mid Z=0) = 1/2$; $\mathbb{P}^{\text{Obs}}(Nt) = 1 - \mathbb{P}^{\text{Obs}}(T=1\mid Z=1) = 1/2$; and $\mathbb{P}^{\text{Obs}}(Z=1) = 1/2$. As a result, we can write $\omega(z \mid x) = 2z$.

Because $\mathbb{P}^{\text{Obs}}(At) = 0$, we directly observe $\mathbb{P}^{\text{True}}(Y(1,1) \mid Co) = \mathbb{P}^{\text{Obs}}(Y \mid Z=1, T=1)$ and $E[Y(1,1) \mid Co] = 0$. The partial identification problem is to bound $\mathbb{P}^{\text{True}}(Y(0,1) \mid Co)$.

Under the structural model, we have $1 \leq \frac{d\mathbb{P}^{\text{True}}(Y(0,0)|Nt)}{d\mathbb{P}^{\text{True}}(Y(0,1)|Nt)} = \frac{d\mathbb{P}^{\text{True}}(Y(0,0)|Nt)}{d\mathbb{P}^{\text{Obs}}(Y|Z=1,T=0)} < \infty$, which implies that the distribution of $Y(0,0) \mid Nt$ is the distribution of $Y \mid Z=1, T=0$. To see this, if $\frac{d\mathbb{P}^{\text{True}}(Y(0,0)|Nt)}{d\mathbb{P}^{\text{True}}(Y(0,1)|Nt)} > 1$ with positive probability, then $\frac{d\mathbb{P}^{\text{True}}(Y(0,0)|Nt)}{d\mathbb{P}^{\text{True}}(Y(0,1)|Nt)} < 1$ with some other positive probability, which contradicts the lower bound, so $\frac{d\mathbb{P}^{\text{True}}(Y(0,0)|Nt)}{d\mathbb{P}^{\text{True}}(Y(0,1)|Nt)} = 1$ almost everywhere. Notice analogously that under the structural model, $0 \leq \frac{d\mathbb{P}^{\text{True}}(Y(0,1)|Co)}{d\mathbb{P}^{\text{True}}(Y(0,0)|Co)} \leq 1$, which implies that the distribution of $Y(0,1) \mid Co$ is the same as the distribution of $Y(0,0) \mid Co$. As a result, $E[Y(0,1) \mid Co]$ is point-identified:

$$\begin{aligned} E[Y \mid Z=0, T=0] &= \mathbb{P}^{\text{Obs}}(Co \mid Z=0, T=0)E[Y(0,0) \mid Co] + \mathbb{P}^{\text{Obs}}(Nt)E[Y(0,0) \mid Nt] \\ 0 &= 0.5E[Y(0,1) \mid Co] + 0.5E[Y \mid Z=1, T=0] = 0.5E[Y(0,1) \mid Co]. \end{aligned}$$

The sharp bounds are the singleton $\{0\}$.

Now consider the statistical bounds of Proposition 4. The likelihood ratios of interest from Lemma 5 are:

$$\begin{aligned}
&\frac{d\mathbb{P}^{\text{Target}}(Y \mid T=1, Z=1)}{d\mathbb{P}^{\text{Obs}}(Y \mid T=1, Z=1)} = \overbrace{\omega(1)}^{=1} + \overbrace{\omega(0)}^{=0} \overbrace{\frac{d\mathbb{P}^{\text{True}}(Y(1,0) \mid Co)}{d\mathbb{P}^{\text{True}}(Y(1,1) \mid Co)}}^{\in[1,\infty)} \\
&\frac{d\mathbb{P}^{\text{Target}}(Y \mid T=1, Z=0)}{d\mathbb{P}^{\text{Obs}}(Y \mid T=1, Z=0)} = 1 \\
&\frac{d\mathbb{P}^{\text{Target}}(Y \mid T=0, Z=0)}{d\mathbb{P}^{\text{Obs}}(Y \mid T=0, Z=0)} = \overbrace{\omega(0)}^{=0} + \overbrace{\omega(1)}^{=1} \overbrace{\frac{d\mathbb{P}^{\text{True}}(Y(0,1) \mid Co)}{d\mathbb{P}^{\text{True}}(Y(0,0) \mid Co)}}^{\in[1,\infty)} \\
&\quad + \underbrace{\omega(1)}_{=1} \underbrace{\frac{\mathbb{P}^{\text{Obs}}(Nt)}{1-\mathbb{P}^{\text{Obs}}(At)}}_{=1/2} \underbrace{\frac{d\mathbb{P}^{\text{Obs}}(Y \mid T=0, Z=1)}{d\mathbb{P}^{\text{Obs}}(Y \mid T=0, Z=0)}}_{=1} \left\{1 - \underbrace{\frac{d\mathbb{P}^{\text{True}}(Y(0,1) \mid Co)}{d\mathbb{P}^{\text{True}}(Y(0,0) \mid Co)}}_{\in[1,\infty)} \underbrace{\frac{d\mathbb{P}^{\text{True}}(Y(0,0) \mid Nt)}{d\mathbb{P}^{\text{True}}(Y(0,1) \mid Nt)}}_{\in[1,\infty)}\right\} \\
&\frac{d\mathbb{P}^{\text{Target}}(Y \mid T=0, Z=1)}{d\mathbb{P}^{\text{Obs}}(Y \mid T=0, Z=1)} = \underbrace{\omega(1)}_{=1} + \underbrace{\omega(0)}_{=0} \underbrace{\frac{d\mathbb{P}^{\text{True}}(Y(0,0) \mid Nt)}{d\mathbb{P}^{\text{True}}(Y(0,1) \mid Nt)}}_{\in[1/2,\infty)}.
\end{aligned}$$

Therefore we obtain the likelihood ratio equalities $\frac{d\mathbb{P}^{\text{Target}}(Y|T=1,Z=1)}{d\mathbb{P}^{\text{Obs}}(Y|T=1,Z=1)} = \frac{d\mathbb{P}^{\text{Target}}(Y|T=1,Z=0)}{d\mathbb{P}^{\text{Obs}}(Y|T=1,Z=0)} = \frac{d\mathbb{P}^{\text{Target}}(Y|T=0,Z=1)}{d\mathbb{P}^{\text{Obs}}(Y|T=0,Z=1)} = 1$, while under the pointwise approach, $\frac{d\mathbb{P}^{\text{Target}}(Y|T=0,Z=0)}{d\mathbb{P}^{\text{Obs}}(Y|T=0,Z=0)}$ could seemingly be as large as ∞ and as small as 0. Therefore our approach's partial identification bounds on $E[Y \mid T, Z]$ are equal to the the singleton $\{0\}$ for $(1-T)(1-Z) = 0$ and are equal to the domain of Y, $[-1, 1]$, for $T=0, Z=0$.

The pointwise bounds are

$$\left[-2\mathbb{P}^{\text{Obs}}(Co)\mathbb{P}^{\text{True}}(T=0 \mid Z=0), 2\mathbb{P}^{\text{Obs}}(Co)\mathbb{P}^{\text{True}}(T=0 \mid Z=0)\right].$$

Note that $\mathbb{P}^{\text{Obs}}(Co) = 1/2$ and $\mathbb{P}^{\text{True}}(T=0 \mid Z=0) = 1$, so that the partially identified set is $[-1, 1]$. $\square$

D.3 Proofs for Appendix A

The following lemma is used in the proof of Proposition 6.

Lemma 9. Bounds on identified set. *Consider the general identification setting and suppose $Y \mid R \sim \mathcal{N}(\mu(R), \sigma(R)^2)$. Then for all $\epsilon \in (0,1)$ and $e = e(X)$ denoting the propensity score, the identified set is a subset of*

$$\left[E[\lambda(R)Y] \pm E\left[\sigma(R)\lambda(R)(1-\underline{w}(R))\left(\sqrt{2\log(\bar{w}(R))} + \sqrt{2/\pi} + (1-\underline{w}(R))^{\epsilon}\sqrt{1/(e*\epsilon)}\right)\right]\right],$$

where $[a \pm b]$ denotes the closed interval $[a-b, a+b]$ and log *is the natural logarithm.*

Proof of Lemma 9. We show that the upper bound is at most $E[\lambda(R)\mu(R)]$ plus one-half the proposed width; the lower bound follows symmetrically.

Note that as we argue in Theorem 1, the upper bound for the identified set can be written as:

$$\psi^+ = E_{\mathbb{P}^{\text{Obs}}}\left[\underline{w}(R)\lambda(R)Y + (1-\underline{w}(R))CVaR^+_{\tau(R)}(R)\right].$$

In the Normal-residual case, we can write $CVaR^+_{\tau(R)}(R) = \lambda(R)\mu(R) + \lambda(R)\sigma(R)\frac{\phi(q_{\tau(R)})}{1-\tau(R)}$, where q_τ is the τ^{th} quantile of a standard normal distribution and ϕ is the standard normal CDF. By existing arguments (e.g. Pinelis (2019)), the inverse Mills ratio $\phi(q)/(1-\Phi(q))$ has the upper bound $\sqrt{2/\pi} + q$.

Therefore the APO upper bound can be further bounded as:

$$\begin{aligned}\psi^+ &= E_{\mathbb{P}^{\text{Obs}}}\left[\lambda(R)\mu(R) + (1-\underline{w}(R))\sigma(R)\lambda(R)\frac{\phi(q_{\tau(R)})}{1-\tau(R)}\right] \\ &\le E[\lambda(R)\mu(R)] + E\left[(1-\underline{w}(R))\sigma(R)\lambda(R)\left(\sqrt{2/\pi} + q_{\tau(R)}\right)\right].\end{aligned}$$

It remains to bound $q_{\tau(R)}$. By standard arguments, if $S \sim N(0,1)$, then $P(S > s) \leq exp(-s^2/2)$. We substitute $s = q_{\tau(R)}$ to obtain:

$$\begin{aligned} 1 - \tau(R) = P(S > q_{\tau(R)}) &\leq exp(-q_{\tau(R)}^2/2) \\ \log(1 - \tau(R)) &\leq -q_{\tau(R)}^2/2 \\ \sqrt{\log\left(\frac{1}{1-\tau(R))^2}\right)} &\geq q_{\tau(R)}. \end{aligned}$$

Therefore we have bounded the identified set as:

$$\psi^+ \leq E\left[\lambda(R)\mu(R) + (1 - \underline{w}(R))\sigma(R)\lambda(R)\left(\sqrt{2/\pi} + \sqrt{2(1-\underline{w}(R))^2 \log\left(\frac{1}{1-\tau(R))}\right)}\right)\right].$$

Now we bound the second square root, using the identity:

$$\frac{1}{1-\tau(R)} = \frac{\bar{w}(R) - \underline{w}(R)}{1 - \underline{w}(R)} = 1 + \frac{\bar{w}(R) - 1}{1 - \underline{w}(R)}.$$

Therefore:

$$\begin{aligned} &2(1-\underline{w}(R))^2 \log\left(\frac{1}{1-\tau(R))}\right) \\ &\quad = 2(1-\underline{w}(R))^2 \log(\bar{w}(R) - \underline{w}(R)) - 2(1-\underline{w}(R))^{2-2\epsilon}(1-\underline{w}(R))^{2\epsilon}\log(1-\underline{w}(R)) \\ &\quad \leq 2(1-\underline{w}(R))^2 \log(\bar{w}(R)) + \frac{(1-\underline{w}(R))^{2-2\epsilon}}{e * \epsilon}. \end{aligned}$$

So that we now have the bound:

$$\begin{aligned} \psi^+ &\leq E\left[\lambda(R)\mu(R)\right] \\ &\quad + E\left[\sigma(R)\lambda(R)\left((1-\underline{w}(R))\sqrt{2/\pi} + \sqrt{2(1-\underline{w}(R))^2\log(\bar{w}(R)) + \frac{(1-\underline{w}(R))^{2-2\epsilon}}{e*\epsilon}}\right)\right] \\ &\quad \leq E\left[\lambda(R)\mu(R)\right] \\ &\quad + E\left[\sigma(R)\lambda(R)(1-\underline{w}(R))\left(\sqrt{2\log(\bar{w}(R))} + \sqrt{2/\pi} + (1-\underline{w}(R))^{\epsilon}\sqrt{1/(e*\epsilon)}\right)\right]. \end{aligned}$$

Applying the same argument to the symmetric lower bound completes the proof. □

Proof of Proposition 6. We first show the identified set of $E[Y(1)]$ is unbounded if $c > \eta_1$;

the argument is symmetric for the lower bound of $E[Y(0)]$ if $c > \eta_2$.

By the decomposition in the proof of Lemma 9, the upper bound of the identified set of $E[Y(1)]$ is:

$$\psi^+ = E\left[\lambda(R)\mu(R) + (1-\underline{w}(R))CVaR^+_{\tau(R)}(R)\right],$$

where $\lambda(R) = Z/e(X)$, $\underline{w}(R) = Z\frac{e(X)}{e(X)+c}$, and $\mu(R) = E[Y \mid R]$. We can lower bound the upper bound as:

$$\begin{aligned}
\psi^+ &\geq E\left[\lambda(R)\mu(R) + 1\{e(X) < [\eta_1, c]\}(1-\underline{w}(R))CVaR^+_{\tau(R)}(R)\right] \\
&\geq E\left[\lambda(R)\mu(R) + 1\{e(X) \in [\eta_1, c]\}\frac{1}{2}\frac{Z}{e(X)}E[Y-\mu(R) \mid Y, R \geq Q_1(Y \mid R)]\right] \\
&\geq E\left[\lambda(R)\mu(R) + 1\{e(X) \in [\eta_1, c]\}\frac{1}{2}\frac{\eta_1}{c} * \infty\right],
\end{aligned}$$

where the infinite conditional value at risk happens for all X with $\sigma(X, 1) > 0$, which happens almost surely by assumption. Since $\eta_1 = \inf p \mid P(e(X) > p) > 0$ by definition and $\eta_1 > 0$ by assumption, $E[1\{e(X) \in [\eta_1, c]\}\frac{1}{2}\frac{\eta_1}{c} \mid R] > 0$ so that the identified set would be unbounded.

Now suppose $c < \eta_1$ and we wish to show that the identified set is uniformly bounded. Write D as the upper bound of the support of $\frac{\sigma(X,Z)}{e(X)} + \frac{\sigma(X,Z)}{1-e(X)}$, which by assumption is finite. By Lemma 9, the identified set can be upper bounded as follows:

$$\begin{aligned}
\psi^+ &\leq E[\lambda(R)Y] + E\left[\sigma(R)\lambda(R)(1-\underline{w}(R))\left(\sqrt{2\log(\bar{w}(R))} + \sqrt{2/\pi} + (1-\underline{w}(R))^{\epsilon}\sqrt{1/(e*\epsilon)}\right)\right] \\
&\leq E[\lambda(R)Y] + DE\left[Z\sqrt{2\log(\bar{w}(R))} + \sqrt{2/\pi} + 1/e\right],
\end{aligned}$$

where log is the natural logarithm. It only remains to bound $\sqrt{2}E[Z\log(\bar{w}(R))]$, where $\bar{w}(R) = Ze(X)/(e(X)-c) + (1-Z)(1-e(X))/(1-e(X)-c)$. Suppose the propensity density is $f_{e(X)}(p)$ and is upper bounded by $\bar{f}_{e(X)}$:

$$\begin{aligned}
E[Z\log(\bar{w}(R))] &= \int_{\eta_1}^{1-\eta_2} p\log(p/(p-c))f_{e(X)}(p) \\
&\leq \bar{f}_{e(X)}\int_{\eta_1}^{1-\eta_2} log(p/(p-\eta_1))dp \leq -\bar{f}_{e(X)}\int_{\eta_1}^{1-\eta_2}\log(p-\eta_1)dp \\
&\leq -\bar{f}_{e(X)}\int_0^1 log(t)dt = \bar{f}_{e(X)}.
\end{aligned}$$

Therefore an APO upper bound is:

$$\begin{aligned}
\psi^+ &\leq E[\lambda(R)Y] + E\left[\sigma(R)\lambda(R)(1-\underline{w}(R))\left(\sqrt{2\log(\bar{w}(R))} + \sqrt{2/\pi} + (1-\underline{w}(R))^{\epsilon}\sqrt{1/(e*\epsilon)}\right)\right] \\
&\leq E[\lambda(R)Y] + DE\left[\sqrt{2}\bar{f}_{e(X)} + \sqrt{2/\pi} + 1/e\right].
\end{aligned}$$

Since this bound holds symmetrically for the lower bound of $E[Y(0)]$, writing $B = 2DE\left[\sqrt{2}\bar{f}_{e(X)} + \sqrt{2/\pi} + 1/e\right]$ completes the proof. □

Proof of Lemma 5. We drop conditioning on X for exposition. Observe that:

$$\begin{aligned}
&d\mathbb{P}^{\text{Target}}(Y \mid T=0, Z=0) \\
&\quad = \omega(1)\left(\mathbb{P}^{\text{Obs}}(Nt|T=Z=0)d\mathbb{P}^{\text{True}}(Y(0,1)|Nt) + \mathbb{P}^{\text{Obs}}(Co|T=Z=0)d\mathbb{P}^{\text{True}}(Y(0,1)|Co)\right) \\
&\quad + \omega(0)\left(\mathbb{P}^{\text{Obs}}(Nt|T=Z=0)d\mathbb{P}^{\text{True}}(Y(0,0)|Nt) + \mathbb{P}^{\text{Obs}}(Co|T=Z=0)d\mathbb{P}^{\text{True}}(Y(0,0)|Co)\right) \\
&\quad = \omega(1)\left(\mathbb{P}^{\text{Obs}}(Nt|T=Z=0)d\mathbb{P}^{\text{Obs}}(Y \mid T=0, Z=1) + \mathbb{P}^{\text{Obs}}(Co|T=Z=0)d\mathbb{P}^{\text{True}}(Y(0,1)|Co)\right) \\
&\quad + \omega(0)d\mathbb{P}^{\text{Obs}}(Y \mid T=0, Z=0).
\end{aligned}$$

We continue to analyze the unobserved term $d\mathbb{P}^{\text{True}}(Y(0,1)|Co)$, which we factor as:

$$\frac{d\mathbb{P}^{\text{True}}(Y(0,1)|Co)}{d\mathbb{P}^{\text{Obs}}(Y \mid T=0, Z=0)} = \frac{d\mathbb{P}^{\text{True}}(Y(0,1)|Co)}{d\mathbb{P}^{\text{True}}(Y(0,0) \mid Co)}\frac{d\mathbb{P}^{\text{True}}(Y(0,0) \mid Co)}{d\mathbb{P}^{\text{Obs}}(Y \mid T=0, Z=0)}$$

Note that:

$$\frac{d\mathbb{P}^{\text{True}}(Y(0,0)|Co)}{d\mathbb{P}^{\text{True}}(Y(0,0)|Nt)} = \frac{1}{\mathbb{P}^{\text{Obs}}(Co|Z=T=0)}\left(\frac{d\mathbb{P}^{\text{Obs}}(Y \mid T=0, Z=0)}{d\mathbb{P}^{\text{True}}(Y(0,0)|Nt)} - \mathbb{P}^{\text{Obs}}(Nt|Z=T=0)\right),$$

because

$$\frac{d\mathbb{P}^{\text{Obs}}(Y \mid T=0, Z=0)}{d\mathbb{P}^{\text{True}}(Y(0,0)|Nt)} = \mathbb{P}^{\text{Obs}}(Nt|Z=T=0) + \mathbb{P}^{\text{Obs}}(Co|Z=T=0)\frac{d\mathbb{P}^{\text{True}}(Y(0,0)|Co)}{d\mathbb{P}^{\text{True}}(Y(0,0)|Nt)}.$$

Combining the results,

$$\begin{aligned}
&\frac{d\mathbb{P}^{\text{Target}}(Y|T=0, Z=0)}{d\mathbb{P}^{\text{Obs}}(Y|T=0, Z=0)} \\
&= \omega(0) + \omega(1)\mathbb{P}^{\text{Obs}}(Nt|Z=T=0)\frac{d\mathbb{P}^{\text{Obs}}(Y \mid T=0, Z=1)}{d\mathbb{P}^{\text{Obs}}(Y \mid T=0, Z=0)}
\end{aligned}$$

$$+\omega(1)\frac{d\mathbb{P}^{\text{True}}(Y(0,1)|Co)}{d\mathbb{P}^{\text{True}}(Y(0,0)|Co)}\left(1-\mathbb{P}^{\text{Obs}}\left(Nt|Z=T=0\right)\frac{d\mathbb{P}^{\text{True}}(Y(0,0)|Nt)}{d\mathbb{P}^{\text{True}}(Y(0,1)|Nt)}\frac{d\mathbb{P}^{\text{Obs}}(Y\mid T=0,Z=1)}{d\mathbb{P}^{\text{Obs}}(Y\mid T=0,Z=0)}\right),$$

where

$$\mathbb{P}^{\text{Obs}}\left(Nt|Z=T=0\right)=\mathbb{P}^{\text{Obs}}(Nt\mid T(0)=0)=\frac{\mathbb{P}^{\text{Obs}}(Nt)}{\mathbb{P}^{\text{Obs}}(Co)+\mathbb{P}^{\text{Obs}}(Nt)}.$$

Using an analogous argument,

$$\begin{aligned}\frac{d\mathbb{P}^{\text{Target}}(Y|T=1,Z=1)}{d\mathbb{P}^{\text{Obs}}(Y|T=1,Z=1)}&=\omega(1)+\omega(0)\frac{\mathbb{P}^{\text{Obs}}(At)}{\mathbb{P}^{\text{Obs}}(Co)+\mathbb{P}^{\text{Obs}}(At)}\frac{d\mathbb{P}^{\text{Obs}}(Y\mid T=1,Z=0)}{d\mathbb{P}^{\text{Obs}}(Y\mid T=1,Z=1)}\\&+\omega(0)\frac{d\mathbb{P}^{\text{True}}(Y(1,0)|Co)}{d\mathbb{P}^{\text{True}}(Y(1,1)|Co)}\left(1-\frac{\mathbb{P}^{\text{Obs}}(At)}{\mathbb{P}^{\text{Obs}}(Co)+\mathbb{P}^{\text{Obs}}(At)}\frac{d\mathbb{P}^{\text{True}}(Y(1,1)|At)}{d\mathbb{P}^{\text{True}}(Y(1,0)|At)}\frac{d\mathbb{P}^{\text{Obs}}(Y\mid T=1,Z=0)}{d\mathbb{P}^{\text{Obs}}(Y\mid T=1,Z=1)}\right).\end{aligned}$$

We continue to analyze $d\mathbb{P}^{\text{Target}}\left(Y|T=0,Z=1\right)$.

$$\begin{aligned}d\mathbb{P}^{\text{Target}}\left(Y|T=0,Z=1\right)&=\omega(0)d\mathbb{P}^{\text{True}}\left(Y(0,0)\mid Nt\right)+\omega(1)d\mathbb{P}^{\text{True}}\left(Y(0,1)\mid Nt\right)\\&=\omega(0)\frac{d\mathbb{P}^{\text{True}}\left(Y(0,0)\mid Nt\right)}{d\mathbb{P}^{\text{True}}\left(Y(0,1)\mid Nt\right)}d\mathbb{P}^{\text{True}}\left(Y(0,1)\mid Nt\right)+\omega(1)d\mathbb{P}^{\text{Obs}}\left(Y\mid T=0,Z=1\right)\\&=\left(\omega(0)\frac{d\mathbb{P}^{\text{True}}\left(Y(0,0)\mid Nt\right)}{d\mathbb{P}^{\text{True}}\left(Y(0,1)\mid Nt\right)}+\omega(1)\right)d\mathbb{P}^{\text{Obs}}\left(Y\mid T=0,Z=1\right).\end{aligned}$$

Analogously,

$$\begin{aligned}d\mathbb{P}^{\text{Target}}(Y\mid T=1,Z=0)&=\omega(1)d\mathbb{P}^{\text{True}}(Y(1,1)\mid At)+\omega(0)d\mathbb{P}^{\text{True}}(Y(1,0)\mid At)\\&=\omega(1)\frac{d\mathbb{P}^{\text{True}}(Y(1,1)\mid At)}{d\mathbb{P}^{\text{Obs}}(Y\mid T=1,Z=0)}d\mathbb{P}^{\text{Obs}}(Y\mid T=1,Z=0)+\omega(0)d\mathbb{P}^{\text{Obs}}(Y\mid T=1,Z=0)\\&=\left(\omega(1)\frac{d\mathbb{P}^{\text{True}}(Y(1,1)\mid At)}{d\mathbb{P}^{\text{True}}(Y(1,0)\mid At)}+\omega(0)\right)d\mathbb{P}^{\text{Obs}}(Y\mid T=1,Z=0).\end{aligned}$$

$\square$

D.4 Proofs for Appendix B

We will use some notation for the estimation and inference proofs for Appendix B. It is convenient to define the true and estimated residuals from the quantile $S\equiv\lambda(R)Y-Q_{\tau(R)}(\lambda(R)Y\mid R)$ and $\hat{S}\equiv\hat{\lambda}(R)Y-\hat{Q}_{\hat{\tau}(R)}(\hat{\lambda}(R)Y\mid R)$. The following lemma is used in the proof of Proposition 7.

Lemma 10. *Define the infeasible $\hat{T}^*$ corresponding to Proposition 7 as:*

$$\hat{T}^* \equiv \hat{E}\left[\hat{\lambda}(R)Y + \hat{S}a(\underline{\hat{w}}(R), \hat{\bar{w}}(R), S)\right].$$

Then $\hat{T} \geq \hat{T}^*$ deterministically.

Proof of Lemma 10. Define the true and estimated residuals from the quantile $S \equiv \lambda(R)Y - Q_{\tau(R)}(\lambda(R)Y \mid R)$ and $\hat{S} \equiv \hat{\lambda}(R)Y - \hat{Q}_{\hat{\tau}(R)}(\hat{\lambda}(R)Y \mid R)$.

$$\hat{T} - \hat{T}^* = \hat{E}\left[(\hat{\bar{w}}(R) - \underline{\hat{w}}(R))\hat{S}\left(1\{\hat{S} > 0\} - 1\{S > 0\}\right)\right] \geq 0.$$

By sign-matching the cases in which $1\{\hat{S} > 0\} \neq 1\{S > 0\}$, the result holds deterministically. In particular, when $\hat{S} > S$, the region where $1\{\hat{S} > 0\} - 1\{S > 0\} = 1$ is where $\hat{S} > 0$, so $\hat{S}\left(1\{\hat{S} > 0\} - 1\{S > 0\}\right) \geq 0$. Similarly, when $\hat{S} < S$, the region where $1\{\hat{S} > 0\} - 1\{S > 0\} = -1$ is where $\hat{S} < 0$, so $\hat{S}\left(1\{\hat{S} > 0\} - 1\{S > 0\}\right) \geq 0$. □

Proof of Proposition 7. When $\underline{\hat{w}}$ and $\hat{\bar{w}}$ are consistent, we get $\hat{\tau} \xrightarrow{p} \tau$. The first part of the proposition is immediate by applying the continuous mapping theorem and the law of large numbers for iid observations. Even though we have an indicator function, we only have a kink point rather than a discontinuity, so the function is still continuous.

For the second part of the lemma, observe that, using $\hat{T}^*$ in Lemma 10, $\hat{T} \geq \hat{T}^*$. Then, using $Q_\tau(\lambda(R)Y \mid R)$ to denote the τ-th quantile of Y for the given R and using $\hat{Q}_{\hat{\tau}(R)}(\hat{\lambda}(R)Y \mid R)$ to denote the estimated conditional quantile function,

$$\begin{aligned}
\hat{T}^* &= \hat{E}\left[\hat{\lambda}(R)Y + \hat{S}a(\underline{\hat{w}}(R), \hat{\bar{w}}(R), S)\right] \\
&= \hat{E}\left[\hat{\lambda}(R)Y + \left(\hat{\lambda}(R)Y - \hat{Q}_{\hat{\tau}(R)}(\hat{\lambda}(R)Y \mid R)\right) a(\underline{\hat{w}}(R), \hat{\bar{w}}(R), S)\right] \\
&= \hat{E}\left[\hat{\lambda}(R)Y + \left(\hat{\lambda}(R)Y - \hat{Q}_{\hat{\tau}(R)}(\hat{\lambda}(R)Y \mid R)\right)\left(\left(\hat{\bar{w}}(R) - \underline{\hat{w}}(R)\right) 1\{S > 0\} - (1 - \underline{\hat{w}}(R))\right)\right].
\end{aligned}$$

By applying the continuous mapping theorem and the weak law of large numbers,

$$\hat{T}^* = E\left[\lambda(R)Y + \left(\lambda(R)Y - Q_{\tau(R)}(\lambda(R)Y \mid R)\right)\left((\bar{w}(R) - \underline{w}(R))\, 1\{S > 0\} - (1 - \underline{w}(R))\right)\right](1 + o_P(1)).$$

Under our assumptions, the target object can be written as:

$$
\begin{aligned}
T &= E\left[\lambda(R)Y + Sa(\underline{w}(R), \bar{w}(R), S)\right] \\
&= E\left[\lambda(R)Y + (\lambda(R)Y - Q_{\tau(R)}(\lambda(R)Y \mid R))((\bar{w}(R) - \underline{w}(R))\, 1\left\{S > 0\right\} - (1 - \underline{w}(R)))\right].
\end{aligned}
$$

Note that $\hat{T}^* = T(1 + o_P(1))$. Either T is finite, so that $\hat{T}^* = T + o_P(1)$, or T is infinite, so that $\hat{T}^* = T$ with probability tending to one. In either event, $\hat{T}^* - T = o_P(1)$.

□

D.5 Proofs for Appendix C

Proof of Lemma 7. Define $f_{x|M=0}(c) > 0$ to be the derivative of $F_{X|M=0}(x)$ at c. Then we have:

$$
\begin{aligned}
&\mathbb{P}^{\text{True}}(X = c^+ \mid X = c, M = 0) \\
&= \lim_{\varepsilon \to 0^+} \frac{F_{X|M=0}(c+\varepsilon) - F_{X|M=0}(c)}{F_{X|M=0}(c+\varepsilon) - F_{X|M=0}(c-\varepsilon)} \\
&= \lim_{\varepsilon \to 0^+} \frac{F_{X|M=0}(c+\varepsilon) - F_{X|M=0}(c)}{F_{X|M=0}(c+\varepsilon) - F_{X|M=0}(c) + F_{X|M=0}(c) - F_{X|M=0}(c-\varepsilon)} \\
&= \lim_{\varepsilon \to 0^+} \frac{\frac{F_{X|M=0}(c+\varepsilon) - F_{X|M=0}(c)}{\varepsilon}}{\frac{F_{X|M=0}(c+\varepsilon) - F_{X|M=0}(c)}{\varepsilon} + \frac{F_{X|M=0}(c) - F_{X|M=0}(c-\varepsilon)}{\varepsilon}} \\
&= \frac{\lim_{\varepsilon \to 0^+} \frac{F_{X|M=0}(c+\varepsilon) - F_{X|M=0}(c)}{\varepsilon}}{\lim_{\varepsilon \to 0^+} \frac{F_{X|M=0}(c+\varepsilon) - F_{X|M=0}(c)}{\varepsilon} + \lim_{\varepsilon \to 0^-} \frac{F_{X|M=0}(c+\varepsilon) - F_{X|M=0}(c)}{\varepsilon}} \\
&= \frac{f_{x|M=0}(c)}{2 f_{x|M=0}(c)} = 1/2
\end{aligned}
$$

□

Proof of Lemma 8. For our proofs, it is useful to use η_0 instead. We define:

$$
\eta_0 \equiv \mathbb{P}^{\text{True}}(M = 1 \mid X = c) = \mathbb{P}^{\text{Obs}}(X = c^+ \mid X = c)\eta.
$$

Notice that $\mathbb{P}^{\text{Obs}}(X = c^+ \mid X = c) = 1/(2 - \eta)$, because $\mathbb{P}^{\text{Obs}}(X = c^- \mid X = c) = \mathbb{P}^{\text{True}}(M = 0 \mid X = c)/2 = (1 - \eta)\mathbb{P}^{\text{Obs}}(X = c^+ \mid X = c)$. As a result, $\eta_0 = \eta/(2 - \eta)$, $\eta = 2\eta_0/(1 + \eta_0)$, and $\mathbb{P}^{\text{Obs}}(X = c^+ \mid X = c) = (1 + \eta_0)/2$.

Also notice that in Equation (6), the constraints on manipulation probabilities are defined relative to:

$$\frac{\mathbb{P}^{\text{True}}(M = 1 \mid X = c)}{\mathbb{P}^{\text{True}}(M = 0 \mid X = c)} = \frac{\eta_0}{1 - \eta_0} = \frac{\eta}{2(1 - \eta)}. \tag{11}$$

We begin with $Y(1)$. Since $\mathbb{Q}$ produces the observable distribution $\mathbb{P}^{\text{Obs}}(X = c^+ \mid X = c) d\mathbb{P}^{\text{Obs}}(Y \mid X = c^+)$ and $\mathbb{P}^{\text{Obs}}(X = c^+ \mid X = c) = (1 + \eta_0)/2$, we have:

$$\begin{aligned}
\mathbb{Q}(X = c^+ \mid X = c) d\mathbb{Q}(Y(1) \mid X = c^+) &= d\mathbb{Q}(Y(1) \mid X = c)\mathbb{Q}(X = c^+ \mid X = c, Y(1)) \\
&= d\mathbb{Q}(Y(1) \mid X = c) \left(1 + q_1(Y(1))\right)/2 \\
d\mathbb{Q}(Y(1) \mid X = c) &= \frac{2\mathbb{P}^{\text{Obs}}(X = c^+ \mid X = c)}{1 + q_1(Y(1))} d\mathbb{Q}(Y(1) \mid X = c^+)
\end{aligned}$$

We can also derive the probability of a treated observation being manipulated under $\mathbb{Q}$ through Bayes' Rule:

$$\begin{aligned}
&\mathbb{Q}(M = 0 \mid Y(1), X = c^+) \\
&\quad = \frac{d\mathbb{Q}(Y(1) \mid X = c)\mathbb{Q}(M = 0 \mid X = c, Y(1))\mathbb{Q}(X = c^+ \mid X = c, Y(1), M = 0)}{\mathbb{Q}(X = c^+ \mid X = c) d\mathbb{Q}(Y(1) \mid X = c^+)} \\
&\quad = \frac{d\mathbb{Q}(Y(1) \mid X = c) * (1 - q_1(Y(1)))/2}{d\mathbb{Q}(Y(1) \mid X = c) \left(1 + q_1(Y(1))\right)/2} \\
&\quad = \frac{1 - q_1(Y(1))}{1 + q_1(Y(1))}
\end{aligned}$$

As a result:

$$\begin{aligned}
d\mathbb{Q}(Y(1) \mid X = c, M = 0) &= d\mathbb{Q}(Y(1) \mid X = c^+, M = 0) \\
&= \frac{d\mathbb{Q}(Y(1) \mid X = c^+)\mathbb{Q}(M = 0 \mid X = c^+, Y(1))}{\mathbb{Q}(M = 0 \mid X = c^+)} \\
&= \frac{\frac{1 - q_1(Y(1))}{1 + q_1(Y(1))}}{1 - \eta} d\mathbb{P}^{\text{True}}(Y(1) \mid X = c^+)
\end{aligned}$$

This is our first equality.

We now turn our attention to $Y(0)$. By a similar observed-untreated-outcome argument, we have:

$$\mathbb{Q}(X = c^- \mid X = c) d\mathbb{Q}(Y(0) \mid X = c^-)$$

$$
\begin{aligned}
&= d\mathbb{Q}(Y(0) \mid X = c)\mathbb{Q}(X = c^- \mid X = c, Y(0)) \\
&= d\mathbb{Q}(Y(0) \mid X = c)\mathbb{Q}(M = 0 \mid X = c, Y(0))\mathbb{Q}(X = c^- \mid X = c, Y(0), M = 0) \\
&= d\mathbb{Q}(Y(0) \mid X = c)(1 - q_0(Y(0)))/2
\end{aligned}
$$

Since $\mathbb{Q}(X = c^- \mid X = c) = \mathbb{P}^{\text{Target}}(X = c^- \mid X = c) = \frac{1-\eta_0}{2}$, we can then obtain:

$$
d\mathbb{Q}(Y(0) \mid X = c) = \frac{1 - \eta_0}{1 - q_0(Y(0))} d\mathbb{Q}(Y(0) \mid X = c^-)
$$

We can also split up $d\mathbb{Q}(Y(0) \mid X = c)$ as:

$$
\begin{aligned}
d\mathbb{Q}(Y(0) \mid X = c) &= \eta_0 d\mathbb{Q}(Y(0) \mid X = c, M = 1) + (1 - \eta_0) d\mathbb{Q}(Y(0) \mid X = c, M = 0) \\
&= \eta_0 d\mathbb{Q}(Y(0) \mid X = c, M = 1) + (1 - \eta_0) d\mathbb{Q}(Y(0) \mid X = c^-)
\end{aligned}
$$

So that we can combine terms to obtain:

$$
\begin{aligned}
d\mathbb{Q}(Y(0) \mid X = c, M = 1) &= \frac{1 - \eta_0}{\eta_0} \frac{q_0(Y(0))}{1 - q_0(Y(0))} d\mathbb{Q}(Y(0) \mid X = c^-) \\
\frac{1 - \eta_0}{\eta_0} &= \frac{2(1 - \eta)}{\eta} \\
d\mathbb{Q}(Y(0) \mid X = c, M = 1) &= \frac{2(1 - \eta)}{\eta} \frac{q_0(Y(0))}{1 - q_0(Y(0))} d\mathbb{Q}(Y(0) \mid X = c^-)
\end{aligned}
$$

Which is the final equality after substituting in $d\mathbb{Q}(Y(0) \mid X = c^-) = d\mathbb{P}^{\text{True}}(Y(0) \mid X = c^-)$. $\square$