throughout this paper and present concrete examples of support restrictions. I review the existing method of identifying the DTE given marginal distributions without support restrictions to demonstrate its limits in the presence of support restrictions. I also briefly discuss the optimal transportation approach to describe the key idea of my identification strategy. Section 3 formally characterizes the identification region of the DTE under general support restrictions and derives informative bounds for economic examples from the characterization. Section 4 provides numerical examples to assess the informativeness of my new bounds and analyzes sources of identification gains. Section 5 illustrates the usefulness of these bounds by applying DTE bounds derived in Section 3 to an empirical analysis of the impact distribution of smoking on infant birth weight. Section 6 concludes and discusses interesting extensions.

2 Basic Setup, DTE Bounds and Optimal Transportation Approach

In this section, I present the potential outcomes setup that this study is based on, the notation, and the assumptions used throughout this study. I demonstrate that the bounds on the DTE established without support restrictions are not the best possible bounds in the presence of support restrictions. Then I propose a new method to derive sharp bounds on the DTE based on the optimal transportation framework.

2.1 Basic Setup

The setup that I consider is as follows: the econometrician observes a realized outcome variable Y and a treatment participation indicator D for each individual, where $D = 1$ indicates treatment participation while $D = 0$ nonparticipation. An observed outcome Y can be written as $Y = DY_1 + (1 - D)Y_0$. Only Y_1 is observed for the individual who takes the treatment while only Y_0 is observed for the individual who does not take the treatment, where Y_0 and Y_1 are the potential outcome without and with treatment, respectively. Treatment effects Δ are defined as $\Delta = Y_1 - Y_0$ the difference of potential outcomes. The objective of this study is to identify the distribution function of treatment effects $F_\Delta(\delta) = \Pr(Y_1 - Y_0 \leq \delta)$ from observed pairs (Y, D) for fixed $\delta \in \mathbb{R}$.

To avoid notational confusion, I differentiate between the *distribution* and the *distribution function*. Let μ_0, μ_1 and π denote marginal distributions of Y_0 and Y_1, and their joint distribution, respectively. That is, for any measurable set A_d in $\mathbb{R}$, $\mu_d(A_d) = \Pr\{Y_d \in A_d\}$ for $d \in \{0, 1\}$ and $\pi(A) = \Pr\{(Y_0, Y_1) \in A\}$ for any measurable set A in $\mathbb{R}^2$. In addition, let F_0, F_1 and F denote marginal distribution functions of Y_0 and Y_1, and their joint distribution function, respectively. That is, $F_d(y_d) = \mu_d((-\infty, y_d])$ and