As shown in Figure 6, $A^D = \phi$ for $A = \phi$, and $A^D = [a + \delta, \infty)$ for $A = [a, \infty)$ with $a \in (-\infty, \infty)$. Then, $\mu_0(A) - \mu_1(A^D) = 0$ for $A = \phi$, while $\mu_0(A) - \mu_1(A^D) = F_1(a + \delta) - F_0(a)$ for $A = [a, \infty)$. Therefore, the RHS in (6) reduces to

$$\sup_{a \in \mathbb{R}} \max \left[F_1(a + \delta) - F_0(a), 0 \right],$$

which is equal to the Makarov lower bound. One can derive the Makarov upper bound in the same way.

Now consider the support restriction $\Pr((Y_0, Y_1) \in C) = 1$. Note that this restriction is linear in the entire joint distribution π, since it can be rewritten as $\int \mathbf{1}_C(y_0, y_1)\, d\pi = 1$. The linearity makes it possible to handle this restriction with penalty. In particular, since support restrictions hold with probability one, the corresponding penalty is infinite. Therefore, one can embed $1 - \mathbf{1}_C(y_0, y_1)$ into the cost function with an infinite multiplier $\lambda = \infty$ as follows:

$$\inf_{\pi \in \Pi(\mu_0, \mu_1)} \int \left\{ \mathbf{1}\{y_1 - y_0 < \delta\} + \lambda \left(1 - \mathbf{1}_C(y_0, y_1)\right) \right\} d\pi \tag{9}$$

The minimization problem (9) is well defined with $\lambda = \infty$ as noted in Remark 1. Note that for $\lambda = \infty$, any joint distribution which violates the restriction $\Pr((Y_0, Y_1) \in C) = 1$ would cause infinite total costs in (9) and it is obviously excluded from the potential optimal joint distribution candidates. The optimal joint distribution should thus satisfy the restriction $\Pr((Y_0, Y_1) \in C) = 1$ to avoid infinite costs by not permitting any positive probability density for the region outside of the set C. Similarly, the upper bound on the DTE is written as

$$\begin{aligned} &\sup_{\pi \in \Pi(\mu_0, \mu_1)} \int \left\{ \mathbf{1}\{y_1 - y_0 \le \delta\} - \lambda \left(1 - \mathbf{1}_C(y_0, y_1)\right) \right\} d\pi \\ &= 1 - \inf_{\pi \in \Pi(\mu_0, \mu_1)} \int \left\{ \mathbf{1}\{y_1 - y_0 > \delta\} + \lambda \left(1 - \mathbf{1}_C(y_0, y_1)\right) \right\} d\pi. \end{aligned} \tag{10}$$

To the best of my knowledge, this is the first paper that allows for $\{0, 1, \infty\}$-valued costs. Although the econometrics literature based on the optimal transportation approach has used Lemma 3 for $\{0, 1\}-$valued costs, the problem (9) cannot be solved using Lemma 3. In the next section, I develop a dual representation for (9) in order to characterize sharp bounds on the DTE.

3 Main Results

This section characterizes sharp DTE bounds under general support restrictions by developing a dual representation for problems (9) and (10). I use this characterization to derive sharp DTE bounds for various