Multiple Treatments with Strategic Substitutes*

Jorge F. Balat
Department of Economics
University of Texas at Austin
jbalat@utexas.com

Sukjin Han
Department of Economics
University of Bristol
vincent.han@bristol.ac.uk

First Draft: February 23, 2016
This Draft: October 30, 2020

Abstract

We develop an empirical framework to identify and estimate the effects of treatments on outcomes of interest when the treatments are the result of strategic interaction (e.g., bargaining, oligopolistic entry, peer effects). We consider a model where agents play a discrete game of complete information and strategic substitutability, whose equilibrium actions (i.e., binary treatments) determine a post-game outcome in a nonseparable model with endogeneity. Due to the simultaneity in the first stage, the model as a whole is incomplete and the selection process fails to exhibit the conventional monotonicity. Without imposing parametric restrictions or large support assumptions, this poses challenges in recovering treatment parameters. To address these challenges, we establish a monotonic pattern of the equilibria in the first-stage game in terms of the number of treatments selected. Based on this finding, we derive bounds on the average treatment effects (ATE's) under nonparametric shape restrictions and the existence of excluded exogenous variables. We show that the instrument variation that compensates strategic substitution helps solve the multiple equilibria problem. We apply our method to data on airlines and air pollution in cities in the U.S. We find that (i) the causal effect of each airline on pollution is positive, and (ii) the effect is increasing in the number of firms but at a decreasing rate.

*The authors are grateful to Tim Armstrong, Steve Berry, Andrew Chesher, Áureo de Paula, Phil Haile, Karam Kang, Juhyun Kim, Yuichi Kitamura, Simon Lee, Konrad Menzel, Francesca Molinari, Adam Rosen, Azeem Shaikh, Jesse Shapiro, Dean Spears, Ed Vytlacil, Haiqing Xu, and participants in the 2016 Texas Econometrics Camp, the 2016 North American Summer Meeting of the Econometric Society, Interactions Conference 2016 at Northwestern University, the 2017 Conference on Econometrics and Models of Strategic Interactions (CeMMAP UCL and Vanderbilt), International Association for Applied Econometrics 2017 at Sapporo, the 27th Annual Meeting of the Midwest Econometrics Group at Texas A&M, and seminars at Brown, BU, SNU, Yale, UBC, UNC, and Xiamen for helpful comments and discussions.

JEL Numbers: C14, C31, C36, C57

Keywords: Heterogeneous treatment effects, strategic interaction, endogenous treatments, average treatment effects, multiple equilibria.

1 Introduction

We develop an empirical framework to identify and estimate the heterogeneous effects of treatments on outcomes of interest, where the treatments are the result of agents' strategic interaction (e.g., bargaining, oligopolistic entry, decisions in the presence of peer effects or strategic effects). Treatments are determined as an equilibrium of a game and these strategic decisions of players endogenously affect common or player-specific outcomes. For example, one may be interested in the effects of entry of newspapers on local political behavior, entry of carbon-emitting companies on local air pollution and health outcomes, the presence of potential entrants in nearby markets on pricing or investment decisions of incumbents, the exit decisions of large supermarkets on local health outcomes, or the provision of limited resources when individuals make participation decisions under peer effects and their own gains from the treatment.[1] As reflected in some of these examples, our framework allows us to study the *externalities of strategic decisions*, such as societal outcomes resulting from firm behavior. Ignoring strategic interaction in the treatment selection process may lead to biased, or at least less informative, conclusions about the effects of interest.

We consider a model in which agents play a discrete game of complete information and strategic substitutability, whose equilibrium actions (i.e., a profile of binary endogenous treatments) determine a post-game outcome in a nonseparable model with endogeneity. We are interested in the various treatment effects of this model. In recovering these parameters, the setting of this study poses several challenges. First, the first-stage game posits a structure in which binary dependent variables are simultaneously determined in threshold crossing models, thereby, making the model, as a whole, *incomplete*. This is related to the problem of multiple equilibria in the game. Second, due to this simultaneity, the selection process for each treatment in the profile does not exhibit the conventional monotonic property à la Imbens and Angrist (1994). Furthermore, we want to remain flexible with other components of the model. That is, we make no assumptions on the joint distributions of the unobservables nor parametric restrictions on the players' payoff functions and how

[1] The entry and pollution is our leading example introduced in Section 2; the other examples are discussed in detail in Appendix A.

treatments affect the outcome. In addition, we do not impose any arbitrary equilibrium selection mechanism to deal with the multiplicity of equilibria, nor require that players be symmetric. In nonparametric models with multiplicity and/or endogeneity, identification may be achieved using excluded instruments with large support. Although such a strong requirement can be met in practice, estimation and inference can still be problematic (Andrews and Schafgans (1998), Khan and Tamer (2010)). Thus, we avoid such assumptions for instruments and other exogenous variables.

The first contribution of this study is to establish that under strategic substitutability, regions that predict the equilibria of the treatment selection process in the first-stage game can present a monotonic pattern in terms of the number of treatments selected.[2] The second contribution of this study is to show, after restoring the *generalized monotonicity* in the selection process, how the model structure and the data can provide information about treatment parameters, such as the average treatment effects (ATE's). We first establish the bounds on the ATE and other related parameters with possibly discrete instruments. We also show that tighter bounds on the ATE can be obtained by introducing (possibly discrete) exogenous variables excluded from the first-stage game. This is especially motivated when the outcome variable is affected by externalities generated by the players. We can derive sharp bounds as long as the outcome variable is binary. To deal with the multiple equilibria problem, we assume that instruments vary sufficiently to offset the effect of strategic substitutability. We provide a simple testable implication for the existence of such instrument variation in the case of mutually independent payoff unobservables. This requirement on variation is qualitatively different and substantially weaker than a typical large support assumption. A marked feature of our analyses is that for the sharp bounds on the ATE, player-specific instruments are not necessary.

Partial identification in single-agent nonparametric triangular models with binary endogenous variables has been studied in Shaikh and Vytlacil (2011) and Chesher (2005), among others. Shaikh and Vytlacil (2011) provide bounds on the ATE in this setting. In a more general model, Vytlacil and Yildiz (2007) achieve point identification with an exogenous variable that is excluded from the selection equation and has a large support. Our bound analysis builds on these papers, but we allow for multi-agent strategic interaction as a key component of the model. Some studies have

[2] To estimate payoff parameters, Berry (1992) partly characterizes equilibrium regions. To calculate the bounds on these parameters, Ciliberto and Tamer (2009) simulate their moment inequalities model that are implied by the shape of these regions, especially the regions for multiple equilibria. While their approaches are sufficient for their analyses, full analytical results are critical for the identification analysis in this current study.

extended a single-treatment model to a multiple-treatment setting (e.g., Heckman et al. (2006), Jun et al. (2011)), but their models maintain monotonicity in the selection process and none of them allow simultaneity among the multiple treatments resulting from agents' interaction, as we do in this study.

In interesting recent work, Heckman and Pinto (2018), and Lee and Salanié (2018) extend the monotonicity of the selection process in multi-valued treatments settings. Lee and Salanié (2018) consider more general non-monotonicity than Heckman and Pinto (2018) and do mention entry games as one example of the treatment selection processes they allow. However, they assume known payoffs and bypass the multiplicity of equilibria by assuming a threshold-crossing equilibrium selection mechanism, both of which we do not assume in this study. In addition, Lee and Salanié (2018)'s focus is on the identification of marginal treatment effects with continuous instruments. In another related work, Chesher and Rosen (2017) consider a wide class of generalized instrumental variable models in which our model may fall and propose a systematic method of characterizing sharp identified sets for admissible structures. This present study's characterization of the identified sets is analytical, which help investigate how the identification is related to exogenous variation in the model and to the equilibrium characterization in the treatment selection. Also, calculating the bounds on the treatment parameters using their approach involves projections of identified sets that may require parametric restrictions. Lastly, Han (2019a,b) consider identification of dynamic treatment effects and optimal treatment regimes in a nonparametric dynamic model, in which the dynamic relationship causes non-monotonicity in the determination of each period's outcome and treatment.

Without triangular structures, Manski (1997), Manski and Pepper (2000) and Manski (2013) also propose bounds on the ATE with multiple treatments under various monotonicity assumptions, including an assumption on the sign of the treatment response. We take an alternative approach that is more explicit about treatments interaction while remaining agnostic about the direction of the treatment response. Our results suggest that provided there exist exogenous variation excluded from the selection process, the bounds calculated from this approach can be more informative than those from their approach.

Identification in models for binary games with complete information has been studied in Tamer

(2003), Ciliberto and Tamer (2009), and Bajari et al. (2010), among others.[3] This present study contributes to this literature by considering an evaluation problem with treatments generated by binary games and post-game outcomes that are often not of players' direct concern. As related work that considers post-game outcomes, Ciliberto et al. (2018) introduce a model in which firms make simultaneous decisions of entry and pricing upon entry. Consequently, their model can be seen as a multi-agent extension of a sample selection model. On the other hand, the model considered in this study is a multi-agent extension of a model for endogenous treatments. As counterfactual policy analyses are the main goal, Ciliberto and Tamer (2009) and Ciliberto et al. (2018) recover model primitives as their parameters of interest and impose parametric assumptions. In contrast, our parameters of interest are treatment effects as functionals of the primitives (but excluding the game parameters), and thus, allow our model to remain essentially nonparametric. In addition, a different approach to partial identification under multiplicity is employed, as their approach is not applicable to the particular setting of this study even if we are willing to assume a known distribution for the unobserved payoff types.

To demonstrate the applicability of our method, we take the bounds we propose to data on airline market structure and air pollution in cities in the U.S. Aircrafts and airports land operations are a major source of emissions, and thus, quantifying the causal effect of air transport on pollution is of importance to policy makers. We explicitly allow the market structure to be determined endogenously as the outcome of an entry game. We do not impose any structure on how airline competition affects pollution and allow for heterogenous effects across firms. To implement our application, we combine data from two sources. The first contains airline information from the Department of Transportation, which we use to construct a dataset of airlines' presence in each market. We then merge it with air pollution data in each airport from air monitoring stations compiled by the Environmental Protection Agency. In our preferred specification, our outcome variable is a binary measure of the level of particulate matter in the air.

We consider three sets of ATE exercises to investigate different aspects of the relationship between market structure and pollution. The first quantifies the effects of each airline operating as a monopolist compared to a situation in which the market is not served by any airline. We find that the effect of each airline on pollution is positive and statistically significant and that the effect is

[3] See also Galichon and Henry (2011) and Beresteanu et al. (2011) for a more general setup that includes complete information games as an example.

heterogeneous across airlines. The second set of exercises examines the ATE's of all potential market structures on pollution. We find that the probability of high pollution is increasing with the number of airlines in the market, but at a decreasing rate. Finally, the third set of exercises quantifies the ATE of a single airline under all potential configurations of the market. We observe that in all cases, Delta entering a market has a positive effect on pollution and this effect is decreasing with the number of rivals. The results from the last two set of exercises are consistent with the results of a Cournot oligopolistic model in which incumbents *accommodate* new entrants by reducing the quantity they produce.

This paper is organized as follows. Section 2 summarizes the analysis of this study using a stylized example. Section 3 presents a general theory. Section 3.1 introduces the model and the parameters of interest; Section 3.2 presents the generalized monotonicity for equilibrium regions for many players; and Section 3.4 delivers the partial identification results of this study. Section 4 the empirical application on airlines and pollution. In the Appendix, Section A provides more examples to which our setup can be applied. Section B presents a numerical illustration. Section C contains discussions and four extensions of our main results. Finally, Section D collects the proofs of theorems and lemmas.

2 A Stylized Example

We first illustrate the main results of this study with a stylized example. Suppose we are interested in the effects of airline competition on local air quality (or health). Let Y_i denote the binary indicator of air pollution in market i. Only for illustration, we assume there are two potential airlines. Let $D_{1,i}$ and $D_{2,i}$ be binary variables that indicate the decisions to enter market i by Delta and United, respectively. We allow $D_{1,i}$ and $D_{2,i}$ to be correlated with unobserved characteristics of the local market that affect Y_i. Moreover, we allow $D_{1,i}$ and $D_{2,i}$ to be outcomes from multiple equilibria. The endogeneity and the presence of multiple equilibria are our key challenges in this study.

Let $Y_i(d_1, d_2)$ be the potential air quality had Delta and United's decisions been $(D_1, D_2) = (d_1, d_2)$; for example, $Y_i(1, 1)$ is the potential air quality from duopoly, $Y_i(1, 0)$ is with Delta being a monopolist, and so on. Let X_i be a vector of market characteristics that affect Y_i. Our parameter of interest is the ATE, $E[Y_i(d_1, d_2) - Y_i(d'_1, d'_2)|X_i = x]$, which captures the effect of market structure on pollution. One interesting ATE is $E[Y_i(1, d_2) - Y_i(0, d_2)|X_i = x]$ for each d_2, where we can learn the interaction effects of treatments, e.g., how much the average effect of Delta's entry is affected by United's entry: $E\left[Y_i(1,1)-Y_i(0,1)\right]-E\left[Y_i(1,0)-Y_i(0,0)\right]$ (suppressing X_i). In our empirical application (Section 4), we consider this and other related parameters in a more realistic model, where there are more than two airlines.

Let $Z_{1,i}$ and $Z_{2,i}$ be cost shifters for Delta and United, respectively, which serve as instruments. As a benchmark, we first consider naive bounds analogous to Manski (1990) using excluded instruments which satisfy

$$Y_i(d_1,d_2)\perp(Z_{1,i},Z_{2,i})|X_i \tag{2.1}$$

for all (d_1,d_2). To simplify notation, we suppress the index i henceforth, let $\boldsymbol{D}\equiv(D_1,D_2)$ and $\boldsymbol{Z}\equiv(Z_1,Z_2)$, and write $E[\cdot|w]\equiv E[\cdot|W=w]$ for a generic r.v. W. As an illustration, we focus on calculating bounds on $E[Y(1,1)|X=x]$. Note that

$$\begin{aligned}E[Y(1,1)|x]=E[Y(1,1)|\boldsymbol{z},x]=E[Y|\boldsymbol{D}=(1,1),\boldsymbol{z},x]\Pr[\boldsymbol{D}=(1,1)|\boldsymbol{z},x]\\+\sum_{\boldsymbol{d}'\neq(1,1)}E[Y(1,1)|\boldsymbol{D}=\boldsymbol{d}',\boldsymbol{z},x]\Pr[\boldsymbol{D}=\boldsymbol{d}'|\boldsymbol{z},x],\end{aligned} \tag{2.2}$$

where the first equality is by (2.1). Manski-type bounds can be obtained by observing that the counterfactual term $E[Y(1,1)|\boldsymbol{D}=\boldsymbol{d}',\boldsymbol{z},x]=\Pr[Y(1,1)=1|\boldsymbol{D}=\boldsymbol{d}',\boldsymbol{z},x]$ is bounded above by one and below by zero. By further using the variation in $\boldsymbol{Z}$, which is excluded from $Y(1,1)$, the lower and upper bounds on $E[Y(1,1)|x]$ can be written as

$$\begin{aligned}L_{Manski}(x)&\equiv\sup_{\boldsymbol{z}\in\mathcal{Z}}\Pr[Y=1,\boldsymbol{D}=(1,1)|\boldsymbol{z},x],\\U_{Manski}(x)&\equiv\inf_{\boldsymbol{z}\in\mathcal{Z}}\left\{\Pr[Y=1,\boldsymbol{D}=(1,1)|\boldsymbol{z},x]+1-\Pr[\boldsymbol{D}=(1,1)|\boldsymbol{z}]\right\}.\end{aligned}$$

The goal of our analysis is to derive tighter bounds than $L_{Manski}(x)$ and $U_{Manski}(x)$ by introducing further assumptions motivated by economic theory.

To illustrate, we introduce the following semi-triangular model with linear indices. From the next section, we generalize this model and introduce fully nonparametric models that allow continuous

Y. Consider

$$Y = 1[\mu_1 D_1 + \mu_2 D_2 + \beta X \geq \epsilon], \tag{2.3}$$

$$D_1 = 1[\delta_2 D_2 + \gamma_1 Z_1 \geq U_1], \tag{2.4}$$

$$D_2 = 1[\delta_1 D_1 + \gamma_2 Z_2 \geq U_2], \tag{2.5}$$

where (ϵ, U_1, U_2) are continuously distributed unobservables that can be arbitrarily correlated, (U_1, U_2) are uniform, and assume

$$(\epsilon, U_1, U_2) \perp (Z_1, Z_2)|X, \tag{2.6}$$

$$\delta_1 < 0 \text{ and } \delta_2 < 0, \tag{2.7}$$

$$sgn(\mu_1) = sgn(\mu_2). \tag{2.8}$$

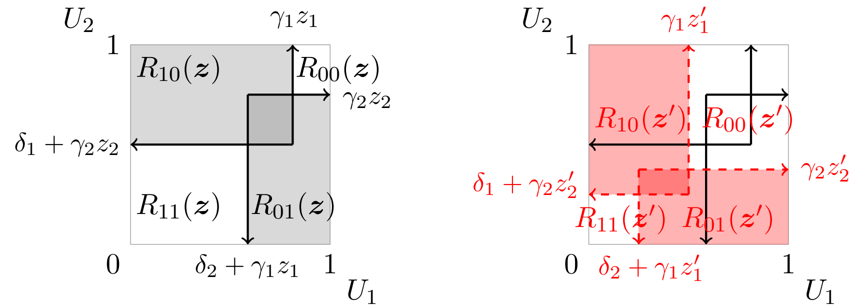
Note that (2.6) replaces (2.1), (2.7) assumes strategic substitutability, and (2.8) is plausible in the current example of air quality and entry. Owing to the first stage simultaneity, (2.4)–(2.5), the model is *incomplete*, i.e., the model primitives and the covariates do not uniquely predict $(Y, \boldsymbol{D})$. In this model, we are *not* interested in the players' payoff parameters (δ_{-s}, γ_s) for $s = 1, 2$, individual parameters (μ_1, μ_2, β) that generate the outcome, nor distributional parameters. Instead, we are interested in the ATE as a function of (μ_1, μ_2, β). This is in contrast to Ciliberto et al. (2018), where payoff and pricing parameters are direct parameters of interest, and thus, our identification question and strategy (especially how we deal with multiple equilibria) are different from theirs.

For two realizations $\boldsymbol{z}, \boldsymbol{z}'$ of $\boldsymbol{Z}$, say low and high entry cost for both airlines, define

$$h(\boldsymbol{z}, \boldsymbol{z}', x) \equiv \Pr[Y = 1|\boldsymbol{z}, x] - \Pr[Y = 1|\boldsymbol{z}', x], \tag{2.9}$$

$$h_{\boldsymbol{d}}(\boldsymbol{z}, \boldsymbol{z}', x) \equiv \Pr[Y = 1, \boldsymbol{D} = \boldsymbol{d}|\boldsymbol{z}, x] - \Pr[Y = 1, \boldsymbol{D} = \boldsymbol{d}|\boldsymbol{z}', x] \tag{2.10}$$

for $d \in \{(0,0), (1,0), (0,1), (1,1)\} \equiv \mathcal{D}$. We show that (2.9)–(2.10) recover useful information about the outcome index function $(\mu_1 D_1 + \mu_2 D_2 + \beta X)$, which in turn is helpful in constructing bounds



(a) When $\boldsymbol{Z} = \boldsymbol{z}$ (b) When $\boldsymbol{Z} = \boldsymbol{z}'$

Figure 1: Change in Equilibrium Regions with Compensating Strategic Substitutability.

on the ATE. Note that

$$
\begin{aligned}
h(\boldsymbol{z}, \boldsymbol{z}', x) &= h_{11}(\boldsymbol{z}, \boldsymbol{z}', x) + h_{10}(\boldsymbol{z}, \boldsymbol{z}', x) + h_{01}(\boldsymbol{z}, \boldsymbol{z}', x) + h_{00}(\boldsymbol{z}, \boldsymbol{z}', x) \\
&= \Pr[Y = 1, \boldsymbol{D} = (1,1)|\boldsymbol{z}, x] - \Pr[Y = 1, \boldsymbol{D} = (1,1)|\boldsymbol{z}', x] \\
&+ \Pr[Y = 1, \boldsymbol{D} = (1,0)|\boldsymbol{z}, x] - \Pr[Y = 1, \boldsymbol{D} = (1,0)|\boldsymbol{z}', x] \\
&+ \Pr[Y = 1, \boldsymbol{D} = (0,1)|\boldsymbol{z}, x] - \Pr[Y = 1, \boldsymbol{D} = (0,1)|\boldsymbol{z}', x] \\
&+ \Pr[Y = 1, \boldsymbol{D} = (0,0)|\boldsymbol{z}, x] - \Pr[Y = 1, \boldsymbol{D} = (0,0)|\boldsymbol{z}', x], \qquad (2.11)
\end{aligned}
$$

where $\boldsymbol{D} = (1,0)$ and $(0,1)$ are the airlines' decisions that may arise as multiple equilibria. The increase in cost (from $\boldsymbol{z}$ to $\boldsymbol{z}'$) will make the operation of these airlines less profitable in some markets, depending on the values of the unobservables $\boldsymbol{U} = (U_1, U_2)$. This will result in a change in the market structure in those markets. Specifically, markets "on the margin" may experience one of the following changes in structure as cost increases: (a) from duopoly to Delta-monopoly; (b) from duopoly to United-monopoly; (c) from Delta-monopoly to no entrant; (d) from United-monopoly to no entrant; and (e) from duopoly to no entrant. These changes are depicted in Figure 1, where each $R_{d_1,d_2}(\boldsymbol{z})$ denotes the maximal region that predicts (d_1, d_2), given $\boldsymbol{Z} = \boldsymbol{z}$.[4]

These changes (a)–(e) are a consequence of the monotonic pattern of equilibrium regions, which is formally established in Theorem 3.1 of Section 3.2.

In general, besides these five scenarios, there may be markets that used to be Delta-monopoly but become United-monopoly and vice versa, i.e., markets that exhibit *non-monotonic* behaviors; see Section C.2 for details. Owing to possible multiple equilibria, we are agnostic about these

[4]See Section D.1 in the Appendix for a formal definition. The figure is drawn in a way that γ_1 and γ_2 are negative.

latter types of changes except in extreme cases (i.e., one equilibrium is selected with probability one). We generally do not know the equilibrium selection mechanism in play, much less about how such mechanism changes as cost $\boldsymbol{Z}$ changes. The key idea in this study is to overcome the non-monotonicity by shifting the cost sufficiently so that there is no market that switches from one monopoly to another. We show that the shift in cost that compensates the strategic substitutability does just that, as is depicted in Figure 1. In this figure, we assume $\delta_2 + \gamma_1 z_1 > \gamma_1 z_1'$ and $\delta_1 + \gamma_2 z_2 > \gamma_2 z_2'$. In other words, we assume compensating strategic substitutability: $|\gamma_s(z_s' - z_s)| \geq |\delta_{-s}|$ for $s = 1, 2$. Importantly, we do not require infinite variation in $\boldsymbol{Z}$.[5] In fact, we show that the compensating strategic substitutability is implied by the following condition, which can be tested using the data: there exist $\boldsymbol{z}, \boldsymbol{z}' \in \mathcal{Z}$ such that $\Pr[\boldsymbol{D} = (1,1)|\boldsymbol{z}] + \Pr[\boldsymbol{D} = (0,0)|\boldsymbol{z}'] > 2 - \sqrt{2}$. Suppose $\boldsymbol{z}, \boldsymbol{z}'$ satisfy the compensating strategic substitutability. Then, by (2.6), we can derive from (2.11) that (suppressing $X = x$ for simplicity)

$$
\begin{aligned}
h(\boldsymbol{z}, \boldsymbol{z}') = &\Pr[\epsilon \leq \mu_1 + \mu_2, \boldsymbol{U} \in \Delta_a \cup \Delta_b \cup \Delta_e] \\
&- \Pr[\epsilon \leq \mu_1, \boldsymbol{U} \in \Delta_a] + \Pr[\epsilon \leq \mu_1, \boldsymbol{U} \in \Delta_c] \\
&- \Pr[\epsilon \leq \mu_2, \boldsymbol{U} \in \Delta_b] + \Pr[\epsilon \leq \mu_2, \boldsymbol{U} \in \Delta_d] \\
&- \Pr[\epsilon \leq 0, \boldsymbol{U} \in \Delta_c \cup \Delta_d \cup \Delta_e], \qquad (2.12)
\end{aligned}
$$

where Δ_i ($i \in \{a, ..., e\}$) are disjoint and each Δ_i characterizes those markets on the margin described above: Δ_a corresponds to the set of $\boldsymbol{U}$'s that experience (a), Δ_b corresponds to (b), and so on. Once (2.12) is derived, it is easy to see that

$$
sgn\{h(\boldsymbol{z}, \boldsymbol{z}', x)\} = sgn(\mu_1) = sgn(\mu_2), \qquad (2.13)
$$

which is formally shown in Lemma 3.1(i). See Section D.4 in the Appendix for a proof in this specific two-player case, which simplifies the general proof. The result (2.13) is helpful for our bound analysis. Again, focus on $E[Y(1,1)|x]$ and suppose $h(\boldsymbol{z}, \boldsymbol{z}', x) > 0$. Then, $\mu_1 > 0$ and $\mu_2 > 0$,

[5] Of course, changing each Z_s from $-\infty$ to ∞ will trivially achieve our requirement of having no market that switches from one monopoly to another.

and thus, we can derive the lower bound on, e.g., $E[Y(1,1)|\boldsymbol{D}=(1,0),\boldsymbol{z},x]$ in (2.2) as

$$\begin{aligned} E[Y(1,1)|\boldsymbol{D}=(1,0),\boldsymbol{z},x] &= \Pr[\epsilon \leq \mu_1+\mu_2+\beta x|\boldsymbol{D}=(1,0),\boldsymbol{z},x] \\ &\geq \Pr[\epsilon \leq \mu_1+\beta x|\boldsymbol{D}=(1,0),\boldsymbol{z},x] \\ &= E[Y|\boldsymbol{D}=(1,0),\boldsymbol{z},x], \end{aligned} \tag{2.14}$$

which is larger than zero, the previous naive lower bound. Similarly, we can calculate the lower bounds on all $E[Y(1,1)|\boldsymbol{D}=\boldsymbol{d},\boldsymbol{z},x]$ for $\boldsymbol{d}\neq(1,1)$. Consequently, by (2.2), we have $E[Y(1,1)|x]\geq \Pr[Y=1|\boldsymbol{z},x]$, i.e., the lower bound on $E[Y(1,1)|x]$ is

$$\tilde{L}(x)\equiv\sup_{\boldsymbol{z}}\Pr[Y=1|\boldsymbol{z},x].$$

Note that $\tilde{L}(x)\geq L_{Manski}(x)$. In this case, $\tilde{U}(x)=U_{Manski}(x)$. In Section 3.4, we show that $\tilde{L}(x)$ and $\tilde{U}(x)$ are sharp under (2.3)–(2.8).

We can further tighten the bounds if we have exogenous variables that are excluded from the entry decisions (i.e., from the D_1 and D_2 equations). The existence of such variables is not necessary but helpful in tightening the bounds, and can be motivated by the notion of externalities. That is, there can exist factors that affect Y but do not enter the players' first-stage payoff functions. Modify (2.6) and assume

$$(\epsilon,U_1,U_2)\perp(Z_1,Z_2,X), \tag{2.15}$$

where conditioning on other (possibly endogenous) covariates is suppressed. Here, X_i can be the characteristics of the local market that directly affect pollution or health levels, such as weather shocks or the share of pollution-related industries in the local economy. We assume that conditional on other covariates, these factors affect the outcome but the airlines do not take them into account in their decisions.

To exploit the variation in X (in addition to the variation in $\boldsymbol{Z}$), let $(x,\tilde{x},\tilde{\tilde{x}})$ be (possibly different) realizations of X, and define

$$\tilde{h}(\boldsymbol{z},\boldsymbol{z}',x,\tilde{x},\tilde{\tilde{x}})\equiv h_{00}(\boldsymbol{z},\boldsymbol{z}',x)+h_{10}(\boldsymbol{z},\boldsymbol{z}',\tilde{x})+h_{01}(\boldsymbol{z},\boldsymbol{z}',\tilde{x})+h_{11}(\boldsymbol{z},\boldsymbol{z}',\tilde{\tilde{x}}). \tag{2.16}$$

Under (2.15) and analogous to (2.13), we can show that if

$$sgn\{\tilde{h}(\boldsymbol{z}, \boldsymbol{z}', x', x', x)\} = sgn(-\mu_1) = sgn(-\mu_2) \tag{2.17}$$

is positive (negative), then $sgn\{\mu_1 + \beta(x - x')\} = sgn\{\mu_2 + \beta(x - x')\}$ is positive (negative). This is formally shown in Lemma 3.1(ii).

As before, suppose $h(\boldsymbol{z}, \boldsymbol{z}', x) > 0$, and thus, $\mu_1 > 0$ and $\mu_2 > 0$ by (2.13). Now, if $\tilde{h}(\boldsymbol{z}, \boldsymbol{z}', x', x', x) < 0$, then $\mu_1 + \beta x < \beta x'$ and $\mu_2 + \beta x < \beta x'$. Therefore, we can derive

$$\begin{aligned}
E[Y(1,1)|\boldsymbol{D} = (1,0), \boldsymbol{z}, x] &= \Pr[\epsilon \leq \mu_1 + \mu_2 + \beta x | \boldsymbol{D} = (1,0), \boldsymbol{z}, x] \\
&\leq \Pr[\epsilon \leq \mu_1 + \beta x' | \boldsymbol{D} = (1,0), \boldsymbol{z}, x'] \\
&= \Pr[Y = 1 | \boldsymbol{D} = (1,0), \boldsymbol{z}, x'],
\end{aligned}$$

where the second equality also uses (2.15) and (2.4)–(2.5). Similarly, we have $E[Y(1,1)|\boldsymbol{D} = (0,1), \boldsymbol{z}, x] \leq \Pr[Y = 1|\boldsymbol{D} = (0,1), \boldsymbol{z}, x']$, and consequently, the upper bound on $E[Y(1,1)|x]$ becomes

$$U(x) \equiv \inf_{\boldsymbol{z} \in \mathcal{Z}} \left\{ \Pr[Y = 1, \boldsymbol{D} = (1,1)|\boldsymbol{z}, x] + \Pr[Y = 1, \boldsymbol{D} \in \{(1,0), (0,1)\}|\boldsymbol{z}, x'] + \Pr[\boldsymbol{D} = (0,0)|\boldsymbol{z}, x] \right\}$$

by (2.2), and the lower bound is $L(x) = \tilde{L}(x)$. Note that we can further take infimum over x' such that $\tilde{h}(\boldsymbol{z}, \boldsymbol{z}', x', x', x) < 0$.

To summarize our illustration, our lower and upper bounds, $L(x)$ and $U(x)$, on $E[Y(1,1)|x]$ achieve

$$\begin{aligned}
L(x) &= \tilde{L}(x) \geq L_{Manski}(x), \\
U(x) &\geq \tilde{U}(x) = U_{Manski}(x),
\end{aligned}$$

where the inequalities are strict if $\sum_{\boldsymbol{d} \neq (1,1)} \Pr[Y = 1, \boldsymbol{D} = \boldsymbol{d}|\boldsymbol{z}, x] > 0$ and $\Pr[Y = 0, \boldsymbol{D} \in \{(1,0), (0,1)\}|\boldsymbol{z}, x'] > 0$. Similarly, we can derive lower and upper bounds on other $E[Y(\boldsymbol{d})|x]$'s for $\boldsymbol{d} \neq (1,1)$, and eventually construct bounds on any ATE. The gain from our approach is also exhibited in Figure 5 in Section B, where we use the same data generating process as in this section and calculate different bounds on the ATE, $E[Y(1,1)|x] - E[Y(0,0)|x]$.

3 General Theory

3.1 Setup

Let $\boldsymbol{D} \equiv (D_1, ..., D_S) \in \mathcal{D} \subseteq \{0,1\}^S$ be an S-vector of observed binary treatments and $\boldsymbol{d} \equiv (d_1, ..., d_S)$ be its realization, where S is fixed. We assume that $\boldsymbol{D}$ is predicted as a pure strategy Nash equilibrium of a complete information game with S players who make entry decisions or individuals who choose to receive treatments. Let Y be a scalar observed post-game outcome that results from profile $\boldsymbol{D}$ of endogenous treatments. It can be an outcome common to all players or an outcome specific to each player. Let $(X, Z_1, ..., Z_S)$ be observed exogenous covariates. We consider a model of a semi-triangular system:

$$Y = \theta(\boldsymbol{D}, X, \epsilon_{\boldsymbol{D}}), \tag{3.1}$$

$$D_s = 1\left[\nu^s(\boldsymbol{D}_{-s}, Z_s) \geq U_s\right], \qquad s \in \{1, ..., S\}, \tag{3.2}$$

where s is indices for players or interchangeably for treatments, and $\boldsymbol{D}_{-s} \equiv (D_1, ..., D_{s-1}, D_{s+1}, ..., D_S)$. Without loss of generality, we normalize the scalar U_s to be distributed as $Unif(0,1)$, and $\nu^s : \mathbb{R}^{S-1+d_{z_s}} \rightarrow (0,1]$ and $\theta : \mathbb{R}^{S+d_x+d_\epsilon} \rightarrow \mathbb{R}$ are unknown functions that are nonseparable in their arguments. We allow the unobservables $(\epsilon_{\boldsymbol{D}}, U_1, ..., U_S)$ to be arbitrarily dependent on one another. Although the notation suggests that the instruments Z_s's are player/treatment-specific, they are not necessarily required to be so for the analyses in this study; see Section C.4 for a discussion. The exogenous variables X are variables excluded from all the equations for D_s. The existence of X is not necessary but useful for the bound analysis of the ATE. There may be covariates W common to all the equations for Y and D_s, which is suppressed for succinctness. Implied from the complete information game, player s's decision D_s depends on the decisions of all others $\boldsymbol{D}_{-s}$ in $\mathcal{D}_{-s}$, and thus, $\boldsymbol{D}$ is determined by a simultaneous system. The unit of observation, a market or geographical region, is indexed by i and is suppressed in all the expressions.

Assumption P. *A pure-strategy Nash equilibrium exists in* (3.2).

This assumption is not innocuous but clarifies the definition of the treatment parameters. It may be possible to extend the setup of this study to incorporate mixed strategies, following the argument in Ciliberto and Tamer (2009). This extension, however, will change the definition of the treatment parameters and thus the identification strategy. We leave this as interesting future work.

The potential outcome of receiving treatments $\boldsymbol{D} = \boldsymbol{d}$ can be written as

$$Y(\boldsymbol{d}) = \theta(\boldsymbol{d}, X, \epsilon_{\boldsymbol{d}}), \qquad \boldsymbol{d} \in \mathcal{D},$$

and $\epsilon_{\boldsymbol{D}} = \sum_{\boldsymbol{d} \in \mathcal{D}} 1[\boldsymbol{D} = \boldsymbol{d}]\epsilon_{\boldsymbol{d}}$. We are interested in the ATE and related parameters. Using the average structural function (ASF) $E[Y(\boldsymbol{d})|x]$, the ATE can be written as

$$E[Y(\boldsymbol{d}) - Y(\boldsymbol{d}')|x] = E[\theta(\boldsymbol{d}, x, \epsilon_{\boldsymbol{d}}) - \theta(\boldsymbol{d}', x, \epsilon_{\boldsymbol{d}'})], \tag{3.3}$$

for $\boldsymbol{d}, \boldsymbol{d}' \in \mathcal{D}$. Another parameter of interest is $E[Y(\boldsymbol{d}) - Y(\boldsymbol{d}')|D = \boldsymbol{d}'', z, x]$ for $\boldsymbol{d}, \boldsymbol{d}', \boldsymbol{d}'' \in \mathcal{D}$. This parameter is related to the average treatment effect on the treated (ATT), but unlike the ATT or the treatment of the untreated in the single-treatment case, $\boldsymbol{d}''$ does not necessarily equal $\boldsymbol{d}$ or $\boldsymbol{d}'$ here. One might also be interested in the sign of the ATE, which in this multi-treatment case is essentially establishing an ordering among the ASF's.

As an example of the ATE, we may choose $\boldsymbol{d} = (1, ..., 1)$ and $\boldsymbol{d}' = (0, ..., 0)$ to measure the cancelling-out effect or more general nonlinear effects. Another example would be choosing $\boldsymbol{d} = (1, \boldsymbol{d}_{-s})$ and $\boldsymbol{d}' = (0, \boldsymbol{d}_{-s})$ for given $\boldsymbol{d}_{-s}$, where we use the notation $\boldsymbol{d} = (d_s, \boldsymbol{d}_{-s})$ by switching the order of the elements for convenience. Sometimes, we instead want to focus on learning about complementarity between two treatments, while averaging over the remaining $S-2$ treatments; see Section C.3.

In identifying these treatment parameters, suppose we attempt to recover the effect of a single treatment D_s in model (3.1)–(3.2) *conditional on* $\boldsymbol{D}_{-s} = \boldsymbol{d}_{-s}$, and then recover the effects of multiple treatments by transitively using these effects of single treatments. This strategy is not valid since $\boldsymbol{D}_{-s}$ is a function of D_s and also because of multiplicity. Therefore, the approaches in the literature with single-treatment, single-agent triangular models are not directly applicable.

3.2 Monotonicity in Equilibria

As an important step of the analysis, we establish that the equilibria of the treatment selection process in the first-stage game present a monotonic pattern when the instruments move. Specifically, we consider the regions in the space of the unobservables that predict equilibria and establish their monotonic pattern in terms of instruments. The analytical characterization of the equilibrium regions when there are more than two players ($S > 2$) can generally be complicated (Ciliberto and Tamer (2009, p. 1800)); however, under a mild uniformity assumption (Assumption M1), our result is obtained under strategic substitutability. Let $\mathcal{Z}_s$ be the support of Z_s. We make the following assumptions on the first-stage nonparametric payoff function for each $s \in \{1, ..., S\}$.

Assumption SS. *For every $z_s \in \mathcal{Z}_s$, $\nu^s(\boldsymbol{d}_{-s}, z_s)$ is strictly decreasing in each element of $\boldsymbol{d}_{-s}$.*

Assumption M1. *For any given $z_s, z_s' \in \mathcal{Z}_s$, either $\nu^s(\boldsymbol{d}_{-s}, z_s) \geq \nu^s(\boldsymbol{d}_{-s}, z_s')$ $\forall \boldsymbol{d}_{-s} \in \mathcal{D}_{-s}$, or $\nu^s(\boldsymbol{d}_{-s}, z_s) \leq \nu^s(\boldsymbol{d}_{-s}, z_s')$ $\forall \boldsymbol{d}_{-s} \in \mathcal{D}_{-s}$.*

Assumption SS asserts that the agents' treatment decisions are produced in a game with strategic substitutability. The strictness of the monotonicity is not important for our purpose but convenient in making statements about the equilibrium regions. In the language of Ciliberto and Tamer (2009), we allow for heterogeneity in the *fixed competitive effects* (i.e., how each of other entrants affects one's payoff), as well as heterogeneity in how each player is affected by other entrants, which is ensured by the nonseparability between $\boldsymbol{d}_{-s}$ and z_s in $\nu^s(\boldsymbol{d}_{-s}, z_s)$; this heterogeneity is related to the *variable competitive effects*. Assumption M1 is required in this multi-agent setting, and the uniformity is across $\boldsymbol{d}_{-s}$. Note that this assumption is weaker than a conventional monotonicity assumption that $\nu^s(\boldsymbol{d}_{-s}, z_s)$ is either non-decreasing or non-increasing in z_s for all $\boldsymbol{d}_{-s}$. Assumption M1 is justifiable especially when z_s is chosen to be of the same kind for all players. For example, in an entry game, if z_s is chosen to be each player's cost shifters, the payoffs would decrease in their costs for any given opponents.

As the first main result of this study, we establish the geometric property of the equilibrium regions. For $j = 0, ..., S$, let $\boldsymbol{R}_j(\boldsymbol{z}) \subset \mathcal{U} \equiv (0, 1]^S$ denote the region that predicts all equilibria with j treatments selected or j entrants, defined as a subset of the space of the entry unobservables $\boldsymbol{U} \equiv (U_1, ..., U_S)$; see Section D.1 for a formal definition. Then, define the region of all equilibria with *at most* j entrants as

$$\boldsymbol{R}^{\leq j}(\boldsymbol{z}) \equiv \bigcup_{k=0}^{j} \boldsymbol{R}_k(\boldsymbol{z}).$$

Although this region is hard to express explicitly in general, it has a simple feature that serves our purpose. For given j, choose $z_s, z_s' \in \mathcal{Z}_s$ such that

$$\Pr[\boldsymbol{D} = (1, ..., 1)|\boldsymbol{Z} = (z_s, \boldsymbol{z}_{-s})] > \Pr[\boldsymbol{D} = (1, ..., 1)|\boldsymbol{Z} = (z_s', \boldsymbol{z}_{-s})] \tag{3.4}$$

for all s. This condition is to merely fix $\boldsymbol{z}, \boldsymbol{z}'$ that change the *joint propensity score*, and the direction of change is without loss of generality. Such $\boldsymbol{z}, \boldsymbol{z}'$ exist by the relevance of the instruments, which is assumed below. Let $\mathcal{Z}$ be the support of $\boldsymbol{Z} \equiv (Z_1, ..., Z_S)$.

Theorem 3.1. *Under Assumptions P, SS, and M1 and for $\boldsymbol{z}, \boldsymbol{z}' \in \mathcal{Z}$ that satisfy (3.4), we have*

$$\boldsymbol{R}^{\leq j}(\boldsymbol{z}) \subseteq \boldsymbol{R}^{\leq j}(\boldsymbol{z}') \ \forall j. \tag{3.5}$$

Theorem 3.1 establishes a generalized version of monotonicity in the treatment selection process. This theorem plays a crucial role in calculating the bounds on the treatment parameters and in showing the sharpness of the bounds.[6] It is worth noting that although the monotonicity is restored for sets defined in terms of the number of entrants, the ATE we recover is *aware of the identity of entrants* as shown below.

Remark 3.1. *Strategic substitutability (Assumption SS) is important in obtaining the generalized monotonicity (Theorem 3.1). With strategic complementarity, multiple equilibria occur among equilibria with different numbers of entrants, and therefore the machinery used in our identification analysis no longer applies.*

[6] Relatedly, Berry (1992) derives the probability of the event that the number of entrants is less than a certain value, which can be written as $\Pr[\boldsymbol{U} \in \boldsymbol{R}^{\leq j}(\boldsymbol{z})]$ using our notation. However, his result is not sufficient for our study and relies on stronger assumptions, such as restricting the payoff functions to only depend on the number of opponents.

3.3 Main Assumptions

To characterize the bounds on the treatment parameters, we make the following assumptions. Unless otherwise noted, the assumptions hold for each $s \in \{1, ..., S\}$.

Assumption IN. $(X, \boldsymbol{Z}) \perp (\epsilon_{\boldsymbol{d}}, \boldsymbol{U})$ $\forall \boldsymbol{d} \in \mathcal{D}$.

Assumption E. *The distribution of* $(\epsilon_{\boldsymbol{d}}, \boldsymbol{U})$ *has strictly positive density with respect to Lebesgue measure on* $\mathbb{R}^{S+1}$ $\forall \boldsymbol{d} \in \mathcal{D}$.

Assumption EX. *For each* $\boldsymbol{d}_{-s} \in \mathcal{D}_{-s}$, $\nu^s(\boldsymbol{d}_{-s}, Z_s)|X$ *is nondegenerate.*

Assumptions IN, EX and all the following analyses can be understood as conditional on W, the common covariates in X and $\boldsymbol{Z} = (Z_1, ..., Z_S)$. Assumption EX is related to the exclusion restriction and the relevance condition of the instruments Z_s.

We now impose a shape restriction on the outcome function $\theta(\boldsymbol{d}, x, \epsilon_{\boldsymbol{d}})$ via restrictions on

$$\vartheta(\boldsymbol{d}, x; \boldsymbol{u}) \equiv E[\theta(\boldsymbol{d}, x, \epsilon_{\boldsymbol{d}})|\boldsymbol{U} = \boldsymbol{u}]$$

a.e. $\boldsymbol{u}$. This restriction on the conditional mean is weaker than the one directly imposed on $\theta(\boldsymbol{d}, x, \epsilon_{\boldsymbol{d}})$. Let $\mathcal{X}$ be the support of X. Recall that we use the notation $\boldsymbol{d} = (d_s, \boldsymbol{d}_{-s})$ by switching the order of the elements for convenience.

Assumption M. *For every* $x \in \mathcal{X}$, *either* $\vartheta(1, \boldsymbol{d}_{-s}, x; \boldsymbol{u}) \geq \vartheta(0, \boldsymbol{d}_{-s}, x; \boldsymbol{u})$ *a.e.* $\boldsymbol{u}$ $\forall \boldsymbol{d}_{-s} \in \mathcal{D}_{-s}$ $\forall s$ *or* $\vartheta(1, \boldsymbol{d}_{-s}, x; \boldsymbol{u}) \leq \vartheta(0, \boldsymbol{d}_{-s}, x; \boldsymbol{u})$ *a.e.* $\boldsymbol{u}$ $\forall \boldsymbol{d}_{-s} \in \mathcal{D}_{-s}$ $\forall s$. *Also,* $Y \in [\underline{Y}, \overline{Y}]$.

Assumption M holds in, but is not restricted to, the leading case of binary Y with a threshold crossing model that satisfies uniformity.

Assumption M*. *(i)* $\theta(\boldsymbol{d}, x, \epsilon_{\boldsymbol{d}}) = 1[\mu(\boldsymbol{d}, x) \geq \epsilon_{\boldsymbol{d}}]$ *where* $\epsilon_{\boldsymbol{d}}$ *is scalar and* $F_{\epsilon_{\boldsymbol{d}}|\boldsymbol{U}} = F_{\epsilon_{\boldsymbol{d}'}|\boldsymbol{U}}$ *for any* $\boldsymbol{d}, \boldsymbol{d}' \in \mathcal{D}$; *(ii) for every* $x \in \mathcal{X}$, *either* $\mu(1, \boldsymbol{d}_{-s}, x) \geq \mu(0, \boldsymbol{d}_{-s}, x)$ $\forall \boldsymbol{d}_{-s} \in \mathcal{D}_{-s}$ $\forall s$ *or* $\mu(1, \boldsymbol{d}_{-s}, x) \leq \mu(0, \boldsymbol{d}_{-s}, x)$ $\forall \boldsymbol{d}_{-s} \in \mathcal{D}_{-s}$ $\forall s$.

Assumption M* implies Assumption M. The second statement in Assumption M is satisfied with binary Y.[7] The first statement in Assumption M can be stated in two parts, corresponding to (i) and (ii) of Assumption M*: (a) for every x and $\boldsymbol{d}_{-s}$, either $\vartheta(1, \boldsymbol{d}_{-s}, x; \boldsymbol{u}) \geq \vartheta(0, \boldsymbol{d}_{-s}, x; \boldsymbol{u})$ a.e. $\boldsymbol{u}$,

[7] Another example would be when $Y \in [0, 1]$, as in Example 1.

or $\vartheta(1, \boldsymbol{d}_{-s}, x; \boldsymbol{u}) \leq \vartheta(0, \boldsymbol{d}_{-s}, x; \boldsymbol{u})$ a.e. $\boldsymbol{u}$; (b) for every x, each inequality statement in (a) holds for all $\boldsymbol{d}_{-s}$. For an outcome function with a scalar index, $\theta(\boldsymbol{d}, x, \epsilon_{\boldsymbol{d}}) = \tilde{\theta}(\mu(\boldsymbol{d}, x), \epsilon_{\boldsymbol{d}})$, part (a) is implied by $\epsilon_{\boldsymbol{d}} = \epsilon_{\boldsymbol{d}'} = \epsilon$ (or more generally, $F_{\epsilon_{\boldsymbol{d}}|\boldsymbol{U}} = F_{\epsilon_{\boldsymbol{d}'}|\boldsymbol{U}}$) for any $\boldsymbol{d}, \boldsymbol{d}' \in \mathcal{D}$ and $E[\tilde{\theta}(t, \epsilon_{\boldsymbol{d}})|\boldsymbol{U} = \boldsymbol{u}]$ being strictly increasing (decreasing) in t a.e. $\boldsymbol{u}$.[8] Functions that satisfy the latter assumption include strictly monotonic functions, such as transformation models $\tilde{\theta}(t, \epsilon) = r(t + \epsilon)$ with $r(\cdot)$ being possibly unknown strictly increasing, or their special case $\tilde{\theta}(t, \epsilon) = t + \epsilon$, allowing continuous dependent variables; and functions that are not strictly monotonic, such as models for limited dependent variables, $\tilde{\theta}(t, \epsilon) = 1[t \geq \epsilon]$ or $\tilde{\theta}(t, \epsilon) = 1[t \geq \epsilon](t - \epsilon)$. However, there can be functions that violate the latter assumption but satisfy part (a). For example, consider a threshold crossing model with a random coefficient: $\theta(\boldsymbol{d}, x, \epsilon) = 1[\phi(\epsilon)\boldsymbol{d}\beta^\top \geq x\gamma^\top]$, where $\phi(\epsilon)$ is nondegenerate. When $\beta_s \geq 0$, then $E[\theta(1, \boldsymbol{d}_{-s}, x, \epsilon) - \theta(0, \boldsymbol{d}_{-s}, x, \epsilon)|\boldsymbol{U} = \boldsymbol{u}] = \Pr\left[\frac{x\gamma^\top}{\beta_s + \boldsymbol{d}_{-s}\beta_{-s}^\top} \leq \phi(\epsilon) \leq \frac{x\gamma^\top}{\boldsymbol{d}_{-s}\beta_{-s}^\top}|\boldsymbol{U} = \boldsymbol{u}\right]$, and thus, nonnegative a.e. $\boldsymbol{u}$, and vice versa. Part (a) also does not impose any monotonicity of θ in $\epsilon_{\boldsymbol{d}}$ (e.g., $\epsilon_{\boldsymbol{d}}$ can be a vector). It is worth noting that part (a) is not consistent with the possibility that agents make treatment decisions by directly observing the potential outcomes (i.e., the structure of Roy models). For instance, $F_{\epsilon_{\boldsymbol{d}}|\boldsymbol{U}} = F_{\epsilon_{\boldsymbol{d}'}|\boldsymbol{U}}$ (for any $\boldsymbol{d}, \boldsymbol{d}' \in \mathcal{D}$) does not hold in Roy models, because U_s needs to be a function of $\epsilon_{\boldsymbol{d}}$'s (possibly for all $\boldsymbol{d} \in \mathcal{D}$) under the structure. However, ruling out this structure is consistent with the notion of externality we propose in this paper.

Part (b) of Assumption M imposes uniformity, as we deal with more than one treatment. Uniformity is required across different values of $\boldsymbol{d}_{-s}$ and s. For instance, in the empirical application of this study, this assumption seems reasonable, since an airline's entry is likely to increase the expected pollution regardless of the identity or the number of existing airlines. On the other hand, in Example 1 in the Appendix regarding media and political behavior, this assumption may rule out the "over-exposure" effect (i.e., too much media exposure diminishes the incumbent's chance of being re-elected). In any case, knowledge on the direction of the monotonicity is not necessary in this assumption, unlike Manski (1997) or Manski (2013), where the semi-monotone treatment response is assumed for possible multiple treatments.

Lastly, we require that there exists variation in $\boldsymbol{Z}$ that offsets the effect of strategic substitutability. Similar as before, using the notation $\boldsymbol{d}_{-s} = (d_{s'}, \boldsymbol{d}_{-(s,s')})$ where $\boldsymbol{d}_{-(s,s')}$ is $\boldsymbol{d}$ without s-th and

[8] A single-treatment version of the latter assumption appears in Vytlacil and Yildiz (2007) (Assumption A-4), which is weaker than assuming $\tilde{\theta}(t, \epsilon)$ is strictly increasing (decreasing) a.e. ϵ; see Vytlacil and Yildiz (2007) for related discussions.

s'-th elements, note that Assumption SS can be restated as $\nu^s(0, \boldsymbol{d}_{-(s,s')}, z_s) > \nu^s(1, \boldsymbol{d}_{-(s,s')}, z_s)$ for every z_s. Given this, we assume the following compensating strategic substitutability.

Assumption EQ. *There exist $\boldsymbol{z}, \boldsymbol{z}' \in \mathcal{Z}$, such that $\nu^s(0, \boldsymbol{d}_{-(s,s')}, z'_s) \leq \nu^s(1, \boldsymbol{d}_{-(s,s')}, z_s)$ $\forall \boldsymbol{d}_{-(s,s')}$ $\forall s, s'$.*

For example, in an entry game with Z_s being cost shifters, Assumption EQ may hold with $z'_s > z_s$ $\forall s$. In this example, players may become less profitable with an increase in cost from government regulation. In particular, players' decreased profits cannot be overturned by the market being less competitive, as one player is absent due to unprofitability. Recall that Assumption EQ is illustrated in Figure 1 with $\nu^s(0, z'_s) = \gamma_s z'_s < \nu^s(1, z_s) = \delta_{-s} + \gamma_s z_s$ for $s = 1, 2$. Assumption EQ is key for our analysis. To see this, let $R^M_j(\cdot)$ denote the region that predicts multiple equilibria with j treatments selected or j entrants. In the proof of a lemma that follows, we show that Assumption EQ holds if and only if $R^M_j(\boldsymbol{z}) \cap R^M_j(\boldsymbol{z}') = \emptyset$. That is, we can at least ensure that there is no market where firms' decisions change from one realization of multiple equilibria to another realization of multiple equilibria with the same number of entrants. To the extent of our analysis, this liberates us from concerns about a possible change in equilibrium selection when changing $\boldsymbol{Z}$.[9] Assumption EQ has a simple testable sufficient condition, provided that the unobservables in the payoffs are mutually independent. Let $\boldsymbol{d}^j \in \mathcal{D}^j$ denote an equilibrium profile with j treatments selected or j entrants, i.e., a vector of j ones and $S - j$ zeros, where $\mathcal{D}^j$ is a set of all equilibrium profiles with j treatments selected.

Assumption EQ*. *There exist $\boldsymbol{z}, \boldsymbol{z}' \in \mathcal{Z}$, such that*

$$\Pr[\boldsymbol{D} = \boldsymbol{d}^j | \boldsymbol{z}] + \Pr[\boldsymbol{D} = \boldsymbol{d}^{j-2} | \boldsymbol{z}'] > 2 - \sqrt{2}. \tag{3.6}$$

for all $\boldsymbol{d}^j \in \mathcal{D}^j$, $\boldsymbol{d}^{j-2} \in \mathcal{D}^{j-2}$ and $2 \leq j \leq S$.

When $S = 2$, the condition is stated as $\Pr[\boldsymbol{D} = (1,1)|\boldsymbol{z}] + \Pr[\boldsymbol{D} = (0,0)|\boldsymbol{z}'] > 2 - \sqrt{2}$. As is detailed in the proof, this essentially restricts the sum of radii of two circular isoquant curves to be less than the length of the diagonal of $\mathcal{U}$: $(1 - \Pr[\boldsymbol{D} = (1,1)|\boldsymbol{z}]) + (1 - \Pr[\boldsymbol{D} = (0,0)|\boldsymbol{z}']) < \sqrt{2}$. This ensures the required variation in Assumption EQ.

[9] In Section C.5, we discuss an assumption, partial conditional symmetry, which can be imposed alternative to Assumption EQ.

Lemma 3.1. *Under Assumptions P, SS, M1, and $U_s \perp U_t$ for all $s \neq t$, Assumption EQ* implies Assumption EQ.*

The mutual independence of U_s's (conditional on W) is useful in inferring the relationship between players' interaction and instruments from the observed choices of players. The intuition for the sufficiency of Assumption EQ* is as follows. As long as there is no dependence in unobserved types, (3.6) dictates that the variation of $\boldsymbol{Z}$ is large enough to offset strategic substitutability, because otherwise, the payoffs of players cannot move in the same direction, and thus, will not result in the same decisions. The requirement of $\boldsymbol{Z}$ variation in (3.6) is significantly weaker than the large support assumption invoked for an identification at infinity argument to overcome the problem of multiple equilibria.

3.4 Partial Identification of the ATE

Under the above assumptions, we now present a generalized version of the sign matching results (2.13) and (2.17) in Section 2. We need to introduce additional notation. For realizations x of X and $\boldsymbol{z}, \boldsymbol{z}'$ of $\boldsymbol{Z}$, define

$$h(\boldsymbol{z}, \boldsymbol{z}', x) \equiv E[Y|\boldsymbol{z}, x] - E[Y|\boldsymbol{z}', x], \tag{3.7}$$

$$\begin{aligned} h_{\boldsymbol{d}^j}(\boldsymbol{z}, \boldsymbol{z}', x) \equiv\ & E[Y|\boldsymbol{D} = \boldsymbol{d}^j, \boldsymbol{z}, x] \Pr[\boldsymbol{D} = \boldsymbol{d}^j|\boldsymbol{z}] \\ & - E[Y|\boldsymbol{D} = \boldsymbol{d}^j, \boldsymbol{z}', x] \Pr[\boldsymbol{D} = \boldsymbol{d}^j|\boldsymbol{z}']. \end{aligned} \tag{3.8}$$

Since $\sum_{j=0}^{S} \sum_{\boldsymbol{d}^j \in \mathcal{D}^j} \Pr[\boldsymbol{D} = \boldsymbol{d}^j|\cdot] = 1$, $h(\boldsymbol{z}, \boldsymbol{z}', x) = \sum_{j=0}^{S} \sum_{\boldsymbol{d}^j} h_{\boldsymbol{d}^j}(\boldsymbol{z}, \boldsymbol{z}', x)$. Let $\tilde{\boldsymbol{x}} = (x_0, ..., x_S)$ be an $(S+1)$-dimensional array of (possibly different) realizations of X, i.e., each x_j for $j = 0, ..., S$ is a realization of X, and define

$$\tilde{h}(\boldsymbol{z}, \boldsymbol{z}', \tilde{\boldsymbol{x}}) \equiv \sum_{j=0}^{S} \sum_{\boldsymbol{d}^j \in \mathcal{D}^j} h_{\boldsymbol{d}^j}(\boldsymbol{z}, \boldsymbol{z}', x_j).$$

For $1 \leq k \leq j$, define a *reduction* of $\boldsymbol{d}^j = (d_1^j, ..., d_S^j)$ as $\boldsymbol{d}^{j-k} = (d_1^{j-k}, ..., d_S^{j-k})$, such that $d_s^{j-k} \leq d_s^j$ $\forall s$. Symmetrically, for $1 \leq k \leq S - j$, define an *extension* of $\boldsymbol{d}^j$ as $\boldsymbol{d}^{j+k} = (d_1^{j+k}, ..., d_S^{j+k})$, such that $d_s^{j+k} \geq d_s^j$ $\forall s$. For example, given $\boldsymbol{d}^2 = (1, 1, 0)$, a reduction $\boldsymbol{d}^1$ is either $(1, 0, 0)$ or $(0, 1, 0)$ but not $(0, 0, 1)$, a reduction $\boldsymbol{d}^0$ is $(0, 0, 0)$, and an extension $\boldsymbol{d}^3$ is $(1, 1, 1)$. As seen from this example, a particular reduction and extension depends on the identity of entrants. Let $\mathcal{D}^{<}(\boldsymbol{d}^j)$ and $\mathcal{D}^{>}(\boldsymbol{d}^j)$ be the set of all reductions and extensions of $\boldsymbol{d}^j$, respectively, and let $\mathcal{D}^{\leq}(\boldsymbol{d}^j) \equiv \mathcal{D}^{<}(\boldsymbol{d}^j) \cup \{\boldsymbol{d}^j\}$ and $\mathcal{D}^{\geq}(\boldsymbol{d}^j) \equiv \mathcal{D}^{>}(\boldsymbol{d}^j) \cup \{\boldsymbol{d}^j\}$. These sets will be used in expressing the bounds on the ATE. Recall $\vartheta(\boldsymbol{d}, x; \boldsymbol{u}) \equiv E[\theta(\boldsymbol{d}, x, \epsilon)|\boldsymbol{U} = \boldsymbol{u}]$. Now, we state the main lemma of this section.

Lemma 3.2. *In model (3.1)–(3.2), suppose Assumptions P, SS, M1, IN, E, EX, and M hold, and $h(\boldsymbol{z}, \boldsymbol{z}', x)$ and $h(\boldsymbol{z}, \boldsymbol{z}', \tilde{\boldsymbol{x}})$ are well-defined. For $\boldsymbol{z}, \boldsymbol{z}'$ such that (3.4) and Assumption EQ hold, and for $j = 1, ..., S$, it satisfies that*

(i) sgn$\{h(\boldsymbol{z}, \boldsymbol{z}', x)\}$ = sgn$\left\{\vartheta(\boldsymbol{d}^j, x; \boldsymbol{u}) - \vartheta(\boldsymbol{d}^{j-1}, x; \boldsymbol{u})\right\}$ a.e. $\boldsymbol{u}$ $\forall \boldsymbol{d}^{j-1} \in \mathcal{D}^{<}(\boldsymbol{d}^j)$;

(ii) for $\iota \in \{-1, 0, 1\}$, if sgn$\{\tilde{h}(\boldsymbol{z}, \boldsymbol{z}', \tilde{\boldsymbol{x}})\}$ = sgn$\{-\vartheta(\boldsymbol{d}^k, x_k; \boldsymbol{u}) + \vartheta(\boldsymbol{d}^{k-1}, x_{k-1}; \boldsymbol{u})\} = \iota$ $\forall \boldsymbol{d}^{k-1} \in \mathcal{D}^{<}(\boldsymbol{d}^k)$ $\forall k \neq j$ $(k \geq 1)$, then sgn$\{\vartheta(\boldsymbol{d}^j, x_j; \boldsymbol{u}) - \vartheta(\boldsymbol{d}^{j-1}, x_{j-1}; \boldsymbol{u})\} = \iota$ a.e. $\boldsymbol{u}$ $\forall \boldsymbol{d}^{j-1} \in \mathcal{D}^{<}(\boldsymbol{d}^j)$.

Parts (i) and (ii) parallel (2.13) and (2.17), respectively. Using Lemma 3.2, we can learn about the ATE. First, note that the sign of the ATE is identified by Lemma 3.2(i), since $E[Y(\boldsymbol{d})|x] = E[\vartheta(\boldsymbol{d}, x; \boldsymbol{U})]$. Next, we establish the bounds on $E[Y(\boldsymbol{d}^j)|x]$ for given $\boldsymbol{d}^j$ for some $j = 0, ..., S$.

We first present the bounds using the variation in $\boldsymbol{Z}$ only, i.e., by using Lemma 3.2(i). To this end, we fix $X = x$ and suppress it in all relevant expressions. To gain efficiency we define the integrated version of h as

$$H(x) \equiv E\left[h(\boldsymbol{Z}, \boldsymbol{Z}', x)\,\middle|(\boldsymbol{Z}, \boldsymbol{Z}') \in \mathcal{Z}_{EQ,j}\ \forall j = 0, ..., S-1\right], \tag{3.9}$$

where $\mathcal{Z}_{EQ,j}$ is the set of $(\boldsymbol{z}, \boldsymbol{z}')$ that satisfy (3.4) and Assumption EQ given j, and $h(\boldsymbol{z}, \boldsymbol{z}', x) = 0$ whenever it is not well-defined. We focus on the case $H(x) > 0$; $H(x) < 0$ is symmetric and $H(x) = 0$ is straightforward. Using Lemma 3.2(i), one can readily show that $L_{\boldsymbol{d}^j}(x) \leq E[Y(\boldsymbol{d}^j)|x] \leq U_{\boldsymbol{d}^j}(x)$ with

$$U_{\boldsymbol{d}^j}(x) \equiv \inf_{\boldsymbol{z} \in \mathcal{Z}} \left\{ \Pr[Y = 1, \boldsymbol{D} \in \mathcal{D}^{\geq}(\boldsymbol{d}^j)|\boldsymbol{z}, x] + \Pr[\boldsymbol{D} \in \mathcal{D} \backslash \mathcal{D}^{\geq}(\boldsymbol{d}^j)|\boldsymbol{z}, x] \right\}, \tag{3.10}$$

$$L_{\boldsymbol{d}^j}(x) \equiv \sup_{\boldsymbol{z} \in \mathcal{Z}} \left\{ \Pr[Y = 1, \boldsymbol{D} \in \mathcal{D}^{\leq}(\boldsymbol{d}^j)|\boldsymbol{z}, x] \right\}. \tag{3.11}$$

We can simplify these bounds and show that they are sharp under the following assumption.

Assumption C. *(i) $\mu_{\boldsymbol{d}}(\cdot)$ and $\nu_{\boldsymbol{d}_{-s}}(\cdot)$ are continuous; (ii) $\mathcal{Z}$ is compact.*

Under Assumption C, for given $\boldsymbol{d}^j$, there exist vectors $\bar{\boldsymbol{z}} \equiv (\bar{z}_1, ..., \bar{z}_S)$ and $\underline{\boldsymbol{z}} \equiv (\underline{z}_1, ..., \underline{z}_S)$ that satisfy

$$\begin{aligned} \bar{\boldsymbol{z}} &= \arg\max_{\boldsymbol{z}\in\mathcal{Z}} \max_{\boldsymbol{d}\in\mathcal{D}^{\geq}(\boldsymbol{d}^j)} \Pr[\boldsymbol{D} = \boldsymbol{d}|\boldsymbol{z}], \\ \underline{\boldsymbol{z}} &= \arg\min_{\boldsymbol{z}\in\mathcal{Z}} \min_{\boldsymbol{d}\in\mathcal{D}^{\geq}(\boldsymbol{d}^j)} \Pr[\boldsymbol{D} = \boldsymbol{d}|\boldsymbol{z}]. \end{aligned} \tag{3.12}$$

The following is the first main result of this study, which establishes the sharp bounds on $E[Y(\boldsymbol{d}^j)|x]$, where $X = x$ is fixed in the model.

Theorem 3.2. *Given model (3.1)–(3.2) with fixed $X = x$, suppose Assumptions P, SS, M1, IN, E, EX, M*, EQ and C hold. In addition, suppose $H(x)$ is well-defined and $H(x) \geq 0$. Then, the bounds $U_{\boldsymbol{d}^j}$ and $L_{\boldsymbol{d}^j}$ in (3.10) and (3.11) simplify to*

$$\begin{aligned} U_{\boldsymbol{d}^j}(x) &= \Pr[Y = 1, \boldsymbol{D} \in \mathcal{D}^{\geq}(\boldsymbol{d}^j)|\bar{\boldsymbol{z}}, x] + \Pr[\boldsymbol{D} \in \mathcal{D}\backslash\mathcal{D}^{\geq}(\boldsymbol{d}^j)|\bar{\boldsymbol{z}}, x], \\ L_{\boldsymbol{d}^j}(x) &= \Pr[Y = 1, \boldsymbol{D} \in \mathcal{D}^{\leq}(\boldsymbol{d}^j)|\underline{\boldsymbol{z}}, x], \end{aligned}$$

and these bounds are sharp.

With binary Y (Assumption M*), sharp bounds on the mean treatment parameters can be obtained, which is reminiscent of the findings of studies that consider single-treatment models (e.g., Shaikh and Vytlacil (2011)).

When the variation of X is additionally exploited in the model, the bounds will be narrower than the bounds in Theorem 3.2. We now proceed with this case, utilizing Lemma 3.2 (i) and (ii). First, analogous to (3.9), we define the integrated version of $\tilde{h}(\boldsymbol{z}, \boldsymbol{z}', \tilde{\boldsymbol{x}})$ as

$$\tilde{H}(\tilde{\boldsymbol{x}}) \equiv E\left[\tilde{h}(\boldsymbol{Z}, \boldsymbol{Z}', \tilde{\boldsymbol{x}}) \,\middle|\, (\boldsymbol{Z}, \boldsymbol{Z}') \in \mathcal{Z}_{EQ,j} \ \forall j = 0, ..., S-1\right],$$

where $\tilde{h}(\boldsymbol{z}, \boldsymbol{z}', \tilde{\boldsymbol{x}}) = 0$ whenever it is not well-defined. Then, we define the following sets of two consecutive elements (x_j, x_{j-1}) of $\boldsymbol{x}$ that satisfy the conditions in Lemma 3.2: for $j = 1, ..., S$, define $\mathcal{X}^0_{j,j-1}(\iota) \equiv \{(x_j, x_{j-1}) : sgn\{\tilde{H}(\tilde{\boldsymbol{x}})\} = \iota, x_0 = \cdots = x_S\}$ and for $t \geq 1$,

$$\mathcal{X}^t_{j,j-1}(\iota) \equiv \{(x_j, x_{j-1}) : sgn\{\tilde{H}(\tilde{\boldsymbol{x}})\} = \iota, (x_k, x_{k-1}) \in \mathcal{X}^{t-1}_{k,k-1}(-\iota) \ \forall k \neq j\} \cup \mathcal{X}^{t-1}_{j,j-1}(\iota),$$

where the sets are understood to be empty whenever $\tilde{h}(\boldsymbol{z}, \boldsymbol{z}', \tilde{\boldsymbol{x}})$ is not well-defined for any $p_{M^{\leq j}}(\boldsymbol{z}) <$

$p_{M^{\leq j}}(\boldsymbol{z}') \; \forall j$. Note that $\mathcal{X}^t_{j,j-1}(\iota) \subset \mathcal{X}^{t+1}_{j,j-1}(\iota)$ for any t. Define $\mathcal{X}_{j,j-1}(\iota) \equiv \lim_{t\to\infty} \mathcal{X}^t_{j,j-1}(\iota)$.[10] Then, by Lemma 3.2, if $(x_j, x_{j-1}) \in \mathcal{X}_{j,j-1}(\iota)$, then

$$sgn\{\vartheta(\boldsymbol{d}^j, x_j; \boldsymbol{u}) - \vartheta(\boldsymbol{d}^{j-1}, x_{j-1}; \boldsymbol{u})\} = \iota \text{ a.e. } \boldsymbol{u} \; \forall \boldsymbol{d}^{j-1} \in \mathcal{D}^{<}(\boldsymbol{d}^j). \tag{3.13}$$

In conclusion, for bounds on the ATE $E[Y(\boldsymbol{d}^j)|x]$, we can introduce the sets $\mathcal{X}^L_{\boldsymbol{d}^j}(x; \boldsymbol{d}')$ and $\mathcal{X}^U_{\boldsymbol{d}^j}(x; \boldsymbol{d}')$ for $\boldsymbol{d}' \neq \boldsymbol{d}^j$ as follows: for $\boldsymbol{d}' \in \mathcal{D}^{<}(\boldsymbol{d}^j) \cup \mathcal{D}^{>}(\boldsymbol{d}^j)$,

$$\begin{aligned}
\mathcal{X}^L_{\boldsymbol{d}^j}(x; \boldsymbol{d}') \equiv & \left\{x_{j'} : (x_k, x_{k-1}) \in \mathcal{X}_{k,k-1}(-1) \cup \mathcal{X}_{k,k-1}(0) \text{ for } j'+1 \leq k \leq j, x_j = x\right\} \\
& \cup \left\{x_{j'} : (x_k, x_{k-1}) \in \mathcal{X}_{k,k-1}(1) \cup \mathcal{X}_{k,k-1}(0) \text{ for } j+1 \leq k \leq j', x_j = x\right\}, \qquad (3.14) \\
\mathcal{X}^U_{\boldsymbol{d}^j}(x; \boldsymbol{d}') \equiv & \left\{x_{j'} : (x_k, x_{k-1}) \in \mathcal{X}_{k,k-1}(1) \cup \mathcal{X}_{k,k-1}(0) \text{ for } j'+1 \leq k \leq j, x_j = x\right\} \\
& \cup \left\{x_{j'} : (x_k, x_{k-1}) \in \mathcal{X}_{k,k-1}(-1) \cup \mathcal{X}_{k,k-1}(0) \text{ for } j+1 \leq k \leq j', x_j = x\right\}. \qquad (3.15)
\end{aligned}$$

The following theorem summarizes our results:

Theorem 3.3. *In model (3.1)–(3.2), suppose the assumptions of Lemma 3.2 hold. Then the sign of the ATE is identified, and the upper and lower bounds on the ASF and ATE with* $\mathbf{d}, \tilde{\boldsymbol{d}} \in \mathcal{D}$ *are*

$$L_{\boldsymbol{d}}(x) \leq E[Y(\mathbf{d})|x] \leq U_{\boldsymbol{d}}(x)$$

and $L_{\boldsymbol{d}}(x) - U_{\tilde{\boldsymbol{d}}}(x) \leq E[Y(\mathbf{d}) - Y(\tilde{\mathbf{d}})|x] \leq U_{\boldsymbol{d}}(x) - L_{\tilde{\boldsymbol{d}}}(x)$, *where for any given* $\boldsymbol{d}^j \in \mathcal{D}^j \subset \mathcal{D}$ *for some* j,

$$\begin{aligned}
U_{\boldsymbol{d}^j}(x) \equiv \inf_{\boldsymbol{z}\in\mathcal{Z}} \Bigg\{ & E[Y|\boldsymbol{D} = \boldsymbol{d}^j, \boldsymbol{z}, x] \Pr[\boldsymbol{D} = \boldsymbol{d}^j|\boldsymbol{z}] + \Pr[\boldsymbol{D} \in \mathcal{D}^j \backslash \{\boldsymbol{d}^j\}|\boldsymbol{z}]\overline{Y} \\
& + \sum_{\boldsymbol{d}' \in \mathcal{D}^{<}(\boldsymbol{d}^j) \cup \mathcal{D}^{>}(\boldsymbol{d}^j)} \inf_{x' \in \mathcal{X}^U_{\boldsymbol{d}^j}(x;\boldsymbol{d}')} E[Y|\boldsymbol{D} = \boldsymbol{d}', \boldsymbol{z}, x'] \Pr[\boldsymbol{D} = \boldsymbol{d}'|\boldsymbol{z}] \Bigg\}, \\
L_{\boldsymbol{d}^j}(x) \equiv \sup_{\boldsymbol{z}\in\mathcal{Z}} \Bigg\{ & E[Y|\boldsymbol{D} = \boldsymbol{d}^j, \boldsymbol{z}, x] \Pr[\boldsymbol{D} = \boldsymbol{d}^j|\boldsymbol{z}] + \Pr[\boldsymbol{D} \in \mathcal{D}^j \backslash \{\boldsymbol{d}^j\}|\boldsymbol{z}]\underline{Y} \\
& + \sum_{\boldsymbol{d}' \in \mathcal{D}^{<}(\boldsymbol{d}^j) \cup \mathcal{D}^{>}(\boldsymbol{d}^j)} \sup_{x' \in \mathcal{X}^L_{\boldsymbol{d}^j}(x;\boldsymbol{d}')} E[Y|\boldsymbol{D} = \boldsymbol{d}', \boldsymbol{z}, x'] \Pr[\boldsymbol{D} = \boldsymbol{d}'|\boldsymbol{z}] \Bigg\}.
\end{aligned}$$

[10] In practice, the formula for $\mathcal{X}^t_{j,j-1}$ provides a natural algorithm to construct the set $\mathcal{X}_{j,j-1}$ for the computation of the bounds. The calculation of each $\mathcal{X}^t_{j,j-1}$ is straightforward, as it is a search over a two-dimensional space for (x_j, x_{j-1}) once the set $\mathcal{X}^{t-1}_{j,j-1}$ from the previous step is obtained. Practitioners can employ truncation $t \leq T$ for some T and use $\mathcal{X}^T_{j,j-1}$ as an approximation for $\mathcal{X}_{j,j-1}$.

See Sections 4 and B for concrete examples of the expression of $U_{\boldsymbol{d}^j}(x)$ and $L_{\boldsymbol{d}^j}(x)$. The terms $\Pr[\boldsymbol{D}=\boldsymbol{d}'|\boldsymbol{z}]\overline{Y}$ and $\Pr[\boldsymbol{D}=\boldsymbol{d}'|\boldsymbol{z}]\underline{Y}$ appear in the expression of the bounds because Lemma 3.2 cannot establish an order between $\vartheta(\boldsymbol{d},x;\boldsymbol{u})$'s for $\boldsymbol{d}\in\mathcal{D}^j$, which is related to the complication due to multiple equilibria, which occurs for $\boldsymbol{d}\in\mathcal{D}^j$. When the variation in $\boldsymbol{Z}$ is only used in deriving the bounds, $\mathcal{X}_{k,k-1}(\iota)$ should simply reduce to $\mathcal{X}^0_{k,k-1}(\iota)$ in the definition of $\mathcal{X}^L_{\boldsymbol{d}^j}(x;\boldsymbol{d}')$ and $\mathcal{X}^U_{\boldsymbol{d}^j}(x;\boldsymbol{d}')$. When Y is binary with no X, such bounds are equivalent to (3.10) and (3.11). The variation in X given $\boldsymbol{Z}$ yields substantially narrower bounds than the sharp bounds established in Theorem 3.2 under Assumption C. However, the resulting bounds are not automatically implied to be sharp from Theorem 3.2, since they are based on a different DGP and the additional exclusion restriction.

Remark 3.2. *Maintaining that Y is binary, sharp bounds on the ATE with variation in X can be derived assuming that the signs of $\vartheta(\mathbf{d},x;\mathbf{u})-\vartheta(\mathbf{d}',x';\mathbf{u})$ are identified for $\mathbf{d}\in\mathcal{D}$ and $\mathbf{d}'\in\mathcal{D}^<(\mathbf{d})$ and $x,x'\in\mathcal{X}$ via Lemma 3.2. To see this, define*

$$\tilde{\mathcal{X}}^U_{\boldsymbol{d}}(x;\boldsymbol{d}')\equiv\left\{x':\vartheta(\mathbf{d},x;\mathbf{u})-\vartheta(\mathbf{d}',x';\mathbf{u})\leq 0 \ a.e.\ \mathbf{u}\right\},$$
$$\tilde{\mathcal{X}}^L_{\boldsymbol{d}}(x;\boldsymbol{d}')\equiv\left\{x':\vartheta(\mathbf{d},x;\mathbf{u})-\vartheta(\mathbf{d}',x';\mathbf{u})\geq 0 \ a.e.\ \mathbf{u}\right\},$$

which are identified by assumption. Then, by replacing $\mathcal{X}^i_{\boldsymbol{d}}(x;\boldsymbol{d}')$ with $\tilde{\mathcal{X}}^i_{\boldsymbol{d}}(x;\boldsymbol{d}')$ (for $i\in\{U,L\}$) in Theorem 3.3, we may be able to show that the resulting bounds are sharp. Since Lemma 3.2 implies that $\mathcal{X}^i_{\boldsymbol{d}^j}(x;\boldsymbol{d}')\subset\tilde{\mathcal{X}}^i_{\boldsymbol{d}^j}(x;\boldsymbol{d}')$ but not necessarily $\mathcal{X}^i_{\boldsymbol{d}^j}(x;\boldsymbol{d}')\supset\tilde{\mathcal{X}}^i_{\boldsymbol{d}^j}(x;\boldsymbol{d}')$, these modified bounds and the original bounds in Theorem 3.3 do not coincide. This contrasts with the result of Shaikh and Vytlacil (2011) for a single-treatment model, and the complication lies in the fact that we deal with an incomplete model with a vector treatment. When there is no X, Lemma 3.2(i) establishes equivalence between the two signs, and thus, $\mathcal{X}^i_{\boldsymbol{d}^j}(x;\boldsymbol{d}')=\tilde{\mathcal{X}}^i_{\boldsymbol{d}^j}(x;\boldsymbol{d}')$ for $i\in\{U,L\}$, which results in Theorem 3.2. Relatedly, we can also exploit variation in W, namely, variables that are common to both X and $\boldsymbol{Z}$ (with or without exploiting excluded variation of X). This is related to the analysis of Chiburis (2010) and Mourifié (2015) in a single-treatment setting. One caveat of this approach is that, similar to these papers, we need to additionally assume that $W\perp(\epsilon,\boldsymbol{U})$.

Remark 3.3. *When X does not have enough variation, we can calculate the bounds on the ATE. To see this, suppose we do not use the variation in X and suppose $H(x)\geq 0$. Then $\vartheta(\boldsymbol{d}^j,x;\boldsymbol{u})\geq\vartheta(\boldsymbol{d}^{j-1},x;\boldsymbol{u})$ $\forall\boldsymbol{d}^{j-1}\in\mathcal{D}^<(\boldsymbol{d}^j)$ $\forall j=1,...,S$ by Lemma 3.2(i) and by transitivity, $\vartheta(\boldsymbol{d}',x;\boldsymbol{u})\geq$*

$\vartheta(\boldsymbol{d}, x; \boldsymbol{u})$ with $\boldsymbol{d}'$ being an extension of $\boldsymbol{d}$. Therefore, we have

$$E[Y(\boldsymbol{d})|x] \leq E[Y|\boldsymbol{D} = \boldsymbol{d}, \boldsymbol{z}, x] \Pr[\boldsymbol{D} = \boldsymbol{d}|\boldsymbol{z}] + \sum_{\boldsymbol{d}' \in \mathcal{D}^{>}(\boldsymbol{d})} E[Y|\boldsymbol{D} = \boldsymbol{d}', \boldsymbol{z}, x] \Pr[\boldsymbol{D} = \boldsymbol{d}'|\boldsymbol{z}]$$
$$+ \sum_{\boldsymbol{d}' \in \mathcal{D} \setminus \mathcal{D}^{\geq}(\boldsymbol{d})} E[Y(\boldsymbol{d}^{j})|\boldsymbol{D} = \boldsymbol{d}', \boldsymbol{z}, x] \Pr[\boldsymbol{D} = \boldsymbol{d}'|\boldsymbol{z}]. \tag{3.16}$$

Without using variation in X, we can bound the last term in (3.16) by $Y \in [\underline{Y}, \overline{Y}]$. This is done above with $\theta(\boldsymbol{d}, x, \epsilon) = 1[\mu(\boldsymbol{d}, x) \geq \epsilon_{\boldsymbol{d}}]$ and $\vartheta(\boldsymbol{d}, x; \boldsymbol{u}) = F_{\epsilon|\boldsymbol{U}}(\mu(\boldsymbol{d}, x)|\boldsymbol{u})$.

4 Empirical Application: Airline Markets and Pollution

In this section, we take the bounds proposed in Section 3.4 to data on airline market structure and air pollution in the top 100 metropolitan statistical areas in the U.S.

In 2013, aircrafts were responsible for about 3 percent of total U.S. carbon dioxide emissions and nearly 9 percent of carbon dioxide emissions from the U.S. transportation sector, and it is one of the fastest growing sources.[11] Airplanes remain the single largest source of carbon dioxide emissions within the U.S. transportation sector, which is not yet subject to greenhouse gas regulations. In addition to aircrafts, airport land operations are also a big source of pollution: 43 of the 50 largest airports are in ozone nonattainment areas and 12 are in particulate matter nonattainment areas.[12]

There is growing literature showing the effects of air pollution on various health outcomes (see, Schlenker and Walker (2015), Chay and Greenstone (2003), Knittel et al. (2011)). In particular, Schlenker and Walker (2015) show that the causal effect of airport pollution on the health of local residents—using a clever instrumental variable approach—is sizable. Their study focuses on the 12 major airports in California and implicitly assume that the level of competition (or market structure) is fixed. Using high-frequency data, they exploit weather shocks in the East coast—that might affect airport activity in California through network effects—to quantify the effect of airport pollution on respiratory and cardiovascular health complications. In contrast, we take the link between airport pollution and health outcomes as given and are interested in quantifying the effects

[11] See `https://www.c2es.org/content/reducing-carbon-dioxide-emissions-from-aircraft/7/`

[12] Ozone is not emitted directly but is formed when nitrogen oxides and hydrocarbons react in the atmosphere in the presence of sunlight. In United States environmental law, a non-attainment area is an area considered to have air quality worse than the National Ambient Air Quality Standards as defined in the Clean Air Act.

of different (endogenous) market structures of the airline industry on air pollution.[13]

In our analysis we combine data from two sources. The first contains airline information from the Department of Transportation. These data have been used extensively in the literature to analyze the airline industry (see, e.g., Borenstein (1989), Berry (1992), Ciliberto and Tamer (2009), and more recently, Li et al. (2018) and Ciliberto et al. (2018)). The second source contains air pollution data in each airport from air monitoring stations compiled by the Environmental Protection Agency. We discuss the definition and construction of the variables in Section E of the Appendix. We experimented with two measures of pollution (fine particulate matter and ozone levels) which we discuss in the Appendix and obtain qualitatively and quantitatively similar results in all cases.[14] In order to save space, we only show results using particulate matter concentration as our outcome variable.

We assume that, in each market, airlines choose to be "in" or "out" in a simultaneous entry game of perfect information, as introduced in Section 3.1. Therefore, we treat market structure –i.e., the profile of airlines that operate in a market– as our endogenous treatment. We then model air pollution as a function of the market structure as in equation (3.1), where the vector $\boldsymbol{D}$ represents the market structure, and X includes market specific covariates that affect pollution directly (i.e., not through airline activity). We use the share of pollution-related activity in the local economy.[15] This is likely to be excluded from (3.2) if we condition on the size of the market.[16] Hence, we allow for market-level covariates, $\boldsymbol{W}$, which affect both the participation decisions and pollution (e.g., the size of the market as measured by population or the level of economic activity). Finally, we consider two sets of instruments: the airport presence of an airline as in Berry (1992) and a firm-market proxy for cost as in Ciliberto and Tamer (2009). For reasons we discuss below, we use the latter in our baseline results.

To simplify the estimation, we discretize all continuous variables into binary variables. We

[13]In this section, we refer to market structure as the particular configuration of airlines present in the market. In other words, market structure not only refers to the number of firms competing in a given market but to the actual identities of the firms. Thus, we will regard a market in which, say, United and American operate as different from a market in which Southwest and Delta operate, despite both markets having two carriers.

[14]This is not surprising given that the two pollution measures are highly correlated.

[15]Note that our definition of market is a city-pair; hence, all of our variables are, in fact, weighted averages over the two cities.

[16]The idea here is that the size of the market, among other things, determines whether a firm might enter it, but not the type of economic activity in the cities. Hence, conditional on the market's GDP, a market with a higher share of polluting industries will have a higher level of pollution but this share would not affect the airline market structure. That is, we are implicitly assuming that pollution activity generates the same business air travel as non-polluting activities, conditional on size.

experimented with several specifications of the covariates, X and $\boldsymbol{W}$, and instruments, $\boldsymbol{Z}$. In particular, we tried different discretizations of each variable (including allowing for more than two points in their supports and different cutoffs). Clearly, there is a limit to how finely we can cut the data even with a large sample size such as ours. The coarser discretization occurs when each covariate (and instrument) is binary and this seems to produce reasonable results; hence, we stick with this discretization in all of our exercises and use the median of each variable as the threshold.[17][18] Of particular interest is the discretization of the instruments. Recall that our procedure relies on there being enough variation in the instruments as stated in Assumption EQ. When the payoff unobservables are mutually independent, there is a simple testable implication for the existence of such instrument variation as we have discussed in Section 3.3. When we consider the binary version of our cost instrument, the conditions in Assumption EQ* are satisfied in our data by setting the vectors $\boldsymbol{z}$ and $\boldsymbol{z}'$ to be low cost and high cost for all airlines, respectively.[19]

Before moving to the results, we discuss the plausibility of the other identifying assumptions. Assumptions SS and M1 relate to the shape of the profit functions. The strategic substitutability assumption (SS) is likely to hold in our case, as firms' profits are likely to decrease with competition (see, e.g., the argument in Section 3.1 in Berry (1992)). The instrument monotonicity assumption (M1) is also likely to hold since as firms costs move from low to high their profits are expected to decrease regardless of the market structure. Assumptions IN and EX are related to the validity of the instruments and the exogeneity of X. Given our choice of $\boldsymbol{Z}$ and X, both assumptions are plausible as we argue in Section E of the Appendix in which we introduce the data. Finally, Assumption M*(i) is imposed as a threshold crossing model for pollution. Assumption M*(ii) is also likely to be satisfied in our context, by assuming that after entering firms play a Cournot game to set their prices and quantities (i.e., the number of passengers flown in each market). Then it is expected that more passengers will be served in equilibrium as a result of an additional firm entering the market, and thus pollution worsens, regardless of the market structure.

To illustrate our estimation procedure, we consider three types of ATE exercises. The first

[17] We find that, in general, our qualitative results are robust to alternative definitions of the thresholds like the mean, median, or mode of each continuous variable.

[18] Also, after some experimentation, we obtained reasonable results when both X and $\boldsymbol{W}$ are scalars: share of pollution related industries in the market and total GDP in the market, respectively. Hence, we consider this parsimonious model in our baseline specification.

[19] When we use the (binary) market presence instrument, there are instances in which the instrument variation is not enough to satisfy Assumption EQ*. Thus, in what follows we use only use our cost instrument.

examines the effects of a single (monopolist) airline on pollution vis-a-vis a market that is not served by any airline. The second set of exercises examine the total effect of the industry on pollution under all possible market configurations. Finally, the third type of exercises examine how the (marginal) effect of a given airline changes when the firm faces different levels of competition. Notice that in all cases we quantify "reduced-form" effects, in that they summarize structural effects resulting from a given market structure. The idea is that following entry, firms compete by choosing their prices, frequency, and which airplanes to operate. Different market structures will have different impacts on these variables, which in turn, affect the level of pollution.

4.1 Estimation and Results

Using the notation from Section 3.1, let the elements of the treatment vector $\boldsymbol{d} = (d_{\text{DL}}, d_{\text{AA}}, d_{\text{UA}}, d_{\text{WN}}, d_{\text{med}}, d_{\text{low}})$ be either 0 or 1, indicating whether each firm (American (AA), Delta (DL), United (UA), Southwest (WN), a medium-size airline, and a low-cost carrier) is active in the market. We compute the upper and lower bounds on the ATE using the result from Theorem 3.3 and the fact that our Y variable is binary. Specifically, given two treatment vectors $\boldsymbol{d}$ and $\tilde{\boldsymbol{d}}$ we can bound the ATE

$$L(\boldsymbol{d}, \tilde{\boldsymbol{d}}; x, w) \leq E[Y(\boldsymbol{d}) - Y(\tilde{\boldsymbol{d}})|x, w] \leq U(\boldsymbol{d}, \tilde{\boldsymbol{d}}; x, w)$$

where the upper bound can be characterized by

$$\begin{aligned} U(\boldsymbol{d}, \tilde{\boldsymbol{d}}; x, w) \equiv \Pr[Y = 1, \boldsymbol{D} = \boldsymbol{d}|\boldsymbol{z}, x, w] + \sum_{\boldsymbol{d}' \in \mathcal{D}^j \setminus \{\boldsymbol{d}\}} \Pr[\boldsymbol{D} = \boldsymbol{d}'|\boldsymbol{Z} = \boldsymbol{z}, W = w] \\ + \sum_{\boldsymbol{d}' \in \mathcal{D}^<(\boldsymbol{d}) \cup \mathcal{D}^>(\boldsymbol{d})} \Pr[Y = 1, \boldsymbol{D} = \boldsymbol{d}'|\boldsymbol{Z} = \boldsymbol{z}, X = x'(\boldsymbol{d}'), W = w] \\ - \Pr[Y = 1, \boldsymbol{D} = \tilde{\boldsymbol{d}}|\boldsymbol{Z} = \boldsymbol{z}, X = x, W = w] \\ - \sum_{\boldsymbol{d}'' \in \mathcal{D}^<(\tilde{\boldsymbol{d}}) \cup \mathcal{D}^>(\tilde{\boldsymbol{d}})} \Pr[Y = 1, \boldsymbol{D} = \boldsymbol{d}''|\boldsymbol{Z} = \boldsymbol{z}, X = x''(\boldsymbol{d}''), W = w] \end{aligned}$$

for every $\boldsymbol{z}$, $x'(\boldsymbol{d}') \in \mathcal{X}^U_{\boldsymbol{d}}(x;\boldsymbol{d}')$ for $\boldsymbol{d}' \neq \boldsymbol{d}$, and $x''(\boldsymbol{d}'') \in \mathcal{X}^L_{\tilde{\boldsymbol{d}}}(x;\boldsymbol{d}'')$ for $\boldsymbol{d}'' \neq \tilde{\boldsymbol{d}}$ and the lower bound can be characterized by

$$\begin{aligned}
L(\boldsymbol{d},\tilde{\boldsymbol{d}};x,w) \equiv\ & \Pr[Y=1,\boldsymbol{D}=\boldsymbol{d}|\boldsymbol{Z}=\boldsymbol{z},X=x,W=w] \\
& + \sum_{\boldsymbol{d}'\in\mathcal{D}^<(\boldsymbol{d})\cup\mathcal{D}^>(\boldsymbol{d})} \Pr[Y=1,\boldsymbol{D}=\boldsymbol{d}'|\boldsymbol{Z}=\boldsymbol{z},X=x'(\boldsymbol{d}'),W=w] \\
& - \Pr[Y=1,\boldsymbol{D}=\tilde{\boldsymbol{d}}|\boldsymbol{Z}=\boldsymbol{z},X=x,W=w] - \sum_{\boldsymbol{d}''\in\mathcal{D}^j\setminus\{\tilde{\boldsymbol{d}}\}} \Pr[\boldsymbol{D}=\boldsymbol{d}''|\boldsymbol{Z}=\boldsymbol{z},W=w] \\
& - \sum_{\boldsymbol{d}''\in\mathcal{D}^<(\tilde{\boldsymbol{d}})\cup\mathcal{D}^>(\tilde{\boldsymbol{d}})} \Pr[Y=1,\boldsymbol{D}=\boldsymbol{d}''|\boldsymbol{Z}=\boldsymbol{z},X=x''(\boldsymbol{d}''),W=w]
\end{aligned}$$

for every $\boldsymbol{z}$, $x'(\boldsymbol{d}') \in \mathcal{X}^L_{\boldsymbol{d}}(x;\boldsymbol{d}')$ for $\boldsymbol{d}' \neq \boldsymbol{d}$, and $x''(\boldsymbol{d}'') \in \mathcal{X}^U_{\tilde{\boldsymbol{d}}}(x;\boldsymbol{d}'')$ for $\boldsymbol{d}'' \neq \tilde{\boldsymbol{d}}$. We estimate the population objects above using their sample counterparts. We also compute confidence sets by deriving unconditional moment inequalities from our conditional moment inequalities and implementing the Generalized Moment Selection test proposed by Andrews and Soares (2010). The confidence sets are obtained by inverting the test.[20]

Monopoly Effects. Here we examine the ATE of a change in market structure from no airline serving a market to a monopolist serving it. Intuitively, we want to understand the change in the probability of being a high-pollution market when an airline starts operating on it. Recall that we allow each firm to have different effects on pollution; hence, we estimate the effects of each one of the six firms in our data becoming a monopolist. Thus, we are interested in the ATE's of the form $E[Y(\boldsymbol{d}_{\text{monop}}) - Y(\boldsymbol{d}_{\text{noserv}})|X,W]$ where $\boldsymbol{d}_{\text{monop}}$ is one of the six vectors in which only one element is 1 and the rest are 0's, and $\boldsymbol{d}_{\text{noserv}}$ is a vector of all 0's. The results are shown in Figure 2, where the solid black intervals are our estimates of the identified sets and the thin red lines are the 95% confidence sets. We see that all ATE's are positive and statistically significant different from 0, except for the low-cost carriers. While there no major differences on the effects of the major carriers, with the exception of Delta which seems to induce a higher increase in the probability of high pollution, the medium and low-cost carriers induce a smaller effect.

Total Market Structure Effect. We now turn to our second set of exercises. Here, we quantify the effect of the airline industry on the likelihood of a market having high levels of pollution. To do so, we estimate ATE's of the form $E[Y(\boldsymbol{d})-Y(\boldsymbol{d}_{\text{noserv}})|X,W]$ for all potential market configurations

[20] For details of this procedure, see Dickstein and Morales (2018).

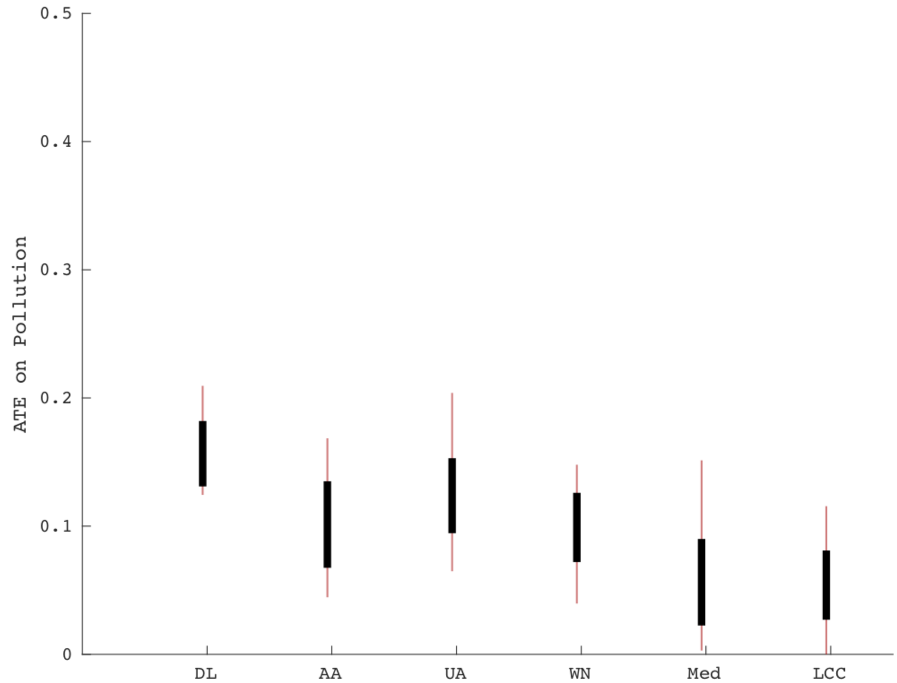


Figure 2: Effect of a Monopolistic Market Structure

This plot shows the ATE's of a change in market structure from no airline serving a market to a monopolist serving it. The solid black intervals are our estimates of the identified sets and the thin red lines are the 95% confidence sets.

$\boldsymbol{d}$ and where $\boldsymbol{d}_{\text{noserv}}$ is defined as before. Figure 3 depicts the results. The left-most set of intervals corresponds to the 6 different monopolistic market structures, and by construction, coincide with those from Figure 2. The next set corresponds to all possible duopolistic structures (15 possibilities), and so on. Not surprisingly, we observe that the effect on the probability of being a high-pollution market is increasing in the number of firms operating in the market. More interesting is the non-linearity of the effect: the effect increases at a decreasing rate. This would be consistent with a model in which firms *accommodate* new entrants by decreasing their frequency, which is analogous to the prediction of a Cournot competition model, as we increase the number of firms. To further investigate this point, in the next set of exercises, we examine the effect of one firm as we change the competition it faces.

Marginal Carrier Effect. In our last set of exercises, we investigate how the marginal effect (i.e., the effect of introducing one more firm into the market) changes under different market configurations. In particular, we are interested in the effect of Delta entering the market, given that the current market structure (excluding Delta) is $\boldsymbol{d}_{\text{-DL}} \equiv (d_{\text{AA}}, d_{\text{UA}}, d_{\text{WN}}, d_{\text{med}}, d_{\text{low}})$. Thus, we

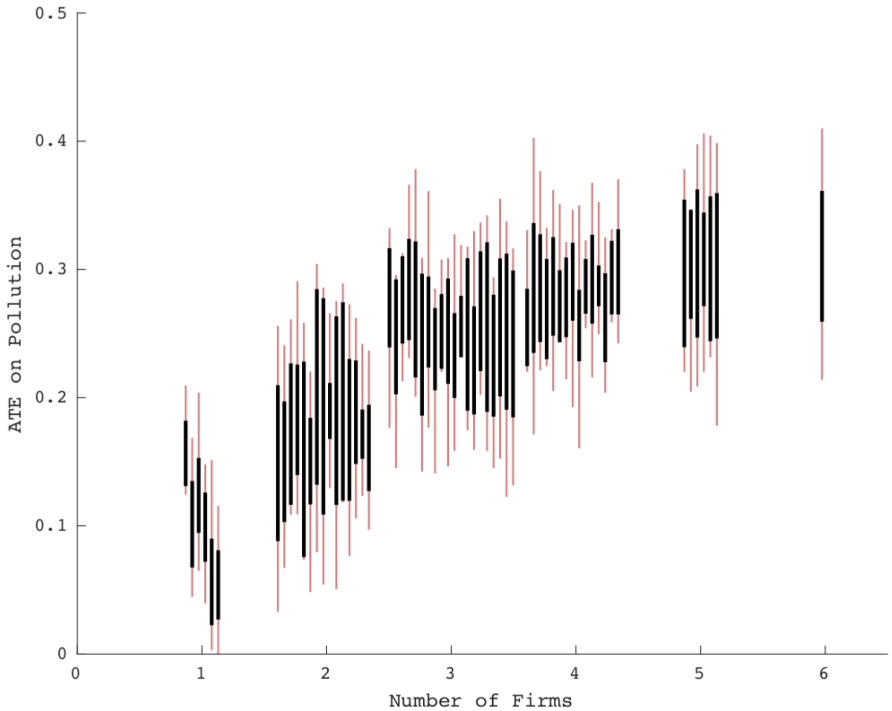


Figure 3: Total Market Structure Effect

This plot shows the ATE's of the airline industry under all possible market configurations. The solid black intervals are our estimates of the identified sets and the thin red lines are the 95% confidence sets. The bars in each cluster correspond to all possible market configurations, respectively.

estimate $E[Y((1, \boldsymbol{d}_{\text{-DL}})) - Y((0, \boldsymbol{d}_{\text{-DL}}))|X, W]$.[21]

Figure 4 shows the identified sets and confidence sets of the marginal effect of Delta on the probability of high pollution under all possible market configuration for Delta's rivals. In the Figure, the left-most exercise is the effect of Delta as a monopolist and coincides, by construction, with the left-most exercise in Figure 2. The second exercise (from the left) corresponds to the additional effect of Delta on pollution when there is already one firm operating in the market, which yields five different possibilities. The next exercise shows the effect of Delta when there are two firms already operating in the market yielding 10 possibilities, and so on. Again, the estimated marginal ATE's in all cases are positive and statistically significant. Interestingly, although we cannot entirely reject the null hypothesis that all the effects are the same, it seems that the marginal effect of Delta is decreasing in the number of rivals it faces. Intuitively, this suggests a situation in which after Delta enters the market, it (or its rivals) operates with a frequency that is decreasing with the number of rivals (again, as we would expect in a Cournot competition model) and is consistent with the

[21] We obtain qualitatively similar results when estimating the marginal effects of the other five carriers, and hence, we omit the graphs to save space.

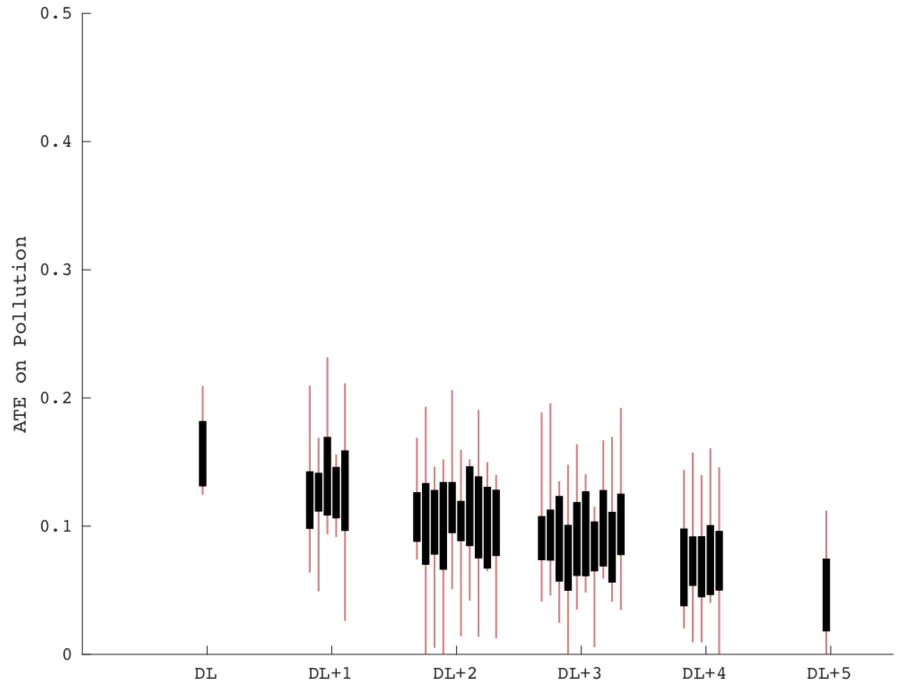


Figure 4: Marginal Effect of Delta under Different Market Structures

This plot shows the ATE's of Delta entering the market given all possible rivals market configurations. The solid black intervals are our estimates of the identified sets and the thin red lines are the 95% confidence sets. The bars in each cluster correspond to all possible market configurations, respectively.

findings in our previous set of exercises.

The conclusions from the *total market* and *marginal* ATE's are also interesting from a policy perspective. For example, a merger of two airlines in which duplicate routes are eliminated would imply a decrease in total pollution in the affected markets, but by less than what one would have naively anticipated from removing one airline while keeping everything else constant. In other words, there are two effects of removing an airline from a market. The first is a direct affect: pollution decreases by the amount of pollution by the carrier that is no longer present in the market. However, the remaining firms in the market will react strategically to the new market structure. In our exercises, we find that this indirect effect implies an increase in pollution. The overall effect is a net decrease in pollution. Moreover, given the non-linearities of the ATE's we estimate it looks like the overall effect, while negative, might be negligible in markets with four or more competitors.

While it is unclear that merger analysis, which is typically concerned with price increases post-merge or cost savings of the merging firms, should also consider externalities such as pollution, (social) welfare analysis should. Hence, our findings may serve as guidance to policy discussion on air traffic regulation.

References

Andrews, D. W. and M. M. Schafgans (1998): "Semiparametric estimation of the intercept of a sample selection model," *The Review of Economic Studies*, 65, 497–517.

Andrews, D. W. K. and G. Soares (2010): "Inference for Parameters Defined by Moment Inequalities Using Generalized Moment Selection," *Econometrica*, 78, 119–157.

Bajari, P., H. Hong, and S. P. Ryan (2010): "Identification and estimation of a discrete game of complete information," *Econometrica*, 78, 1529–1568.

Beresteanu, A., I. Molchanov, and F. Molinari (2011): "Sharp identification regions in models with convex moment predictions," *Econometrica*, 79, 1785–1821.

Berry, S. T. (1992): "Estimation of a Model of Entry in the Airline Industry," *Econometrica: Journal of the Econometric Society*, 889–917.

Borenstein, S. (1989): "Hubs and High Fares: Dominance and Market Power in the US Airline Industry," *RAND Journal of Economics*, 20, 344–365.

Chay, K. Y. and M. Greenstone (2003): "The Impact of Air Pollution on Infant mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession," *The Quarterly Journal of Economics*, 1121–1167.

Chesher, A. (2005): "Nonparametric identification under discrete variation," *Econometrica*, 73, 1525–1550.

Chesher, A. and A. Rosen (2017): "Generalized instrumental variable models," *Econometrica, forthcoming.*

Chiburis, R. C. (2010): "Semiparametric bounds on treatment effects," *Journal of Econometrics*, 159, 267–275.

Ciliberto, F., C. Murry, and E. Tamer (2018): "Market Structure and Competition in Airline Markets," *University of Virginia, Penn State University, Harvard University.*

Ciliberto, F. and E. Tamer (2009): "Market structure and multiple equilibria in airline markets," *Econometrica*, 77, 1791–1828.

Dickstein, M. J. and E. Morales (2018): "What do Exporters Know?" *The Quarterly Journal of Economics*, 133, 1753—1801.

Federal Aviation Administration (2015): "Aviation Emissions, Impacts & Mitigation: A Primer," *FAA, Office of Environment and Energy.*

Foster, A. and M. Rosenzweig (2008): "Inequality and the sustainability of agricultural productivity growth: Groundwater and the Green Revolution in rural India," in *Prepared for the India Policy Conference at Stanford University.*

Galichon, A. and M. Henry (2011): "Set identification in models with multiple equilibria," *The Review of Economic Studies*, 78, 1264–1298.

Gentzkow, M., J. M. Shapiro, and M. Sinkinson (2011): "The effect of newspaper entry and exit on electoral politics," *The American Economic Review*, 101, 2980–3018.

Goolsbee, A. and C. Syverson (2008): "How do incumbents respond to the threat of entry? Evidence from the major airlines," *The Quarterly Journal of Economics*, 123, 1611–1633.

Han, S. (2019a): "Identification in Nonparametric Models for Dynamic Treatment Effects," *UT Austin.*

——— (2019b): "Optimal Dynamic Treatment Regimes from Studies with Imperfect Compliance," *UT Austin.*

Heckman, J. and R. Pinto (2018): "Unordered monotonicity," *Econometrica*, 86, 1–35.

Heckman, J. J., S. Urzua, and E. Vytlacil (2006): "Understanding instrumental variables in models with essential heterogeneity," *The Review of Economics and Statistics*, 88, 389–432.

Heckman, J. J. and E. Vytlacil (2005): "Structural equations, treatment effects, and econometric policy evaluation1," *Econometrica*, 73, 669–738.

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

Jun, S. J., J. Pinkse, and H. Xu (2011): "Tighter bounds in triangular systems," *Journal of Econometrics*, 161, 122–128.

Khan, S. and E. Tamer (2010): "Irregular identification, support conditions, and inverse weight estimation," *Econometrica*, 78, 2021–2042.

Knittel, C. R., D. Miller, and N. J. Sanders (2011): "Caution, drivers! Children present. Traffic, pollution, and infant health," *working paper.*

Lee, S. and B. Salanié (2018): "Identifying effects of multivalued treatments," *Econometrica, forthcoming.*

Li, S., J. Mazur, Y. Park, J. Roberts, A. Sweeting, and J. Zhang (2018): "Endogenous and Selective Service Choices After Airline Mergers," *working paper.*

Manski, C. F. (1990): "Nonparametric bounds on treatment effects," *The American Economic Review*, 80, 319–323.

——— (1997): "Monotone treatment response," *Econometrica: Journal of the Econometric Society*, 1311–1334.

——— (2013): "Identification of treatment response with social interactions," *The Econometrics Journal*, 16, S1–S23.

Manski, C. F. and J. V. Pepper (2000): "Monotone instrumental variables: With an application to the returns to schooling," *Econometrica*, 68, 997–1010.

Mourifié, I. (2015): "Sharp bounds on treatment effects in a binary triangular system," *Journal of Econometrics*, 187, 74–81.

Schlenker, W. and W. R. Walker (2015): "Airports, air pollution, and contemporaneous health," *The Review of Economic Studies*, 83, 768–809.

Sekhri, S. (2014): "Wells, water, and welfare: the impact of access to groundwater on rural poverty and conflict," *American Economic Journal: Applied Economics*, 6, 76–102.

Shaikh, A. M. and E. J. Vytlacil (2011): "Partial identification in triangular systems of equations with binary dependent variables," *Econometrica*, 79, 949–955.

Tamer, E. (2003): "Incomplete simultaneous discrete response model with multiple equilibria," *The Review of Economic Studies*, 70, 147–165.

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

Vytlacil, E. and N. Yildiz (2007): "Dummy endogenous variables in weakly separable models," *Econometrica*, 75, 757–779.

Walker, R. E., C. R. Keane, and J. G. Burke (2010): "Disparities and access to healthy food in the United States: A review of food deserts literature," *Health & place*, 16, 876–884.

A More Examples

Example 1 (Media and political behavior). *In this example, the interest is in how media affects political participation or electoral competitiveness. In county or market i, either $Y_i \in [0,1]$ can denote voter turnout, or $Y_i \in \{0,1\}$ can denote whether an incumbent is re-elected or not. Let $D_{s,i}$ denote the market entry decision by local newspaper type s, which is correlated with unobserved characteristics of the county. In this example, $Z_{s,i}$ can be the neighborhood counties' population size and income, which is common to all players ($Z_{1,i} = \cdots = Z_{S,i}$). Lastly, X_i can include changes in voter ID regulations. Using a linear panel data model, Gentzkow et al. (2011) show that the number of newspapers in the market significantly affects the voter turnout but find no evidence whether it affects the re-election of incumbents. More explicit modeling of the strategic interaction among newspaper companies can be important to capture competition effects on political behavior of the readers.*

Example 2 (Incumbents' response to potential entrants). *In this example, we are interested in how market i's incumbents respond to the threat of entry of potential competitors. Let Y_i be an incumbent firm's pricing or investment decision and $D_{s,i}$ be an entry decision by firm s in "nearby" markets, which can be formally defined in each context. For example, in airline entry, nearby markets are defined as city pairs that share the endpoints with the city pair of an incumbent (Goolsbee and Syverson (2008)). That is, potential entrants are airlines that operate in one (or both) of the endpoints of the incumbent's market i, but who have not connected these endpoints. Then the* *parameter $E[Y_i(\boldsymbol{d}) - Y_i(\boldsymbol{d}')]$ captures the incumbent's response to the threat, specifically whether it responds by lowering the price or making an investment. As in Example 1, $Z_{s,i}$ are cost shifters and X_i are other factors affecting price of the incumbent, excluded from nearby markets, conditional of W_i. The characteristics of the incumbent's market can be a candidate of X_i, such as the distance between the endpoints of the incumbent's market in the airline example.*

Example 3 (Food desert)**.** *Let Y_i denote a health outcome, such as diabetes prevalence, in region i, and $D_{s,i}$ be the exit decision by large supermarket s in the region. Then $E[Y_i(\boldsymbol{d}) - Y_i(\boldsymbol{d}')]$ measures the effects of absence of supermarkets on health of the residents. Conditional on other factors W_i, the instrument $Z_{s,i}$ can include changes in local government's zoning plans and X_i can include the region's health-related variables, such as the number of hospitals and the obesity rate. This problem is related to the literature on "food desert" (e.g., Walker et al. (2010)).*

Example 4 (Ground water and agriculture)**.** *In this example, we are interested in the impact of access to groundwater on economic outcomes in rural areas (Foster and Rosenzweig (2008)). In each Indian village i, symmetric wealthy farmers (of the same caste) make irrigation decisions $D_{s,i}$, i.e., whether or not to buy motor pumps, in the presence of peer effects and learning spillovers. Since ground water is a limited resource that is seasonally recharged and depleted, other farmers' entry may negatively affects one's payoff. The adoption of the technology affects Y_i, which can be the average of local wages of peasants or prices of agricultural products, or a village development or poverty level. In this example, continuous or binary instrument $Z_{s,i}$ can be the depth to groundwater, which is exogenously given (Sekhri (2014)), or provision of electricity for pumping in a randomized field experiment. X_i can be village-level characteristics that villagers do not know ex ante or do not concern about.*[22]

[22]Especially in this example, the number of players/treatments S_i is allowed to vary across villages. We assume in this case that players/treatments are symmetric (in a sense that becomes clear later) and $\nu^1(\cdot) = \cdots = \nu^{S_i}(\cdot) = \nu(\cdot)$.

B Numerical Study

To illustrate the main results of this study in a simulation exercise, we calculate the bounds on the ATE using the following data generating process:

$$Y_{\boldsymbol{d}} = 1[\tilde{\mu}_{\boldsymbol{d}} + \beta X \geq \epsilon],$$

$$D_1 = 1[\delta_2 D_2 + \gamma_1 Z_1 \geq V_1],$$

$$D_2 = 1[\delta_1 D_1 + \gamma_2 Z_2 \geq V_2],$$

where (ϵ, V_1, V_2) are drawn, independent of $(X, \boldsymbol{Z})$, from a joint normal distribution with zero means and each correlation coefficient being 0.5. We draw Z_s $(s = 1, 2)$ and X from a multinomial distribution, allowing Z_s to take two values, $\mathcal{Z}_s = \{-1, 1\}$, and X to take either three values, $\mathcal{X} = \{-1, 0, 1\}$, or fifteen values, $\mathcal{X} = \{-1, -\frac{6}{7}, -\frac{5}{7}, ..., \frac{5}{7}, \frac{6}{7}, 1\}$. Being consistent with Assumption M, we choose $\tilde{\mu}_{11} > \tilde{\mu}_{10}$ and $\tilde{\mu}_{01} > \tilde{\mu}_{00}$. Let $\tilde{\mu}_{10} = \tilde{\mu}_{01}$. With Assumption SS, we choose $\delta_1 < 0$ and $\delta_2 < 0$. Without loss of generality, we choose positive values for γ_1, γ_2, and β. Specifically, $\tilde{\mu}_{11} = 0.25$, $\tilde{\mu}_{10} = \tilde{\mu}_{01} = 0$ and $\tilde{\mu}_{00} = -0.25$. For default values, $\delta_1 = \delta_2 \equiv \delta = -0.1$, $\gamma_1 = \gamma_2 \equiv \gamma = 1$ and $\beta = 0.5$.

In this exercise, we focus on the ATE $E[Y(1,1) - Y(0,0)|X = 0]$, whose true value is 0.2 given our choice of parameter values. For $h(\boldsymbol{z}, \boldsymbol{z}', x)$, we consider $\boldsymbol{z} = (1, 1)$ and $\boldsymbol{z}' = (-1, -1)$. Note that $H(x) = h(\boldsymbol{z}, \boldsymbol{z}', x)$ and $\tilde{H}(x, x', x'') = \tilde{h}(\boldsymbol{z}, \boldsymbol{z}', x, x', x'')$, since Z_s is binary. Then, we can derive the sets $\mathcal{X}^U_{\boldsymbol{d}}(0; \boldsymbol{d}')$ and $\mathcal{X}^L_{\boldsymbol{d}}(0; \boldsymbol{d}')$ for each $\boldsymbol{d} \in \{(1, 1), (0, 0)\}$ and $\boldsymbol{d}' \neq \boldsymbol{d}$ in Theorem 3.3.

Based on our design, $H(0) > 0$, and thus, the bounds when we use Z only are, with $x = 0$,

$$\max_{\boldsymbol{z} \in \mathcal{Z}} \Pr[Y = 1, \boldsymbol{D} = (0, 0)|\boldsymbol{z}, x] \leq \Pr[Y(0, 0) = 1|x] \leq \min_{\boldsymbol{z} \in \mathcal{Z}} \Pr[Y = 1|\boldsymbol{z}, x],$$

and

$$\max_{\boldsymbol{z} \in \mathcal{Z}} \Pr[Y = 1|\boldsymbol{z}, x] \leq \Pr[Y(1, 1) = 1|x] \leq \min_{\boldsymbol{z} \in \mathcal{Z}} \{\Pr[Y = 1, \boldsymbol{D} = (1, 1)|\boldsymbol{z}, x] + 1 - \Pr[\boldsymbol{D} = (1, 1)|\boldsymbol{z}, x]\}.$$

Using both $\boldsymbol{Z}$ and X, we obtain narrower bounds. For example, when $|\mathcal{X}| = 3$, with $\tilde{H}(0, -1, -1) <$

0, the lower bound on $\Pr[Y(0,0)=1|X=0]$ becomes

$$\max_{\boldsymbol{z}\in\mathcal{Z}}\left\{\Pr[Y=1,\boldsymbol{D}=(0,0)|\boldsymbol{z},0]+\Pr[Y=1,\boldsymbol{D}\in\{(1,0),(0,1)\}|\boldsymbol{z},-1]\right\}.$$

With $\tilde{H}(1,1,0)<0$, the upper bound on $\Pr[Y(1,1)=1|X=0]$ becomes

$$\min_{\boldsymbol{z}\in\mathcal{Z}}\left\{\Pr[Y=1,\boldsymbol{D}=(1,1)|\boldsymbol{z},0]+\Pr[Y=1,\boldsymbol{D}\in\{(1,0),(0,1)\}|\boldsymbol{z},1]+\Pr[\boldsymbol{D}=(0,0)|\boldsymbol{z},0]\right\}.$$

For comparison, we calculate the bounds in Manski (1990) using $\boldsymbol{Z}$. These bounds are given by

$$\begin{aligned}&\max_{\boldsymbol{z}\in\mathcal{Z}}\Pr[Y=1,\boldsymbol{D}=(0,0)|\boldsymbol{z},x]\leq\Pr[Y(0,0)=1|x]\\&\leq\min_{\boldsymbol{z}\in\mathcal{Z}}\left\{\Pr[Y=1,\boldsymbol{D}=(0,0)|\boldsymbol{z},x]+1-\Pr[\boldsymbol{D}=(0,0)|\boldsymbol{z}]\right\},\end{aligned}$$

and

$$\begin{aligned}&\max_{\boldsymbol{z}\in\mathcal{Z}}\Pr[Y=1,\boldsymbol{D}=(1,1)|\boldsymbol{z},x]\leq\Pr[Y(1,1)=1|x]\\&\leq\min_{\boldsymbol{z}\in\mathcal{Z}}\left\{\Pr[Y=1,\boldsymbol{D}=(1,1)|\boldsymbol{z},x]+1-\Pr[\boldsymbol{D}=(1,1)|\boldsymbol{z}]\right\}.\end{aligned}$$

We also compare the estimated ATE using a standard linear IV model in which the nonlinearity of the true DGP is ignored:

$$\begin{aligned}Y&=\pi_0+\pi_1D_1+\pi_2D_2+\beta X+\epsilon,\\\begin{pmatrix}D_1\\D_2\end{pmatrix}&=\begin{pmatrix}\gamma_{10}\\\gamma_{20}\end{pmatrix}+\begin{pmatrix}\gamma_{11}&\gamma_{12}\\\gamma_{21}&\gamma_{22}\end{pmatrix}\begin{pmatrix}Z_1\\Z_2\end{pmatrix}+\begin{pmatrix}V_1\\V_2\end{pmatrix}.\end{aligned}$$

Here, the first stage is the reduced-form representation of the linear simultaneous equations model for strategic interaction. Under this specification, the ATE becomes $E[Y(1,1)-Y(0,0)|X=0]=\pi_1+\pi_2$, which is estimated via two-stage least squares (TSLS).

The bounds calculated for the ATE are shown in Figures 5–8. Figure 5 shows how the bounds on the ATE change, as the value of γ changes from 0 to 2.5. The larger γ is, the stronger the instrument $\boldsymbol{Z}$ is. The first conspicuous result is that the TSLS estimate of the ATE is biased because of the problem of misspecification. Next, as expected, Manski's bounds and our proposed bounds converge to the true value of the ATE as the instrument becomes stronger. Overall, our bounds, with or without exploiting the variation of X, are much narrower than Manski's bounds.[23] Notice that the sign of the ATE is identified in the whole range of γ, as predicted by the first part of Theorem 3.3, in contrast to Manski's bounds. Using the additional variation in X with $|\mathcal{X}| = 3$ decreases the width of the bounds, particularly with the smaller upper bounds on the ATE in this simulation design. Figure 6 depicts the bounds using X with $|\mathcal{X}| = 15$, which yields narrower bounds than when $|\mathcal{X}| = 3$, and substantially narrower than those only using $\boldsymbol{Z}$.

Figure 7 shows how the bounds change as the value of β changes from 0 to 1.5, where a larger β corresponds to a stronger exogenous variable X. The jumps in the upper bound are associated with the sudden changes in the signs of $\tilde{H}(-1, 0, -1)$ and $\tilde{H}(0, 1, 1)$. At least in this simulation design, the strength of X is not a crucial factor for obtaining narrower bounds. In fact, based on other simulation results (omitted in the paper), we conclude that the number of values X can take matters more than the dispersion of X (unless we pursue point identification of the ATE).

Finally, Figure 8 shows how the width of the bounds is related to the extent to which the opponents' actions D_{-s} affect one's payoff, captured by δ. We vary the value of δ from -2 to 0, and when $\delta = 0$, the players solve a single-agent optimization problem. Thus, heuristically, the bound at this point would be similar to the ones that can be obtained when Shaikh and Vytlacil (2011) is extended to a multiple-treatment setting with no simultaneity. In the figure, as the value of δ becomes smaller, the bounds get narrower.

C Discussions and Extensions

C.1 Point Identification of the ATE

When there exist player-specific excluded instruments with large support, we point identify the ATE's. To invoke an identification-at-infinity argument, the following assumptions are instead

[23] Although we do not make a rigorous comparison of the assumptions here, note that the bounds by Manski and Pepper (2000) under the semi-MTR is expected to be similar to ours. However, their bounds need to assume the direction of the monotonicity.

needed to hold:

$$\gamma_1 \text{ and } \gamma_2 \text{ are nonzero,} \tag{C.1}$$

$$Z_1|(X, Z_2) \text{ and } Z_2|(X, Z_1) \text{ has an everywhere positive Lebesgue density.} \tag{C.2}$$

These assumptions impose a player-specific exclusion restriction and large support. Under (C.1)–(C.2), we can easily show that the ATE in (3.3) is point identified. In this case, the structure we impose, especially on the outcome function (such as the threshold-crossing structure, or more generally Assumption M in Section 3.3 below) is not needed.

The identification strategy is to exploit the large variation of player specific instruments based on (C.1)–(C.2), which simultaneously solves the multiple equilibria and the endogeneity problems. For example, to identify $E[Y(1,1)|x]$, consider

$$\begin{aligned} &E[Y|\boldsymbol{D} = (1,1), \boldsymbol{z}, x] = E[Y(1,1)|\boldsymbol{D} = (1,1), \boldsymbol{z}, x] \\ &= E[Y(1,1)|\delta_2 + \gamma_1 z_1 \geq U_1, \delta_1 + \gamma_2 z_2 \geq U_2, x] \to E[Y(1,1)|x], \end{aligned}$$

where the second equation is by (2.6) and $Y(1,1) = \mu_1 + \mu_2 + \beta X$, and the convergence is by (C.1)–(C.2) with $z_1 \to \infty$ and $z_2 \to \infty$. The identification of $E[Y(0,0)|x]$, $E[Y(1,0)|x]$ and $E[Y(0,1)|x]$ can be achieved by similar reasoning. Note that $\boldsymbol{D} = (1,0)$ or $\boldsymbol{D} = (0,1)$ can be predicted as an outcome of multiple equilibria. However, when either $(z_1, z_2) \to (\infty, -\infty)$ or $(z_1, z_2) \to (-\infty, \infty)$ occurs, a unique equilibrium is guaranteed as a dominant strategy, i.e., $\boldsymbol{D} = (1,0)$ or $\boldsymbol{D} = (0,1)$, respectively.

C.2 Non-Monotonicity of Treatment Selection

In the case of a single binary treatment, the standard selection equation exhibits monotonicity that facilitates various identification strategies (e.g., Imbens and Angrist (1994), Heckman and Vytlacil (2005), Vytlacil and Yildiz (2007) to name a few). Relatedly, Vytlacil (2002) shows the equivalence between imposing the selection equation with threshold-crossing structure and assuming the local ATE (LATE) monotonicity. This equivalence (and thus, previous identification strategies) is inapplicable to our setting due to the simultaneity in the first stage (3.2). To formally state this, let $\boldsymbol{D}(\boldsymbol{z})$ be a potential treatment vector, had $\boldsymbol{Z} = \boldsymbol{z}$ been realized. When cost $\boldsymbol{Z} = (Z_1, Z_2)$ increases from $\boldsymbol{z}$ to $\boldsymbol{z}'$, it may be that some markets witness Delta entering and United going out of business (i.e., $\boldsymbol{D}(\boldsymbol{z}) = (0,1)$ and $\boldsymbol{D}(\boldsymbol{z}') = (1,0)$), while other markets witness the opposite (i.e., $\boldsymbol{D}(\boldsymbol{z}) = (1,0)$ and $\boldsymbol{D}(\boldsymbol{z}') = (0,1)$). The direction of monotonicity is reversed in the two groups of markets, and thus, $\Pr[\boldsymbol{D}(\boldsymbol{z}) \geq \boldsymbol{D}(\boldsymbol{z}')] \neq 1$ and $\Pr[\boldsymbol{D}(\boldsymbol{z}) \leq \boldsymbol{D}(\boldsymbol{z}')] \neq 1$ where the inequality for vectors is pairwise inequalities, which violates the LATE monotonicity.[24] Despite this non-monotonic pattern, Theorem 3.1 restores generalized monotonicity, i.e., monotonicity in terms of the algebra of sets. This generalized monotonicity, combined with the compensating strategic substitutability (**??**), allows us to use a strategy analogous to the single-treatment case for our bound analysis. This also suggests that we can introduce a generalized version of the LATE parameter in the current framework, although we do not pursue it in this study.

Related to our study, Lee and Salanié (2018) introduce a framework for treatment effects with general non-monotonicity of selection, and consider the simultaneous treatment selection as one of the examples. Although they engage in a similar discussion on non-monotonicity, their approach to gain tractability for identification is different from ours. When they allow the identity of players being observed as in our setting, they show that their treatment measurability condition (Assumption 2.1) introduced to restore monotonicity is satisfied, provided they assume a threshold-crossing equilibrium selection mechanism. In contrast, we avoid making assumptions on equilibrium selection, but require compensating variation of instruments. In addition, for this particular example, they assume the first-stage is known (i.e., payoff functions are known), and focus on point identification of the MTE with continuous instruments.

C.3 Partial ATE

Define a *partial counterfactual outcome* as follows: with a partition $\boldsymbol{D} = (\boldsymbol{D}_1, \boldsymbol{D}_2) \in \mathcal{D}_1 \times \mathcal{D}_2 = \mathcal{D}$ and its realization $\boldsymbol{d} = (\boldsymbol{d}_1, \boldsymbol{d}_2)$,

$$Y(\boldsymbol{d}_1, \boldsymbol{D}_2) \equiv \sum_{\boldsymbol{d}_2 \in \mathcal{D}_2} 1[\boldsymbol{D}_2 = \boldsymbol{d}_2] Y(\boldsymbol{d}_1, \boldsymbol{d}_2). \tag{C.3}$$

[24] The same argument applies with a scalar multi-valued treatment $\tilde{D} \in \{1,2,3,4\}$, which has a one-to-one map with $\boldsymbol{D} \in \{(0,0),(0,1),(1,0),(1,1)\}$. Then, some markets can experience $\tilde{D}(\boldsymbol{z}) = 2$ and $\tilde{D}(\boldsymbol{z}') = 3$ while others experience $\tilde{D}(\boldsymbol{z}) = 3$ and $\tilde{D}(\boldsymbol{z}') = 2$, and thus, it is possible to have $\Pr[\tilde{D}(\boldsymbol{z}) \geq \tilde{D}(\boldsymbol{z}')] \neq 1$ and $\Pr[\tilde{D}(\boldsymbol{z}) \leq \tilde{D}(\boldsymbol{z}')] \neq 1$.

This is a counterfactual outcome that is fully observed once $\boldsymbol{D}_1 = \boldsymbol{d}_1$ is realized. Then for each $\boldsymbol{d}_1 \in \mathcal{D}_1$, the partial ASF can be defined as

$$E[Y(\boldsymbol{d}_1, \boldsymbol{D}_2)] = \sum_{\boldsymbol{d}_2 \in \mathcal{D}_2} E[Y(\boldsymbol{d}_1, \boldsymbol{d}_2)|\boldsymbol{D}_2 = \boldsymbol{d}_2] \Pr[\boldsymbol{D}_2 = \boldsymbol{d}_2] \tag{C.4}$$

and the *partial ATE* between $\boldsymbol{d}$ and $\boldsymbol{d}'$ as

$$E[Y(\boldsymbol{d}_1, \boldsymbol{D}_2) - Y(\boldsymbol{d}_1', \boldsymbol{D}_2)]. \tag{C.5}$$

Using this concept, we can consider complementarity concentrated on, e.g., the first two treatments:

$$E\left[Y(1,1,\boldsymbol{D}_2) - Y(0,1,\boldsymbol{D}_2)\right] > E\left[Y(1,0,\boldsymbol{D}_2) - Y(0,0,\boldsymbol{D}_2)\right].$$

C.4 Model with Common Z

Consider model (3.1)–(3.2) but with instruments common to all players/treatments, i.e., $Z_1 = \cdots = Z_S$:

$$\begin{aligned} Y &= \theta(\boldsymbol{D}, X, \epsilon_{\boldsymbol{D}}), \\ D_s &= 1\left[\nu^s(\boldsymbol{D}_{-s}, Z_1) \geq U_s\right], \qquad s \in \{1, ..., S\}. \end{aligned}$$

This setting can be motivated by such instruments as appeared in Example 1. Given this model, Assumptions SS, M1, IN, EX and C will be understood with $Z_1 = \cdots = Z_S$ imposed.[25] Then the bound analysis for the ATE including sharpness will naturally follow. The intuition of this straightforward extension is as follows. As a generalized version of monotonicity in the treatment selection process is restored (Theorem 3.1), model (3.1)–(3.2) can essentially be seen as a triangular model with an ordered-choice type of a first-stage. Therefore an instrument that "shift" the entire first-stage process is sufficient for the purpose of our analyses. Player-specific instruments do introduce an additional source of variation, as it is crucial for the point identification of the ATE that employs identification at infinity.

[25] Assumption EQ may be slightly harder to justify with a common instrument.

C.5 Partial Symmetry: Interaction Within Groups

In some cases, strategic interaction may occur within groups of players (i.e., treatments). In the airline example, it may be the case that larger airlines interact with one another as a group, so do smaller airlines as a different group, but there may be no interaction across the groups.[26] In general for K groups of players/treatments, we consider, with player index $s = 1, ..., S_g$ and group index $g = 1, ..., G$,

$$Y = \theta(\boldsymbol{D}_1, ..., \boldsymbol{D}_G, X, \epsilon_{\boldsymbol{D}}), \tag{C.6}$$

$$D_{g,s} = 1\left[\nu_g^s(\boldsymbol{D}_{g,-s}, Z_{g,s}) \geq U_{g,s}\right], \tag{C.7}$$

where each $\boldsymbol{D}_g \equiv (D_{g,1}, ..., D_{g,S_k})$ is the treatment vector of group g and $\boldsymbol{D} \equiv (\boldsymbol{D}_1, ..., \boldsymbol{D}_G)$. This model generalizes the model (3.1)–(3.2). It can also be seen as a special case of exogenously endowing an incomplete undirected network structure, where players interact with one another within each of complete sub-networks. In this model, each group can differ in the number (S_g) and identity of players (under which the entry decision is denoted by $D_{g,s}$). Also, the unobservables $\boldsymbol{U}_g \equiv (U_{g,1}, ..., U_{g,S})$ can be arbitrarily correlated across groups, in addition to the fact that $U_{g,s}$'s can be correlated within group g and $\boldsymbol{U} \equiv (\boldsymbol{U}_1, ..., \boldsymbol{U}_G)$ can be correlated with $\epsilon_{\boldsymbol{D}}$. This partly relaxes the independence assumption across markets, which is frequently imposed in the entry game literature. When $G = 1$, the model (C.6)–(C.7) coincides to (3.1)–(3.2).

To calculate the bounds on the ATE $E[Y(\boldsymbol{d}) - Y(\boldsymbol{d}')|x]$ we apply the results in Theorem 3.3, by adapting the assumptions in Sections 3.2 and 3.3 to the current extension. Although Assumption EQ can also be adapted accordingly, we consider an alternative assumption that may be valid in the current setting. Under this assumption, Assumption EQ is no longer needed for identification.[27] Let $\boldsymbol{D}_g^- \equiv (\boldsymbol{D}_g, ..., \boldsymbol{D}_{g-1}, \boldsymbol{D}_{g+1}, ..., \boldsymbol{D}_G)$ and let its realization be $\boldsymbol{d}_g^-$.

Assumption SY. *For $g = 1, ..., G$ and every $x \in \mathcal{X}$, $\vartheta(\boldsymbol{d}_g, \boldsymbol{d}_g^-, x; \boldsymbol{u}) = \vartheta(\tilde{\boldsymbol{d}}_g, \boldsymbol{d}_g^-, x; \boldsymbol{u})$ a.e. $\boldsymbol{u}$ for any permutation $\tilde{\boldsymbol{d}}_g$ of $\boldsymbol{d}_g$.*

This assumption is a *partial conditional symmetry* assumption. It requires symmetry in the functions within each group g, as long as the observed characteristics X remain the same. When

[26] We can also easily extend the model so that smaller airlines take larger airlines' entry decisions as given and play their own entry game, which may be more reasonable to assume.

[27] Also, Y can be unbounded and thus the second statement in Assumption M is not needed.

$G = 1$, SY is related to an assumption found in Manski (2013).

Under Assumption SY, the bound on the ASF can be calculated by iteratively applying the previous results to each group. Assumptions SS, EX and M can be modified so that they hold for treatments with within-group interaction. In particular, Assumption EX can be modified as follows: for each $\boldsymbol{d}_{g,-s} \in \mathcal{D}_{g,-s}$, $\nu_g^s(\boldsymbol{d}_{g,-s}, Z_{g,s})|X, \boldsymbol{Z}_g^-$ is nondegenerate, where $\boldsymbol{Z} \equiv (\boldsymbol{Z}_g, \boldsymbol{Z}_g^-)$. That is, there must be group-specific instruments that are excluded from other groups.[28]

We briefly show how to modify the previous bound analysis with binary Y and no X for simplicity.

Analogous to the previous notation, let $\mathcal{D}_g^j$ be the set of equilibria with j entrants in group g and let $\mathcal{D}_g^{\leq j} \equiv \bigcup_{k=0}^{j} \mathcal{D}_g^k$. Suppose $G = 2$, and $\boldsymbol{d}_1 \in \{0,1\}^{S_1}$ and $\boldsymbol{d}_2 \in \{0,1\}^{S_2}$. Consider the ASF $E[Y(\boldsymbol{d})] = E[Y(\boldsymbol{d}_1, \boldsymbol{d}_2)]$ with $\boldsymbol{d}_1 \in \mathcal{D}_1^{j-1}$ and $\boldsymbol{d}_2 \in \mathcal{D}_2^{k-1}$ for some $j = 1, ..., S_1$ and $k = 1, ..., S_2$. To calculate its bounds, we can bound $E[Y(\boldsymbol{d})|\boldsymbol{D} = \boldsymbol{d}', \boldsymbol{z}]$ in an expansion similar to (D.30) for $\tilde{\boldsymbol{d}} \neq \boldsymbol{d}$ by sequentially applying the analysis of Section 3.4 in each group. First, consider $\tilde{\boldsymbol{d}} = (\tilde{\boldsymbol{d}}_1, \boldsymbol{d}_2)$ with $\tilde{\boldsymbol{d}}_1 \in \mathcal{D}_1^j$. We apply Lemma 3.2 for the $\boldsymbol{D}_1$ portion after holding $\boldsymbol{D}_2 = \boldsymbol{d}_2$. Suppose

$$
\begin{aligned}
\Pr[Y = 1|\boldsymbol{D}_2 = \boldsymbol{d}_2, \boldsymbol{Z}_1 = \boldsymbol{z}_1, \boldsymbol{Z}_2 = \boldsymbol{z}_2] - \Pr[Y = 1|\boldsymbol{D}_2 = \boldsymbol{d}_2, \boldsymbol{Z}_1 = \boldsymbol{z}_1', \boldsymbol{Z}_2 = \boldsymbol{z}_2] \geq 0, \\
\Pr[\boldsymbol{D}_1 \in \mathcal{D}_1^{>j-1}|\boldsymbol{Z}_1 = \boldsymbol{z}_1] - \Pr[\boldsymbol{D}_1 \in \mathcal{D}_1^{>j-1}|\boldsymbol{Z}_1 = \boldsymbol{z}_1'] > 0,
\end{aligned}
$$

then we have $\mu(\tilde{\boldsymbol{d}}_1, \boldsymbol{d}_2) \geq \mu(\boldsymbol{d}_1, \boldsymbol{d}_2)$. The proof of Lemma 3.2 can be adapted by holding $\boldsymbol{D}_2 = \boldsymbol{d}_2$ in this case, because there is no strategic interaction across groups and therefore the multiple equilibria problem only occurs within each group. Note that this strategy still allows for dependence between $\boldsymbol{D}_1$ and $\boldsymbol{D}_2$ even after conditioning on $\boldsymbol{Z}$ due to dependence between $\boldsymbol{U}_1$ and $\boldsymbol{U}_2$. Then,

$$
\begin{aligned}
\Pr[Y(\boldsymbol{d}_1, \boldsymbol{d}_2) = 1|\boldsymbol{D} = (\tilde{\boldsymbol{d}}_1, \boldsymbol{d}_2), \boldsymbol{z}] &= \Pr[\epsilon \leq \mu(\boldsymbol{d}_1, \boldsymbol{d}_2)|\boldsymbol{D} = (\tilde{\boldsymbol{d}}_1, \boldsymbol{d}_2), \boldsymbol{z}] \\
&\leq \Pr[\epsilon \leq \mu(\tilde{\boldsymbol{d}}_1, \boldsymbol{d}_2)|\boldsymbol{D} = (\tilde{\boldsymbol{d}}_1, \boldsymbol{d}_2), \boldsymbol{z}] \qquad \text{(C.8)} \\
&= \Pr[Y = 1|\boldsymbol{D} = (\tilde{\boldsymbol{d}}_1, \boldsymbol{d}_2), \boldsymbol{z}].
\end{aligned}
$$

Next, consider $\boldsymbol{d} = (\boldsymbol{d}_1, \boldsymbol{d}_2)$ and $\tilde{\boldsymbol{d}} = (\tilde{\boldsymbol{d}}_1, \tilde{\boldsymbol{d}}_2)$ with $\tilde{\boldsymbol{d}}_2 \in \mathcal{D}_2^k$ and the other elements as previously determined. Then, by applying Lemma 3.2 this time for $\boldsymbol{D}_2$ after holding $\boldsymbol{D}_1 = \tilde{\boldsymbol{d}}_1$, we have

[28] We maintain Assumption R in the current setting since the assumption is equivalent to assuming a rank invariance within each group, i.e., $\epsilon_{\boldsymbol{d}^g, \boldsymbol{d}^{-g}} = \epsilon_{\tilde{\boldsymbol{d}}^g, \boldsymbol{d}^{-g}}$ $\forall \boldsymbol{d}^g, \tilde{\boldsymbol{d}}^g \in \{0,1\}^{S_g}$ and $g = 1, ..., G$.

$\mu(\tilde{\boldsymbol{d}}_1, \tilde{\boldsymbol{d}}_2) \geq \mu(\tilde{\boldsymbol{d}}_1, \boldsymbol{d}_2)$ by supposing

$$\Pr[Y = 1|\boldsymbol{D}_1 = \tilde{\boldsymbol{d}}_1, \boldsymbol{Z}_1 = \boldsymbol{z}_1, \boldsymbol{Z}_2 = \boldsymbol{z}_2] - \Pr[Y = 1|\boldsymbol{D}_1 = \tilde{\boldsymbol{d}}_1, \boldsymbol{Z}_1 = \boldsymbol{z}_1, \boldsymbol{Z}_2 = \boldsymbol{z}_2'] \geq 0,$$
$$\Pr[\boldsymbol{D}_2 \in \mathcal{D}_2^{>j-1}|\boldsymbol{Z}_2 = \boldsymbol{z}_2] - \Pr[\boldsymbol{D}_2 \in \mathcal{D}_2^{>j-1}|\boldsymbol{Z}_2 = \boldsymbol{z}_2'] > 0.$$

Then,

$$\begin{aligned}\Pr[Y(\boldsymbol{d}_1, \boldsymbol{d}_2) = 1|\boldsymbol{D} = (\tilde{\boldsymbol{d}}_1, \tilde{\boldsymbol{d}}_2), \boldsymbol{z}] &\leq \Pr[\epsilon \leq \mu(\tilde{\boldsymbol{d}}_1, \boldsymbol{d}_2)|\boldsymbol{D} = (\tilde{\boldsymbol{d}}_1, \tilde{\boldsymbol{d}}_2), \boldsymbol{z}] \\ &\leq \Pr[\epsilon \leq \mu(\tilde{\boldsymbol{d}}_1, \tilde{\boldsymbol{d}}_2)|\boldsymbol{D} = (\tilde{\boldsymbol{d}}_1, \tilde{\boldsymbol{d}}_2), \boldsymbol{z}] \\ &= \Pr[Y = 1|\boldsymbol{D} = (\tilde{\boldsymbol{d}}_1, \tilde{\boldsymbol{d}}_2), \boldsymbol{z}],\end{aligned} \tag{C.9}$$

where the first inequality is by (C.8). Note that in deriving the upper bound in (C.9), it is important that at least the two groups share the same signs of within-group h's and $\tilde{h}$'s.

C.6 Player-Specific Outcomes

So far, we considered a scalar Y that may represent an outcome common to all players in a given market or a geographical region. The outcome, however, can also be an outcome that is specific to each player. In this regard, consider a vector of outcomes $\boldsymbol{Y} = (Y_1, ..., Y_S)$ where each element Y_s is a player-specific outcome. An interesting example of this setting may be where $\boldsymbol{Y}$ is also an equilibrium outcome from strategic interaction not only through $\boldsymbol{D}$ but also through itself. In this case, it would become important to have a vector of unobservables even after assuming e.g., rank invariance, since we may want to include $\boldsymbol{\epsilon}_{\boldsymbol{D}} = (\epsilon_{1,\boldsymbol{D}}, ..., \epsilon_{S,\boldsymbol{D}})$, where $\epsilon_{s,\boldsymbol{D}}$ is an unobservable directly affecting Y_s.[29] We may also want to include a vector of observables of all players $\boldsymbol{X} = (X_1, ..., X_S)$, where X_s directly affects Y_s. Then, interaction among Y_s can be modeled via a reduced-form representation:

$$Y_s = \theta_s(\boldsymbol{D}, \boldsymbol{X}, \boldsymbol{\epsilon}_{\boldsymbol{D}}), \qquad s \in \{1, ..., S\}.$$

In the entry example, the first-stage scalar unobservable U_s may represent each firm's unobserved fixed cost (while Z_s captures observed fixed cost). The vector of unobservables in the player-specific

[29] In this case, Assumption M should be imposed on $\epsilon_{s,\boldsymbol{D}}$ for each s.

outcome equation represents multiple shocks, such as the player's demand and variable cost shocks, and other firms' variable cost and demand shocks. Unlike in a linear model, it would be hard to argue that these errors are all aggregated as a scalar variable in this nonlinear outcome model, since it is not known in which fashion they enter the equation.

D Proofs

In terms of notation, when no confusion arises, we sometimes change the order of entry and write $\boldsymbol{v} = (v_s, \boldsymbol{v}_{-s})$ for convenience. For a multivariate function $f(\boldsymbol{v})$, the integral $\int_A f(\boldsymbol{v})d\boldsymbol{v}$ is understood as a multi-dimensional integral over a set A contained in the space of $\boldsymbol{v}$. Vectors in this paper are row vectors. Also, we write $Y_{\boldsymbol{d}} \equiv Y(\boldsymbol{d})$ for simplicity in this section.

D.1 Formal Notation for Equilibrium Regions

We begin by introducing some notation for equilibrium profiles. For $k = 1, ..., S$, let $\boldsymbol{e}_k$ be an S-vector of all zeros except for the k-th element which is equal to one, and let $\boldsymbol{e}_0 \equiv (0, ..., 0)$. For $j = 0, ..., S$, define $\boldsymbol{e}^j \equiv \sum_{k=0}^{j} \boldsymbol{e}_k$, which is an S-vector where the first j elements are unity and the rest are zero. For a set of positive integers, define a permutation function $\sigma : \{n_1, ..., n_S\} \to \{n_1, ..., n_S\}$, which has to be a one-to-one function.[30] Let Σ be a set of all possible permutations. Define a set of all possible permutations of $\boldsymbol{e}^j = (e_1^j, ..., e_S^j)$ as

$$\mathcal{D}^j \equiv \left\{ \boldsymbol{d}^j : \boldsymbol{d}^j = (\sigma(e_1^j), ..., \sigma(e_S^j)) \text{ for } \sigma(\cdot) \in \Sigma \right\} \tag{D.1}$$

for $j = 0, ..., S$. Note $\mathcal{D}^j$ is constructed to be a set of all equilibrium profiles with j treatments selected or j entrants, and it partitions $\mathcal{D} = \bigcup_{j=0}^{S} \mathcal{D}^j$. There are $S!/j!(S-j)!$ distinct $\boldsymbol{d}^j$'s in $\mathcal{D}^j$. For example with $S = 3$, $\boldsymbol{d}^2 \in \mathcal{D}^2 = \{(1,1,0), (1,0,1), (0,1,1)\}$ and $\boldsymbol{d}^0 \in \mathcal{D}^0 = \{(0,0,0)\}$. Note $\boldsymbol{d}^0 = \boldsymbol{e}^0 = (0, ..., 0)$ and $\boldsymbol{d}^S = \boldsymbol{e}^S = (1, ..., 1)$.

Let $\boldsymbol{D}(\boldsymbol{z}) \equiv (D_1(\boldsymbol{z}), ..., D_S(\boldsymbol{z}))$ where $\boldsymbol{z} \equiv (z_1, ..., z_S)$ and $D_s(\boldsymbol{z})$ is the potential treatment decision had the player s been assigned $\boldsymbol{Z} = \boldsymbol{z}$. We are interested in characterizing a region R of

[30] For example,

$$\begin{pmatrix} n_1 & n_2 & n_3 & n_4 & n_5 \\ \sigma(n_1) & \sigma(n_2) & \sigma(n_3) & \sigma(n_4) & \sigma(n_5) \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 5 & 3 & 4 \end{pmatrix}.$$

$\boldsymbol{U} \equiv (U_1, ..., U_S)$ in $\mathcal{U} \equiv (0,1]^S$ that satisfies $\boldsymbol{U} \in R \Leftrightarrow \boldsymbol{D}(\boldsymbol{z}) \in \mathcal{D}^j$ for some j. For each equilibrium profile, we define regions of $\boldsymbol{U}$ that are Cartesian products in $\mathcal{U}$. With a slight abuse of notation, let $\boldsymbol{d}^j_{-s} \equiv (\sigma(e^j_1), ..., \sigma(e^j_{s-1}), \sigma(e^j_{s+1}), ..., \sigma(e^j_{S-1}))$ for $0 \leq j \leq S-1$:

$$R_{\boldsymbol{d}^0}(\boldsymbol{z}) \equiv \prod_{s=1}^{S} \left(\nu^s(\boldsymbol{d}^0_{-s}, z_s), 1\right], \qquad R_{\boldsymbol{d}^S}(\boldsymbol{z}) \equiv \prod_{s=1}^{S} \left(0, \nu^s(\boldsymbol{d}^{S-1}_{-s}, z_s)\right]$$

and, given $\boldsymbol{d}^j = (\sigma(e^j_1), ..., \sigma(e^j_S))$ for some $\sigma(\cdot) \in \Sigma$[31] and $j = 1, ..., S-1$,

$$R_{\boldsymbol{d}^j}(\boldsymbol{z}) \equiv \left\{ \boldsymbol{U} : (U_{\sigma(1)}, ..., U_{\sigma(S)}) \in \left\{ \prod_{s=1}^{j} \left(0, \nu^{\sigma(s)}(\boldsymbol{d}^{j-1}_{-\sigma(s)}, z_{\sigma(s)})\right] \right\} \times \left\{ \prod_{s=j+1}^{S} \left(\nu^{\sigma(s)}(\boldsymbol{d}^j_{-\sigma(s)}, z_{\sigma(s)}), 1\right] \right\} \right\} \tag{D.2}$$

For example, for $\sigma(\cdot)$ such that $\boldsymbol{d}^1 = (\sigma(1), \sigma(0), \sigma(0)) = (0, 1, 0)$,

$$R_{010}(\boldsymbol{z}) = \left(\nu^1(1,0,z_1), 1\right] \times \left(0, \nu^2(0,0,z_2)\right] \times \left(\nu^3(0,1,z_3), 1\right].$$

Define the region of all equilibria with j treatments selected or j entrants as

$$\boldsymbol{R}_j(\boldsymbol{z}) \equiv \bigcup_{\boldsymbol{d}^j \in \mathcal{D}^j} R_{\boldsymbol{d}^j}(\boldsymbol{z}). \tag{D.3}$$

D.2 Proof of Theorem 3.1

We prove the theorem by showing the following lemma:

Lemma D.1. *Under Assumptions SS, for $j = 0, ..., S-1$, $\boldsymbol{R}^{\leq j}(\boldsymbol{z})$ is expressed as a union across $\sigma(\cdot) \in \Sigma$ of Cartesian products, each of which is a product of intervals that are either $(0,1]$ or $\left(\nu^{\sigma(s)}(d^j_{-\sigma(s)}, z_{\sigma(s)}), 1\right]$ for some $s = 1, ...S$.*

Given this lemma, (3.5) holds by Assumption M1, because for given s, $\left(\nu^s(d^j_{-s}, z_s), 1\right] \subseteq \left(\nu^s(d^j_{-s}, z'_s), 1\right]$ for any d^j_{-s} where the direction of inclusion is given by (3.4). Now we prove Lemma D.1.

Consider $d^j_s = 1$ for an s-th element d^j_s in $\boldsymbol{d}^j$ ($j \geq 1$). Then there exists $\boldsymbol{d}^{j-1}$ such that $d^{j-1}_s = 0$.

[31] Sometime we use the notation $\boldsymbol{d}^j_\sigma$ to emphasize the permutation function $\sigma(\cdot)$ from which $\boldsymbol{d}^j$ is generated.

Suppose not. Then $d_s^{j-1} = 1\ \forall \boldsymbol{d}^{j-1}$, and thus we can construct $\boldsymbol{d}^{j-1}$ that is equal to $\boldsymbol{d}^j$, which is contradiction. Therefore, in calculating $\boldsymbol{R}_j(\boldsymbol{z}) \cup \boldsymbol{R}_{j-1}(\boldsymbol{z})$, according to (D.2), what is involved is the union of intervals associated with $d_s^j = 1$ and $d_s^{j-1} = 0$, while sharing the same opponent d_{-s}^{j-1}: $\left(0, \nu^s(d_{-s}^{j-1}, z_s)\right] \cup \left(\nu^s(d_{-s}^{j-1}, z_s), 1\right] = (0,1]$. This implies that $\boldsymbol{R}_j(\boldsymbol{z}) \cup \boldsymbol{R}_{j-1}(\boldsymbol{z})$ is not a function of $\boldsymbol{z}$ through $\nu^s(d_{-s}^{j-1}, z_s)$ for any, and nor is $\boldsymbol{R}^{\leq j}(\boldsymbol{z}) \equiv \bigcup_{k=0}^{j} \boldsymbol{R}_k(\boldsymbol{z})$. On the other hand, when $d_s^j = 0$ for given s, the associated interval is $\left(\nu^s(d_{-s}^j, z_s), 1\right]$ as shown in (D.2). Therefore, $\boldsymbol{R}^{\leq j}(\boldsymbol{z}) \equiv \bigcup_{k=0}^{j} \boldsymbol{R}_k(\boldsymbol{z})$ is a function of $\boldsymbol{z}$ only through $\nu^s(d_{-s}^j, z_s)$ for some s. This proves Lemma D.1.

D.3 Proof of Lemma 3.1

The following proposition is useful later:

Proposition D.1. *Let R and Q be sets defined by Cartesian products: $R = \prod_{s=1}^{S} r_s$ and $Q = \prod_{s=1}^{S} q_s$ where r_s and q_s are intervals in $\mathbb{R}$. Then $R \cap Q = \prod_{s=1}^{S} r_s \cap q_s$.*

The proof of this proposition follows directly from the definition of R and Q.

The first part proves that Assumption EQ is equivalent to $R_{\boldsymbol{d}^j}(\boldsymbol{z}) \cap R_{\tilde{\boldsymbol{d}}^j}(\boldsymbol{z}') = \emptyset$ for all $\boldsymbol{d}^j \neq \tilde{\boldsymbol{d}}^j$ and j. For any $\boldsymbol{d}^j$ and $\tilde{\boldsymbol{d}}^j$ ($\boldsymbol{d}^j \neq \tilde{\boldsymbol{d}}^j$), the expression of $R_{\boldsymbol{d}^j}(\boldsymbol{z}) \cap R_{\tilde{\boldsymbol{d}}^j}(\boldsymbol{z}')$ can be inferred as follows. Under Assumption M1, we can simplify the notation of the payoff function as $\nu_j^s(z_s) \equiv \nu^s(\boldsymbol{d}_{-s}^j, z_s)$ when we compare it for different values of z_s. First, there exists s^* such that $d_{s^*}^j = 1$ and $\tilde{d}_{s^*}^j = 0$ (without loss of generality), otherwise it contradicts $\boldsymbol{d}^j \neq \tilde{\boldsymbol{d}}^j$. That is, $U_{s^*} \in \left(0, \nu_{j-1}^{s^*}(z_{s^*})\right]$ in $R_{\boldsymbol{d}^j}(\boldsymbol{z})$ and $U_{s^*} \in \left(\nu_j^{s^*}(z'_{s^*}), 1\right]$ in $R_{\tilde{\boldsymbol{d}}^j}(\boldsymbol{z}')$. For other $s \neq s^*$, the pair is realized to be one of the four types: (i) $d_s^j = 1$ and $\tilde{d}_s^j = 0$; (ii) $d_s^j = 0$ and $\tilde{d}_s^j = 1$; (iii) $d_s^j = 1$ and $\tilde{d}_s^j = 1$; (iv) $d_s^j = 0$ and $\tilde{d}_s^j = 0$. Then the corresponding pair of intervals for $R_{\boldsymbol{d}^j}(\boldsymbol{z})$ and $R_{\tilde{\boldsymbol{d}}^j}(\boldsymbol{z}')$, respectively, falls into one of the four types: (i) $\left(0, \nu_{j-1}^s(z_s)\right]$ and $\left(\nu_j^s(z'_s), 1\right]$; (ii) $\left(\nu_j^s(z_s), 1\right]$ and $\left(0, \nu_{j-1}^s(z'_s)\right]$; (iii) $\left(0, \nu_{j-1}^s(z_s)\right]$ and $\left(0, \nu_{j-1}^s(z'_s)\right]$; (iv) $\left(\nu_j^s(z_s), 1\right]$ and $\left(\nu_j^s(z'_s), 1\right]$. Then by Proposition D.1, $R_{\boldsymbol{d}^j}(\boldsymbol{z}) \cap R_{\tilde{\boldsymbol{d}}^j}(\boldsymbol{z}')$ is a product of the intersections of the interval pairs. But the intersection resulting from $\left(0, \nu_{j-1}^{s^*}(z_{s^*})\right]$ and $\left(\nu_j^{s^*}(z'_{s^*}), 1\right]$ is empty if and only if $\nu_{j-1}^{s^*}(z_{s^*}) \leq \nu_j^{s^*}(z'_{s^*})$. Therefore, $R_{\boldsymbol{d}^j}(\boldsymbol{z}) \cap R_{\tilde{\boldsymbol{d}}^j}(\boldsymbol{z}') = \emptyset$ for all $\boldsymbol{d}^j$ and $\tilde{\boldsymbol{d}}^j$ ($\boldsymbol{d}^j \neq \tilde{\boldsymbol{d}}^j$) if and only if $(\boldsymbol{z}, \boldsymbol{z}')$ are such that $\nu_{j-1}^s(z_s) \leq \nu_j^s(z'_s)$ for all s. Additionally, note that $R_{\boldsymbol{d}^j}(\boldsymbol{z}) \cap R_{\tilde{\boldsymbol{d}}^j}(\boldsymbol{z}') = \emptyset$ implies

$$R^*_{\boldsymbol{d}^j}(\boldsymbol{z}) \cap R^*_{\tilde{\boldsymbol{d}}^j}(\boldsymbol{z}') = R^*_{\boldsymbol{d}^j}(\boldsymbol{z}') \cap R^*_{\tilde{\boldsymbol{d}}^j}(\boldsymbol{z}) = \emptyset \tag{D.4}$$

for $\boldsymbol{d}^j \neq \tilde{\boldsymbol{d}}^j$, where $R^*_{\boldsymbol{d}}(\boldsymbol{z})$ is the region that predicts equilibrium $\boldsymbol{d}$.[32] This last display is useful later in other proofs later.

Moreover, note that any region $R^M_j(\boldsymbol{z})$ of multiple equilibria for $\mathcal{D}_j$ given $\boldsymbol{z}$ is defined by the intersection of the following interval pairs (and no more): (i) $\left(0, \nu^s_{j-1}(z_s)\right]$ and $\left(\nu^s_j(z_s), 1\right]$; (ii) $\left(0, \nu^s_{j-1}(z_s)\right]$ and $\left(0, \nu^s_{j-1}(z_s)\right]$; (iii) $\left(\nu^s_j(z_s), 1\right]$ and $\left(\nu^s_j(z_s), 1\right]$. Therefore, by Assumption SS (i.e., $\nu^s_{j-1}(z_s) > \nu^s_j(z_s)$), such a region is defined by the following corresponding intersections: (i) $\left(\nu^s_j(z_s), \nu^s_{j-1}(z_s)\right]$; (ii) $\left(0, \nu^s_{j-1}(z_s)\right]$; (iii) $\left(\nu^s_j(z_s), 1\right]$. Therefore $R^M_j(\boldsymbol{z}) \cap R^M_j(\boldsymbol{z}') = \emptyset$ if and only if $R_{\boldsymbol{d}^j}(\boldsymbol{z}) \cap R_{\tilde{\boldsymbol{d}}^j}(\boldsymbol{z}') = \emptyset$ for $\boldsymbol{d}^j \neq \tilde{\boldsymbol{d}}^j$.

We now prove that, when (3.6) holds, it satisfies $R^M_j(\boldsymbol{z}) \cap R^M_j(\boldsymbol{z}') = \emptyset$ for all j. We first prove the claim for $S = 2$ and then generalize it. The probabilities in (3.6) equal

$$\Pr[\boldsymbol{D} = (1,1)|\boldsymbol{Z} = \boldsymbol{z}] = \Pr[\boldsymbol{U} \in R_{11}(\boldsymbol{z})],$$
$$\Pr[\boldsymbol{D} = (0,0)|\boldsymbol{Z} = \boldsymbol{z}'] = \Pr[\boldsymbol{U} \in R_{00}(\boldsymbol{z}')].$$

Under independent unobserved types, these probabilities are equivalent to the volume of $R_{11}(\boldsymbol{z})$ and $R_{00}(\boldsymbol{z}')$, respectively. We consider two isoquant curves that are subsets of the surface of circles in $\mathcal{U}$: a curve $C_{11}(\boldsymbol{z})$ that is strictly convex from its origin $(0,0)$ and delivers the same volume as $R_{11}(\boldsymbol{z})$ and a curve $C_{00}(\boldsymbol{z}')$ that is strictly convex from its origin $(1,1)$ for $R_{00}(\boldsymbol{z}')$. Note that any region of multiple equilibria lies between the curve and its *opposite* origin. That is, $R^M_1(\boldsymbol{z})$ lies between $C_{11}(\boldsymbol{z})$ and $(1,1)$, and $R^M_1(\boldsymbol{z}')$ lies between $C_{00}(\boldsymbol{z}')$ and $(0,0)$. Therefore, if $C_{11}(\boldsymbol{z}) \cap C_{00}(\boldsymbol{z}') = \emptyset$ then $R^M_1(\boldsymbol{z}) \cap R^M_1(\boldsymbol{z}') = \emptyset$, because the curves are strictly convex.

The remaining argument is to prove that $C_{11}(\boldsymbol{z}) \cap C_{00}(\boldsymbol{z}') = \emptyset$. In order for this to be true, the sum of the radii of $C_{11}(\boldsymbol{z})$ and $C_{00}(\boldsymbol{z}')$ should not be great than $\sqrt{2}$, the length of the *space diagonal* of $\mathcal{U} = (0,1]^2$. But note that the radius can be identified from the data by considering an extreme scenario along each isoquant curve. First, consider the situation that player 1 is unprofitable to enter irrespective of player 2's decisions with $\boldsymbol{z}$. Then $\mathcal{U} = \tilde{R}_{11}(\boldsymbol{z}) \cup \tilde{R}_{10}(\boldsymbol{z})$ and it is easy to see that $1 - \Pr[\boldsymbol{U} \in \tilde{R}_{11}(\boldsymbol{z})]$ the radius of $C_{11}(\boldsymbol{z})$. Second, consider a situation that player 1 is profitable to enter irrespective of player 2's decisions with $\boldsymbol{z}'$. Then $\mathcal{U} = \tilde{R}_{00}(\boldsymbol{z}') \cup \tilde{R}_{01}(\boldsymbol{z}')$ and $1 - \Pr[\boldsymbol{U} \in \tilde{R}_{00}(\boldsymbol{z}')]$

[32] Note that $R^*_{\boldsymbol{d}}(\boldsymbol{z})$ is unknown to the econometrician even if all the players' payoffs had been known, since the equilibrium selection rule is unknown. This is in contrast to $R_{\boldsymbol{d}}(\boldsymbol{z})$ defined in Section D.1, which is purely determined by the payoffs $\nu^s_{\boldsymbol{d}_{-s}}(z_s)$,

is the radius of $C_{00}(\boldsymbol{z}')$. Therefore, $C_{11}(\boldsymbol{z}) \cap C_{00}(\boldsymbol{z}') = \emptyset$ is implied by

$$\begin{aligned}\sqrt{2} &> (1 - \Pr[\boldsymbol{U} \in \tilde{R}_{11}(\boldsymbol{z})]) + (1 - \Pr[\boldsymbol{U} \in \tilde{R}_{00}(\boldsymbol{z}')]) \\ &= (1 - \Pr[\boldsymbol{U} \in R_{11}(\boldsymbol{z})]) + (1 - \Pr[\boldsymbol{U} \in R_{00}(\boldsymbol{z}')]),\end{aligned}$$

where the equality is by the definition of the isoquant curves.

To prove the general case for $S \geq 2$, we iteratively apply the result from the previous case of one less player, starting from $S = 2$. Suppose $S = 3$. Consider $R_{111}(\boldsymbol{z})$ and $R_{001}(\boldsymbol{z}')$. By definition, these regions are analogous to the regions in the $S = 2$ case above on the surface $\{(U_1, U_2, 0)\} \subset \mathcal{U} = (0, 1]^3$. Similarly, the following is the pairs of regions and corresponding surfaces that are analogous to $S = 2$: $R_{110}(\boldsymbol{z})$ and $R_{000}(\boldsymbol{z}')$ on $\{(U_1, U_2, 1)\}$, $R_{111}(\boldsymbol{z})$ and $R_{010}(\boldsymbol{z}')$ on $\{(U_1, 0, U_3)\}$, $R_{101}(\boldsymbol{z})$ and $R_{000}(\boldsymbol{z}')$ on $\{(U_1, 1, U_3)\}$, $R_{111}(\boldsymbol{z})$ and $R_{100}(\boldsymbol{z}')$ on $\{(0, U_2, U_3)\}$, $R_{011}(\boldsymbol{z})$ and $R_{000}(\boldsymbol{z}')$ on $\{(1, U_2, U_3)\}$. But note that any region of multiple equilibria can be partitioned and projected on the regions of multiple equilibria on these surface; see Figures 9 and 10. Therefore, $R^M_j(\boldsymbol{z}) \cap R^M_j(\boldsymbol{z}') = \emptyset$ for all j if

$$\begin{aligned}\sqrt{2} &> (1 - \Pr[\boldsymbol{U} \in R_{\boldsymbol{d}^j}(\boldsymbol{z})]) + (1 - \Pr[\boldsymbol{U} \in R_{\boldsymbol{d}^{j-2}}(\boldsymbol{z}')]) \\ &= (1 - \Pr[\boldsymbol{D} = \boldsymbol{d}^j|\boldsymbol{z}]) + (1 - \Pr[\boldsymbol{D} = \boldsymbol{d}^{j-2}|\boldsymbol{z}'])\end{aligned} \tag{D.5}$$

for all $\boldsymbol{d}^j$ and $\boldsymbol{d}^{j-2}$ and $j \in \{2, 3\}$. Next, for $S = 4$, focusing on the surfaces of the hypercube $\mathcal{U} = (0, 1]^4$, we can apply the result from $S = 3$, and so on. Therefore, in general, $R^M_j(\boldsymbol{z}) \cap R^M_j(\boldsymbol{z}') = \emptyset$ for all j if (D.5) for any $\boldsymbol{d}^j \in \mathcal{D}^j$, $\boldsymbol{d}^{j-2} \in \mathcal{D}^{j-2}$ and $2 \leq j \leq S$.

D.4 Proof of Result (2.13)

Introduce

$$\begin{aligned}h_{11}(\boldsymbol{z}, \boldsymbol{z}') &\equiv \Pr[Y = 1, \boldsymbol{D} = (1, 1)|\boldsymbol{Z} = \boldsymbol{z}] - \Pr[Y = 1, \boldsymbol{D} = (1, 1)|\boldsymbol{Z} = \boldsymbol{z}'], \\ h_{00}(\boldsymbol{z}, \boldsymbol{z}') &\equiv \Pr[Y = 1, \boldsymbol{D} = (0, 0)|\boldsymbol{Z} = \boldsymbol{z}] - \Pr[Y = 1, \boldsymbol{D} = (0, 0)|\boldsymbol{Z} = \boldsymbol{z}'], \\ h_{10}(\boldsymbol{z}, \boldsymbol{z}') &\equiv \Pr[Y = 1, \boldsymbol{D} = (1, 0)|\boldsymbol{Z} = \boldsymbol{z}] - \Pr[Y = 1, \boldsymbol{D} = (1, 0)|\boldsymbol{Z} = \boldsymbol{z}'], \\ h_{01}(\boldsymbol{z}, \boldsymbol{z}') &\equiv \Pr[Y = 1, \boldsymbol{D} = (0, 1)|\boldsymbol{Z} = \boldsymbol{z}] - \Pr[Y = 1, \boldsymbol{D} = (0, 1)|\boldsymbol{Z} = \boldsymbol{z}'].\end{aligned}$$

Then h defined in (2.9) satisfies $h = h_{11} + h_{00} + h_{10} + h_{01}$. Let R^*_{10} and R^*_{01} be the regions that predict $\boldsymbol{D} = (1, 0)$ and $\boldsymbol{D} = (0, 1)$, respectively, which is unknown since the equilibrium selection mechanism is unknown. Suppose $(\boldsymbol{z}, \boldsymbol{z}')$ are such that EQ (or equivalently (**??**)) holds. Also, suppose $(\boldsymbol{z}, \boldsymbol{z}')$ are such that (3.4) holds, then we have $R_{11}(\boldsymbol{z}) \supset R_{11}(\boldsymbol{z}')$ and $R_{00}(\boldsymbol{z}) \subset R_{00}(\boldsymbol{z}')$, respectively, by Theorem 3.1. Define

$$\Delta(\boldsymbol{z}, \boldsymbol{z}') \equiv \{R^*_{10}(\boldsymbol{z}) \cup R^*_{01}(\boldsymbol{z})\} \setminus \boldsymbol{R}_1(\boldsymbol{z}'), \tag{D.6}$$

$$-\Delta(\boldsymbol{z}, \boldsymbol{z}') \equiv \left\{R^*_{10}(\boldsymbol{z}') \cup R^*_{01}(\boldsymbol{z}')\right\} \setminus \boldsymbol{R}_1(\boldsymbol{z}). \tag{D.7}$$

Consider partitions $\Delta(\boldsymbol{z}, \boldsymbol{z}') = \Delta^1(\boldsymbol{z}, \boldsymbol{z}') \cup \Delta^2(\boldsymbol{z}, \boldsymbol{z}')$ and $-\Delta(\boldsymbol{z}, \boldsymbol{z}') = -\Delta^1(\boldsymbol{z}, \boldsymbol{z}') \cup -\Delta^2(\boldsymbol{z}, \boldsymbol{z}')$ such that

$$\Delta^1(\boldsymbol{z}, \boldsymbol{z}') \equiv R^*_{10}(\boldsymbol{z}) \setminus \boldsymbol{R}_1(\boldsymbol{z}'), \quad \Delta^2(\boldsymbol{z}, \boldsymbol{z}') \equiv R^*_{01}(\boldsymbol{z}) \setminus \boldsymbol{R}_1(\boldsymbol{z}'),$$
$$-\Delta^1(\boldsymbol{z}, \boldsymbol{z}') \equiv R^*_{10}(\boldsymbol{z}') \setminus \boldsymbol{R}_1(\boldsymbol{z}), \quad -\Delta^2(\boldsymbol{z}, \boldsymbol{z}') \equiv R^*_{01}(\boldsymbol{z}') \setminus \boldsymbol{R}_1(\boldsymbol{z}).$$

That is, $\Delta^1(\boldsymbol{z}, \boldsymbol{z}')$ and $-\Delta^1(\boldsymbol{z}, \boldsymbol{z}')$ are regions of R^*_{10} exchanged with the regions for $\boldsymbol{D} = (0, 0)$ and $\boldsymbol{D} = (1, 1)$, respectively, and $+\Delta^2(\boldsymbol{z}, \boldsymbol{z}')$ and $-\Delta^2(\boldsymbol{z}, \boldsymbol{z}')$ are for R^*_{01}.

Before proceeding, we introduce the following general rule that is useful later: for a uniform random vector $\tilde{\boldsymbol{U}}$ and two sets B and B' contained in $\tilde{\mathcal{U}}$ and for a r.v. ϵ and set $A \subset \mathcal{E}$,

$$\Pr[\epsilon \in A, \tilde{\boldsymbol{U}} \in B] - \Pr[\epsilon \in A, \tilde{\boldsymbol{U}} \in B'] = \Pr[\epsilon \in A, \tilde{\boldsymbol{U}} \in B \setminus B'] - \Pr[\epsilon \in A, \tilde{\boldsymbol{U}} \in B' \setminus B]. \tag{D.8}$$

Since we do not use the variation in X, we suppress it throughout. Let $\mu_{\boldsymbol{d}} \equiv \mu_0 + \mu_1 d_1 + \mu_2 d_2$ for simplicity. Now, by Assumption IN,

$$\begin{aligned} h_{10}(\boldsymbol{z}, \boldsymbol{z}') &= \Pr[\epsilon \leq \mu_{10}, \boldsymbol{U} \in R^*_{10}(\boldsymbol{z})] - \Pr[\epsilon \leq \mu_{10}, \boldsymbol{U} \in R^*_{10}(\boldsymbol{z}')] \\ &= \Pr[\epsilon \leq \mu_{10}, \boldsymbol{U} \in R^*_{10}(\boldsymbol{z}) \setminus R^*_{10}(\boldsymbol{z}')] - \Pr[\epsilon \leq \mu_{10}, \boldsymbol{U} \in R^*_{10}(\boldsymbol{z}') \setminus R^*_{10}(\boldsymbol{z})] \\ &= \Pr[\epsilon \leq \mu_{10}, \boldsymbol{U} \in \Delta^1(\boldsymbol{z}, \boldsymbol{z}')] - \Pr[\epsilon \leq \mu_{10}, \boldsymbol{U} \in -\Delta^1(\boldsymbol{z}, \boldsymbol{z}')] \end{aligned}$$

where the second equality is by (D.8) and the third equality is by the following derivation:

$$\begin{aligned}R^*_{10}(\boldsymbol{z})\backslash R^*_{10}(\boldsymbol{z}') &= \left[\left\{R^*_{10}(\boldsymbol{z}) \cap \boldsymbol{R}_1(\boldsymbol{z}')^c\right\} \backslash R^*_{10}(\boldsymbol{z}')\right] \cup \left[\left\{R^*_{10}(\boldsymbol{z}) \cap \boldsymbol{R}_1(\boldsymbol{z}')\right\} \backslash R^*_{10}(\boldsymbol{z}')\right] \\ &= \left[\left\{R^*_{10}(\boldsymbol{z}) \cap \boldsymbol{R}_1(\boldsymbol{z}')^c\right\}\right] \cup \left[\left\{R^*_{10}(\boldsymbol{z}') \cap \boldsymbol{R}_1(\boldsymbol{z})\right\} \backslash R^*_{10}(\boldsymbol{z}')\right] \\ &= \Delta^1(\boldsymbol{z}, \boldsymbol{z}'),\end{aligned}$$

where the first equality is by the distributive law and $\mathcal{U} = \boldsymbol{R}_1(\boldsymbol{z}')^c \cup \boldsymbol{R}_1(\boldsymbol{z}')$, the second equality is by $\boldsymbol{R}_1(\boldsymbol{z}')^c = R^*_{10}(\boldsymbol{z}')^c \cap R^*_{01}(\boldsymbol{z}')^c$ (the first term) and by Assumption EQ (the second term), and the last equality is by the definition of $\Delta^1(\boldsymbol{z}, \boldsymbol{z}')$ and $\{R^*_{10}(\boldsymbol{z}') \cap \boldsymbol{R}_1(\boldsymbol{z})\} \backslash R^*_{10}(\boldsymbol{z}')$ being empty. Analogously, one can show that $R^*_{10}(\boldsymbol{z}')\backslash R^*_{10}(\boldsymbol{z}) = -\Delta^1(\boldsymbol{z}, \boldsymbol{z}')$ using Assumption EQ and the definition of $-\Delta^1(\boldsymbol{z}, \boldsymbol{z}')$. Similarly,

$$\begin{aligned}h_{01}(\boldsymbol{z}, \boldsymbol{z}') =& \Pr[\epsilon \leq \mu_{01}, \boldsymbol{U} \in R^*_{01}(\boldsymbol{z})] - \Pr[\epsilon \leq \mu_{01}, \boldsymbol{U} \in R^*_{01}(\boldsymbol{z}')] \\ =& \Pr[\epsilon \leq \mu_{01}, \boldsymbol{U} \in R^*_{01}(\boldsymbol{z})\backslash R^*_{01}(\boldsymbol{z}')] - \Pr[\epsilon \leq \mu_{01}, \boldsymbol{U} \in R^*_{01}(\boldsymbol{z}')\backslash R^*_{01}(\boldsymbol{z})] \\ =& \Pr[\epsilon \leq \mu_{01}, \boldsymbol{U} \in \Delta^2(\boldsymbol{z}, \boldsymbol{z}')] - \Pr[\epsilon \leq \mu_{01}, \boldsymbol{U} \in -\Delta^2(\boldsymbol{z}, \boldsymbol{z}')].\end{aligned}$$

Also, by the definitions of the partitions,

$$\begin{aligned}h_{11}(\boldsymbol{z}, \boldsymbol{z}') =& \Pr[\epsilon \leq \mu_{11}, \boldsymbol{U} \in -\Delta(\boldsymbol{z}, \boldsymbol{z}') \cup A^*] \\ =& \Pr[\epsilon \leq \mu_{11}, \boldsymbol{U} \in -\Delta^1(\boldsymbol{z}, \boldsymbol{z}')] + \Pr[\epsilon \leq \mu_{11}, \boldsymbol{U} \in -\Delta^2(\boldsymbol{z}, \boldsymbol{z}')] \\ &+ \Pr[\epsilon \leq \mu_{11}, \boldsymbol{U} \in A^*]\end{aligned}$$

since $-\Delta(\boldsymbol{z}, \boldsymbol{z}')$ and A^* are disjoint, and

$$\begin{aligned}h_{00}(\boldsymbol{z}, \boldsymbol{z}') =& -\Pr[\epsilon \leq \mu_{00}, \boldsymbol{U} \in \Delta(\boldsymbol{z}, \boldsymbol{z}') \cup A^*] \\ =& -\Pr[\epsilon \leq \mu_{00}, \boldsymbol{U} \in \Delta^1(\boldsymbol{z}, \boldsymbol{z}')] - \Pr[\epsilon \leq \mu_{00}, \boldsymbol{U} \in \Delta^2(\boldsymbol{z}, \boldsymbol{z}')] \\ &- \Pr[\epsilon \leq \mu_{00}, \boldsymbol{U} \in A^*]\end{aligned}$$

since $\Delta(\boldsymbol{z},\boldsymbol{z}')$ and A^* are disjoint. Now combining all the terms yields

$$\begin{aligned}
h(\boldsymbol{z},\boldsymbol{z}') =& \Pr[\epsilon \le \mu_{11}, \boldsymbol{U} \in -\Delta^1(\boldsymbol{z},\boldsymbol{z}')] - \Pr[\epsilon \le \mu_{10}, \boldsymbol{U} \in -\Delta^1(\boldsymbol{z},\boldsymbol{z}')] \\
&+ \Pr[\epsilon \le \mu_{11}, \boldsymbol{U} \in -\Delta^2(\boldsymbol{z},\boldsymbol{z}')] - \Pr[\epsilon \le \mu_{01}, \boldsymbol{U} \in -\Delta^2(\boldsymbol{z},\boldsymbol{z}')] \\
&+ \Pr[\epsilon \le \mu_{10}, \boldsymbol{U} \in \Delta^1(\boldsymbol{z},\boldsymbol{z}')] - \Pr[\epsilon \le \mu_{00}, \boldsymbol{U} \in \Delta^1(\boldsymbol{z},\boldsymbol{z}')] \\
&+ \Pr[\epsilon \le \mu_{01}, \boldsymbol{U} \in \Delta^2(\boldsymbol{z},\boldsymbol{z}')] - \Pr[\epsilon \le \mu_{00}, \boldsymbol{U} \in \Delta^2(\boldsymbol{z},\boldsymbol{z}')]. \\
&+ \Pr[\epsilon \le \mu_{11}, \boldsymbol{U} \in A^*] - \Pr[\epsilon \le \mu_{00}, \boldsymbol{U} \in A^*]
\end{aligned}$$

In this expression, each set of $\boldsymbol{U}$ has a corresponding set in the expression (2.12) of the main text: $-\Delta^1(\boldsymbol{z},\boldsymbol{z}') = \Delta_a$, $-\Delta^2(\boldsymbol{z},\boldsymbol{z}') = \Delta_b$, $\Delta^1(\boldsymbol{z},\boldsymbol{z}') = \Delta_c$, $\Delta^2(\boldsymbol{z},\boldsymbol{z}') = \Delta_d$, and $A^* = \Delta_e$. Then, as already argued in the text, $\mu_{1,\boldsymbol{d}_{-s}} - \mu_{0,\boldsymbol{d}_{-s}}$ share the same signs for all s and $\forall \boldsymbol{d}_{-s} \in \{0,1\}$ and therefore $sgn\{h(\boldsymbol{z},\boldsymbol{z}')\} = sgn\left\{\mu_{1,\boldsymbol{d}_{-s}} - \mu_{0,\boldsymbol{d}_{-s}}\right\}$.

D.5 Proof of Theorem 3.2

To reduce the notation, we suppress the conditioning of $X = x$ throughout the proof. For a set $\tilde{\mathcal{D}} \subset \mathcal{D}$, let $\tilde{p}_{\tilde{\mathcal{D}}}(\boldsymbol{z}) \equiv \Pr[Y = 1, \boldsymbol{D} \in \tilde{\mathcal{D}} | \boldsymbol{Z} = \boldsymbol{z}]$ and $p_{\tilde{\mathcal{D}}}(\boldsymbol{z}) \equiv \Pr[\boldsymbol{D} \in \tilde{\mathcal{D}} | \boldsymbol{Z} = \boldsymbol{z}]$. Then the bounds (3.10) and (3.11) can be rewritten as

$$U_{\boldsymbol{d}^j} = \inf_{\boldsymbol{z}\in\mathcal{Z}} \left\{\tilde{p}_{\mathcal{D}^{\ge}(\boldsymbol{d}^j)}(\boldsymbol{z}) + p_{\mathcal{D}\setminus\mathcal{D}^{\ge}(\boldsymbol{d}^j)}(\boldsymbol{z})\right\}, \qquad L_{\boldsymbol{d}^j} = \sup_{\boldsymbol{z}\in\mathcal{Z}} \tilde{p}_{\mathcal{D}^{\le}(\boldsymbol{d}^j)}(\boldsymbol{z}).$$

Note that $\tilde{p}_{\mathcal{D}^{\ge}(\boldsymbol{d}^j)}(\boldsymbol{z}) = \Pr[Y = 1|\boldsymbol{Z} = \boldsymbol{z}] - \tilde{p}_{\mathcal{D}\setminus\mathcal{D}^{\ge}(\boldsymbol{d}^j)}(\boldsymbol{z})$. Suppose $\boldsymbol{z},\boldsymbol{z}'$ are chosen such that $p_{\boldsymbol{d}}(\boldsymbol{z}) - p_{\boldsymbol{d}}(\boldsymbol{z}') = \Pr[\boldsymbol{U} \in \Delta_{\boldsymbol{d}}(\boldsymbol{z},\boldsymbol{z}')] - \Pr[\boldsymbol{U} \in -\Delta_{\boldsymbol{d}}(\boldsymbol{z},\boldsymbol{z}')] > 0 \ \forall \mathcal{D}^{\ge}(\boldsymbol{d}^j)$, where $\Delta_{\boldsymbol{d}}$ and $-\Delta_{\boldsymbol{d}}$ are defined in (D.20) and (D.21) below. Observe that each term in $U_{\boldsymbol{d}^j}$ satisfies

$$\tilde{p}_{\mathcal{D}^{\ge}(\boldsymbol{d}^j)}(\boldsymbol{z}) - \tilde{p}_{\mathcal{D}^{\ge}(\boldsymbol{d}^j)}(\boldsymbol{z}') = \sum_{\boldsymbol{d}\in\mathcal{D}^{\ge}(\boldsymbol{d}^j)} \left(\Pr[\epsilon \le \mu_{\boldsymbol{d}}, \boldsymbol{U} \in \Delta_{\boldsymbol{d}}(\boldsymbol{z},\boldsymbol{z}')] - \Pr[\epsilon \le \mu_{\boldsymbol{d}}, \boldsymbol{U} \in -\Delta_{\boldsymbol{d}}(\boldsymbol{z},\boldsymbol{z}')]\right),$$

$$\begin{aligned}
p_{\mathcal{D}\setminus\mathcal{D}^{\ge}(\boldsymbol{d}^j)}(\boldsymbol{z}) - p_{\mathcal{D}\setminus\mathcal{D}^{\ge}(\boldsymbol{d}^j)}(\boldsymbol{z}') &= -(p_{\mathcal{D}^{\ge}(\boldsymbol{d}^j)}(\boldsymbol{z}) - p_{\mathcal{D}^{\ge}(\boldsymbol{d}^j)}(\boldsymbol{z}')) \\
&= -\left(\sum_{\boldsymbol{d}\in\mathcal{D}^{\ge}(\boldsymbol{d}^j)} \Pr[\boldsymbol{U} \in \Delta_{\boldsymbol{d}}(\boldsymbol{z},\boldsymbol{z}')] - \Pr[\boldsymbol{U} \in -\Delta_{\boldsymbol{d}}(\boldsymbol{z},\boldsymbol{z}')]\right),
\end{aligned}$$

and thus

$$\tilde{p}_{\mathcal{D}^{\geq}(\boldsymbol{d}^j)}(\boldsymbol{z}) + p_{\mathcal{D}\setminus\mathcal{D}^{\geq}(\boldsymbol{d}^j)}(\boldsymbol{z}) - \left\{\tilde{p}_{\mathcal{D}^{\geq}(\boldsymbol{d}^j)}(\boldsymbol{z}') + p_{\mathcal{D}\setminus\mathcal{D}^{\geq}(\boldsymbol{d}^j)}(\boldsymbol{z}')\right\}$$
$$= -\sum_{\boldsymbol{d}\in\mathcal{D}^{\geq}(\boldsymbol{d}^j)} \left(\Pr[\epsilon > \mu_{\boldsymbol{d}}, \boldsymbol{U} \in \Delta_{\boldsymbol{d}}(\boldsymbol{z},\boldsymbol{z}')] - \Pr[\epsilon > \mu_{\boldsymbol{d}}, \boldsymbol{U} \in -\Delta_{\boldsymbol{d}}(\boldsymbol{z},\boldsymbol{z}')]\right) < 0.$$

Then this relationship creates a *partial ordering* of $\tilde{p}_{\mathcal{D}^{\geq}(\boldsymbol{d}^j)}(\boldsymbol{z})+p_{\mathcal{D}\setminus\mathcal{D}^{\geq}(\boldsymbol{d}^j)}(\boldsymbol{z})$ as a function of $\boldsymbol{z}$. According to this ordering, $\tilde{p}_{\mathcal{D}^{\geq}(\boldsymbol{d}^j)}(\boldsymbol{z})+p_{\mathcal{D}\setminus\mathcal{D}^{\geq}(\boldsymbol{d}^j)}(\boldsymbol{z})$ takes its smallest value as $\max_{\boldsymbol{d}(\boldsymbol{z})\in\mathcal{D}^{\geq}(\boldsymbol{d}^j)} p_{\boldsymbol{d}(\boldsymbol{z})}(\boldsymbol{z})$ takes its largest value. Therefore, by (3.12),

$$U_{\boldsymbol{d}^j} = \inf_{\boldsymbol{z}\in\mathcal{Z}} \left\{\tilde{p}_{\mathcal{D}^{\geq}(\boldsymbol{d}^j)}(\boldsymbol{z}) + p_{\mathcal{D}\setminus\mathcal{D}^{\geq}(\boldsymbol{d}^j)}(\boldsymbol{z})\right\} = \tilde{p}_{\mathcal{D}^{\geq}(\boldsymbol{d}^j)}(\bar{\boldsymbol{z}}) + p_{\mathcal{D}\setminus\mathcal{D}^{\geq}(\boldsymbol{d}^j)}(\bar{\boldsymbol{z}}).$$

By a similar argument, $L_{\boldsymbol{d}^j} = \sup_{\boldsymbol{z}\in\mathcal{Z}} \tilde{p}_{\mathcal{D}^{\leq}(\boldsymbol{d}^j)}(\boldsymbol{z}) = \tilde{p}_{\mathcal{D}^{\leq}(\boldsymbol{d}^j)}(\underline{\boldsymbol{z}})$.

To prove that these bounds on $E[Y_{\boldsymbol{d}^j}]$ are sharp, it suffices to show that for $s_j \in [L_{\boldsymbol{d}^j}, U_{\boldsymbol{d}^j}]$, there exists a density function $f^*_{\epsilon,\boldsymbol{U}}$ such that the following claims hold:

(A) $f^*_{\epsilon|\boldsymbol{U}}$ is strictly positive on $\mathbb{R}$.

(B) The proposed model is consistent with the data: $\forall \boldsymbol{d}$,

$$\Pr[\boldsymbol{D} = \boldsymbol{d}|\boldsymbol{Z} = \boldsymbol{z}] = \Pr[\boldsymbol{U}^* \in \boldsymbol{R}_{\boldsymbol{d}}(\boldsymbol{z})],$$
$$\Pr[Y = 1|\boldsymbol{D} = \boldsymbol{d}, \boldsymbol{Z} = \boldsymbol{z}] = \Pr[\epsilon^* \leq \mu_{\boldsymbol{d}}|\boldsymbol{U}^* \in \boldsymbol{R}_{\boldsymbol{d}}(\boldsymbol{z})],$$

(C) The proposed model is consistent with the specified values of $E[Y_{\boldsymbol{d}^j}]$: $\Pr[\epsilon^* \leq \mu_{\boldsymbol{d}^j}] = s_j$.

An argument similar to the proof of Theorem 3.1 and the partial ordering above establish the monotonicity of the event $\boldsymbol{U} \in \bigcup_{\boldsymbol{d}\in\mathcal{D}^{\geq}(\boldsymbol{d}^j)} \boldsymbol{R}_{\boldsymbol{d}}(\boldsymbol{z})$ (and $\boldsymbol{U} \in \bigcup_{\boldsymbol{d}\in\mathcal{D}^{\leq}(\boldsymbol{d}^j)} \boldsymbol{R}_{\boldsymbol{d}}(\boldsymbol{z})$) w.r.t. $\boldsymbol{z}$. For example, for $\boldsymbol{z}, \boldsymbol{z}'$ chosen above, we have that $p_{\mathcal{D}^{\geq}(\boldsymbol{d}^j)}(\boldsymbol{z}) - p_{\mathcal{D}^{\geq}(\boldsymbol{d}^j)}(\boldsymbol{z}') > 0$, and thus $\bigcup_{\boldsymbol{d}\in\mathcal{D}^{\geq}(\boldsymbol{d}^j)} \boldsymbol{R}_{\boldsymbol{d}}(\boldsymbol{z}) \supset \bigcup_{\boldsymbol{d}\in\mathcal{D}^{\geq}(\boldsymbol{d}^j)} \boldsymbol{R}_{\boldsymbol{d}}(\boldsymbol{z}')$, which implies

$$1\left[\boldsymbol{U} \in \bigcup_{\boldsymbol{d}\in\mathcal{D}^{\geq}(\boldsymbol{d}^j)} \boldsymbol{R}_{\boldsymbol{d}}(\boldsymbol{z})\right] - 1\left[\boldsymbol{U} \in \bigcup_{\boldsymbol{d}\in\mathcal{D}^{\geq}(\boldsymbol{d}^j)} \boldsymbol{R}_{\boldsymbol{d}}(\boldsymbol{z}')\right] = 1\left[\boldsymbol{U} \in \bigcup_{\boldsymbol{d}\in\mathcal{D}^{\geq}(\boldsymbol{d}^j)} \boldsymbol{R}_{\boldsymbol{d}}(\boldsymbol{z})\setminus \bigcup_{\boldsymbol{d}\in\mathcal{D}^{\geq}(\boldsymbol{d}^j)} \boldsymbol{R}_{\boldsymbol{d}}(\boldsymbol{z})\right]. \tag{D.9}$$

Given $1[\boldsymbol{D} \in \mathcal{D}^{\geq}(\boldsymbol{d}^j)] = 1[\boldsymbol{U} \in \bigcup_{\boldsymbol{d}\in\mathcal{D}^{\geq}(\boldsymbol{d}^j)} \boldsymbol{R}_{\boldsymbol{d}}(\boldsymbol{Z})]$, (D.9) is analogous to a scalar treatment decision $\tilde{D} = 1[\tilde{D} = 1] = 1[\tilde{U} \leq \tilde{P}]$ with a scalar instrument $\tilde{P}$, where $1[\tilde{U} \leq p'] - 1[\tilde{U} \leq p] = 1[p \leq \tilde{U} \leq p']$

for $p' > p$. Based on this result and the results for the first part of Theorem 3.2, we can modify the proof of Theorem 2.1(iii) in Shaikh and Vytlacil (2011) to show (A)–(C).

D.6 Proof of Lemma 3.2

We introduce a lemma that establishes the connection between Theorem 3.1 and Lemma 3.2.

Lemma D.2. *Based on the results in Theorem 3.1, $\tilde{h}(\boldsymbol{z}, \boldsymbol{z}', \tilde{\boldsymbol{x}}) \equiv \sum_{j=0}^{S} h_j(\boldsymbol{z}, \boldsymbol{z}', x_j)$ satisfies*

$$\tilde{h}(\boldsymbol{z}, \boldsymbol{z}', \tilde{\boldsymbol{x}}) = \sum_{j=1}^{S} \sum_{(1,\boldsymbol{d}_{-s})\in\mathcal{D}^j} \int_{\Delta_{(1,\boldsymbol{d}_{-s}),(0,\boldsymbol{d}_{-s})}} \{\vartheta((1,\boldsymbol{d}_{-s}), x_j; \boldsymbol{u}) - \vartheta((0,\boldsymbol{d}_{-s}), x_{j-1}; \boldsymbol{u})\} d\boldsymbol{u}, \tag{D.10}$$

where $\Delta_{\boldsymbol{d},\tilde{\boldsymbol{d}}} = \Delta_{\boldsymbol{d},\tilde{\boldsymbol{d}}}(\boldsymbol{z}, \boldsymbol{z}')$ is a partition of $\Delta_{\boldsymbol{d}}(\boldsymbol{z}, \boldsymbol{z}')$ defined below.

As a special case of this lemma, $\tilde{h}(\boldsymbol{z}', \boldsymbol{z}, x, ..., x) = h(\boldsymbol{z}', \boldsymbol{z}, x)$ can be expressed as

$$h(\boldsymbol{z}', \boldsymbol{z}, x) = \sum_{\boldsymbol{d}_{-s}} \int_{\Delta_{(1,\boldsymbol{d}_{-s}),(0,\boldsymbol{d}_{-s})}} \{\vartheta((1,\boldsymbol{d}_{-s}), x; \boldsymbol{u}) - \vartheta((0,\boldsymbol{d}_{-s}), x; \boldsymbol{u})\} d\boldsymbol{u}. \tag{D.11}$$

We prove Lemma D.2 by drawing on the result of Theorem 3.1. We first establish the relationship between $(\boldsymbol{R}_j(\boldsymbol{z}), \boldsymbol{R}_j(\boldsymbol{z}'))$ and $(\boldsymbol{R}_{j-1}(\boldsymbol{z}), \boldsymbol{R}_{j-1}(\boldsymbol{z}'))$, and then establish refined results for individual equilibrium regions. By Theorem 3.1, for $\boldsymbol{z}$ and $\boldsymbol{z}'$ such that (3.4) holds, we have

$$\boldsymbol{R}^j(\boldsymbol{z}) \subseteq \boldsymbol{R}^j(\boldsymbol{z}') \tag{D.12}$$

for $j = 0, ..., S$, including $\boldsymbol{R}^S(\boldsymbol{z}) = \boldsymbol{R}^S(\boldsymbol{z}') = \mathcal{U}$ as a trivial case. For those $\boldsymbol{z}$ and $\boldsymbol{z}'$, introduce notation

$$\Delta_j(\boldsymbol{z}, \boldsymbol{z}') \equiv \boldsymbol{R}_j(\boldsymbol{z}) \backslash \boldsymbol{R}_j(\boldsymbol{z}'), \tag{D.13}$$

$$-\Delta_j(\boldsymbol{z}, \boldsymbol{z}') \equiv \boldsymbol{R}_j(\boldsymbol{z}') \backslash \boldsymbol{R}_j(\boldsymbol{z}), \tag{D.14}$$

and

$$\Delta^j(\boldsymbol{z}, \boldsymbol{z}') \equiv \boldsymbol{R}^j(\boldsymbol{z}) \backslash \boldsymbol{R}^j(\boldsymbol{z}'). \tag{D.15}$$

Note that, for $j = 1, ..., S$,

$$\boldsymbol{R}_j(\cdot) = \boldsymbol{R}^j(\cdot) \backslash \boldsymbol{R}^{j-1}(\cdot), \tag{D.16}$$

since $\boldsymbol{R}^j(\boldsymbol{z}) \equiv \bigcup_{k=0}^{j} \boldsymbol{R}_k(\boldsymbol{z})$. Fix $j = 1, ..., S$. Consider

$$\begin{aligned}
\Delta_j(\boldsymbol{z}, \boldsymbol{z}') &= \left(\boldsymbol{R}^j(\boldsymbol{z}) \cap \boldsymbol{R}^{j-1}(\boldsymbol{z})^c\right) \cap \left(\boldsymbol{R}^j(\boldsymbol{z}') \cap \boldsymbol{R}^{j-1}(\boldsymbol{z}')^c\right)^c \\
&= \left(\boldsymbol{R}^j(\boldsymbol{z}) \cap \boldsymbol{R}^{j-1}(\boldsymbol{z})^c\right) \cap \left(\boldsymbol{R}^j(\boldsymbol{z}')^c \cup \boldsymbol{R}^{j-1}(\boldsymbol{z}')\right) \\
&= \left(\boldsymbol{R}^j(\boldsymbol{z}) \cap \boldsymbol{R}^{j-1}(\boldsymbol{z})^c \cap \boldsymbol{R}^j(\boldsymbol{z}')^c\right) \cup \left(\boldsymbol{R}^j(\boldsymbol{z}) \cap \boldsymbol{R}^{j-1}(\boldsymbol{z})^c \cap \boldsymbol{R}^{j-1}(\boldsymbol{z}')\right) \\
&= \left\{\left(\boldsymbol{R}^j(\boldsymbol{z}) \backslash \boldsymbol{R}^j(\boldsymbol{z}')\right) \cap \boldsymbol{R}^{j-1}(\boldsymbol{z})^c\right\} \cup \left\{\left(\boldsymbol{R}^{j-1}(\boldsymbol{z}') \backslash \boldsymbol{R}^{j-1}(\boldsymbol{z})\right) \cap \boldsymbol{R}^j(\boldsymbol{z})\right\} \\
&= \Delta^{j-1}(\boldsymbol{z}', \boldsymbol{z}) \cap \boldsymbol{R}^j(\boldsymbol{z}),
\end{aligned}$$

where the first equality is by plugging in (D.16) into (D.13), the third equality is by the distributive law, and the last equality is by (D.12) and hence $\left(\boldsymbol{R}^j(\boldsymbol{z}) \backslash \boldsymbol{R}^j(\boldsymbol{z}')\right) \cap \boldsymbol{R}^{j-1}(\boldsymbol{z})^c = \emptyset$. But

$$\Delta^{j-1}(\boldsymbol{z}', \boldsymbol{z}) \backslash \boldsymbol{R}^j(\boldsymbol{z}) = \Delta^{j-1}(\boldsymbol{z}', \boldsymbol{z}) \backslash \left(\Delta^{j-1}(\boldsymbol{z}', \boldsymbol{z}) \cap \boldsymbol{R}^j(\boldsymbol{z})\right).$$

Symmetrically, by changing the role of $\boldsymbol{z}$ and $\boldsymbol{z}'$, consider

$$\begin{aligned}
-\Delta_j(\boldsymbol{z}, \boldsymbol{z}') &= \left(\boldsymbol{R}^j(\boldsymbol{z}') \cap \boldsymbol{R}^{j-1}(\boldsymbol{z}')^c\right) \cap \left(\boldsymbol{R}^j(\boldsymbol{z}) \cap \boldsymbol{R}^{j-1}(\boldsymbol{z})^c\right)^c \\
&= \left\{\left(\boldsymbol{R}^j(\boldsymbol{z}') \backslash \boldsymbol{R}^j(\boldsymbol{z})\right) \cap \boldsymbol{R}^{j-1}(\boldsymbol{z}')^c\right\} \cup \left\{\left(\boldsymbol{R}^{j-1}(\boldsymbol{z}) \backslash \boldsymbol{R}^{j-1}(\boldsymbol{z}')\right) \cap \boldsymbol{R}^j(\boldsymbol{z}')\right\} \\
&= \Delta^j(\boldsymbol{z}', \boldsymbol{z}) \cap \boldsymbol{R}^{j-1}(\boldsymbol{z}')^c,
\end{aligned}$$

where the last equality is by (D.12) that $\boldsymbol{R}^{j-1}(\boldsymbol{z}) \subset \boldsymbol{R}^{j-1}(\boldsymbol{z}')$. But

$$\Delta^j(\boldsymbol{z}', \boldsymbol{z}) \cap \boldsymbol{R}^{j-1}(\boldsymbol{z}')^c = \Delta^j(\boldsymbol{z}', \boldsymbol{z}) \backslash \left(\Delta^j(\boldsymbol{z}', \boldsymbol{z}) \backslash \boldsymbol{R}^{j-1}(\boldsymbol{z}')\right).$$

Note that

$$\Delta^{j-1}(\boldsymbol{z}', \boldsymbol{z}) \backslash \boldsymbol{R}^j(\boldsymbol{z}) = \Delta^j(\boldsymbol{z}', \boldsymbol{z}) \backslash \boldsymbol{R}^{j-1}(\boldsymbol{z}') \equiv A^*, \tag{D.17}$$

because

$$\Delta^{j-1}(\boldsymbol{z}',\boldsymbol{z})\backslash\boldsymbol{R}^j(\boldsymbol{z}) = \boldsymbol{R}^{j-1}(\boldsymbol{z}')\cap\boldsymbol{R}^{j-1}(\boldsymbol{z})^c\cap\boldsymbol{R}^j(\boldsymbol{z})^c = \boldsymbol{R}^{j-1}(\boldsymbol{z}')\cap\boldsymbol{R}^j(\boldsymbol{z})^c$$
$$= \boldsymbol{R}^j(\boldsymbol{z}')\cap\boldsymbol{R}^j(\boldsymbol{z})^c\cap\boldsymbol{R}^{j-1}(\boldsymbol{z}') = \Delta^j(\boldsymbol{z}',\boldsymbol{z})\cap\boldsymbol{R}^{j-1}(\boldsymbol{z}'),$$

where the second equality is by $\boldsymbol{R}^{j-1}(\boldsymbol{z}) \subset \boldsymbol{R}^j(\boldsymbol{z})$ and the third equality is by $\boldsymbol{R}^{j-1}(\boldsymbol{z}') \subset \boldsymbol{R}^j(\boldsymbol{z}')$. In sum,

$$\Delta_j(\boldsymbol{z},\boldsymbol{z}') = \Delta^{j-1}(\boldsymbol{z}',\boldsymbol{z})\backslash A^*, \qquad -\Delta_j(\boldsymbol{z},\boldsymbol{z}') = \Delta^j(\boldsymbol{z}',\boldsymbol{z})\backslash A^*. \tag{D.18}$$

(D.18) shows how the outflow ($\Delta_j(\boldsymbol{z},\boldsymbol{z}')$) and inflow ($-\Delta_j(\boldsymbol{z},\boldsymbol{z}')$) of $\boldsymbol{R}_j$ can be written in terms of the inflows of $\boldsymbol{R}^{j-1}$ and $\boldsymbol{R}^j$, respectively. And figuratively, A^* adjusts for the "leakage" when the change from $\boldsymbol{z}$ to $\boldsymbol{z}'$ is relatively large. Therefore, by (D.18), we have the inflow and outflow match result between $\boldsymbol{R}_j$ and $\boldsymbol{R}_{j-1}$:

$$\Delta_j(\boldsymbol{z},\boldsymbol{z}') = -\Delta_{j-1}(\boldsymbol{z},\boldsymbol{z}') \tag{D.19}$$

Now, we want to decompose this match into matches of flows in individual $\boldsymbol{R}_{\boldsymbol{d}^j}$'s. Define

$$\Delta_{\boldsymbol{d}^j}(\boldsymbol{z},\boldsymbol{z}') \equiv \boldsymbol{R}^*_{\boldsymbol{d}^j}(\boldsymbol{z})\backslash\boldsymbol{R}^*_{\boldsymbol{d}^j}(\boldsymbol{z}'), \tag{D.20}$$
$$-\Delta_{\boldsymbol{d}^j}(\boldsymbol{z},\boldsymbol{z}') \equiv \boldsymbol{R}^*_{\boldsymbol{d}^j}(\boldsymbol{z}')\backslash\boldsymbol{R}^*_{\boldsymbol{d}^j}(\boldsymbol{z}). \tag{D.21}$$

By Assumption EQ (or (D.4)),

$$\Delta_{\boldsymbol{d}^j}(\boldsymbol{z},\boldsymbol{z}') = \boldsymbol{R}^*_{\boldsymbol{d}^j}(\boldsymbol{z})\backslash\boldsymbol{R}_j(\boldsymbol{z}'),$$
$$-\Delta_{\boldsymbol{d}^j}(\boldsymbol{z},\boldsymbol{z}') = \boldsymbol{R}^*_{\boldsymbol{d}^j}(\boldsymbol{z}')\backslash\boldsymbol{R}_j(\boldsymbol{z}),$$

and therefore,

$$\Delta_j(\boldsymbol{z},\boldsymbol{z}') = \bigcup_{\boldsymbol{d}^j}\Delta_{\boldsymbol{d}^j}(\boldsymbol{z},\boldsymbol{z}'),$$
$$-\Delta_j(\boldsymbol{z},\boldsymbol{z}') = \bigcup_{\boldsymbol{d}^j}-\Delta_{\boldsymbol{d}^j}(\boldsymbol{z},\boldsymbol{z}'),$$

since $\boldsymbol{R}_j(\cdot) = \bigcup_{\boldsymbol{d}^j} \boldsymbol{R}^*_{\boldsymbol{d}^j}(\cdot)$. Also, note that $\{\Delta_{\boldsymbol{d}^j}(\boldsymbol{z},\boldsymbol{z}')\}_{\boldsymbol{d}^j}$ are disjoint since $\{\boldsymbol{R}^*_{\boldsymbol{d}^j}(\boldsymbol{z})\}_{\boldsymbol{d}^j}$ are disjoint. Therefore, $\{\Delta_{\boldsymbol{d}^j}(\boldsymbol{z},\boldsymbol{z}')\}_{\boldsymbol{d}^j}$ and $\{-\Delta_{\boldsymbol{d}^j}(\boldsymbol{z},\boldsymbol{z}')\}_{\boldsymbol{d}^j}$ are partitions of $\Delta_j(\boldsymbol{z},\boldsymbol{z}')$ and $-\Delta_j(\boldsymbol{z},\boldsymbol{z}')$, respectively. Then, suppressing $(\boldsymbol{z},\boldsymbol{z}')$, rewrite (D.19) as

$$\bigcup_{\boldsymbol{d}^j} \Delta_{\boldsymbol{d}^j} = \bigcup_{\boldsymbol{d}^{j-1}} -\Delta_{\boldsymbol{d}^{j-1}}.$$

Note that, for any $\boldsymbol{d}^j$ and $\boldsymbol{d}^{j-1}$, $\Delta_{\boldsymbol{d}^j}$ does not necessarily coincide with $-\Delta_{\boldsymbol{d}^{j-1}}$. Therefore, we proceed as follows. For a given $\bar{\boldsymbol{d}}^j$, we further partition $\Delta_{\bar{\boldsymbol{d}}^j}$ by considering $\{\Delta_{\bar{\boldsymbol{d}}^j,\boldsymbol{d}^{j-1}}\}_{\boldsymbol{d}^{j-1}}$ with $\Delta_{\bar{\boldsymbol{d}}^j,\boldsymbol{d}^{j-1}} = \emptyset$ for $\boldsymbol{d}^j \neq \bar{\boldsymbol{d}}^j$ and $\boldsymbol{d}^j \in \mathcal{D}^>(\boldsymbol{d}^{j-1})$. Likewise, for a given $\bar{\boldsymbol{d}}^{j-1}$, partition $-\Delta_{\bar{\boldsymbol{d}}^{j-1}}$ by considering $\{-\Delta_{\bar{\boldsymbol{d}}^{j-1},\boldsymbol{d}^j}\}_{\boldsymbol{d}^j}$ with $-\Delta_{\bar{\boldsymbol{d}}^{j-1},\boldsymbol{d}^j} = \emptyset$ for $\boldsymbol{d}^{j-1} \neq \bar{\boldsymbol{d}}^{j-1}$ and $\boldsymbol{d}^{j-1} \in \mathcal{D}^<(\boldsymbol{d}^j)$. Then,

$$\Delta_{\boldsymbol{d}^j} = \bigcup_{\boldsymbol{d}^{j-1}} \Delta_{\boldsymbol{d}^j,\boldsymbol{d}^{j-1}}, \tag{D.22}$$

$$-\Delta_{\boldsymbol{d}^{j-1}} = \bigcup_{\boldsymbol{d}^j} -\Delta_{\boldsymbol{d}^{j-1},\boldsymbol{d}^j}, \tag{D.23}$$

with

$$\Delta_{\boldsymbol{d}^j,\boldsymbol{d}^{j-1}} = -\Delta_{\boldsymbol{d}^{j-1},\boldsymbol{d}^j}. \tag{D.24}$$

Now, for a given $\boldsymbol{d}^j$ and $j = 1, ..., S-1$,

$$\begin{aligned}
&h_{\boldsymbol{d}^j}(\boldsymbol{z},\boldsymbol{z}',x) \\
&= \int_{\boldsymbol{R}^*_{\boldsymbol{d}^j}(\boldsymbol{z})} \vartheta(\boldsymbol{d}^j, x; \boldsymbol{u}) d\boldsymbol{u} - \int_{\boldsymbol{R}^*_{\boldsymbol{d}^j}(\boldsymbol{z}')} \vartheta(\boldsymbol{d}^j, x; \boldsymbol{u}) d\boldsymbol{u} \\
&= \int_{\Delta_{\boldsymbol{d}^j}} \vartheta(\boldsymbol{d}^j, x; \boldsymbol{u}) d\boldsymbol{u} - \int_{-\Delta_{\boldsymbol{d}^j}} \vartheta(\boldsymbol{d}^j, x; \boldsymbol{u}) d\boldsymbol{u} \\
&= \sum_{\boldsymbol{d}^{j-1}} \int_{\Delta_{\boldsymbol{d}^j,\boldsymbol{d}^{j-1}}} \vartheta(\boldsymbol{d}^j, x; \boldsymbol{u}) d\boldsymbol{u} - \sum_{\boldsymbol{d}^{j+1}} \int_{-\Delta_{\boldsymbol{d}^j,\boldsymbol{d}^{j+1}}} \vartheta(\boldsymbol{d}^j, x; \boldsymbol{u}) d\boldsymbol{u} \\
&= \sum_{\boldsymbol{d}^{j-1}} \int_{\Delta_{\boldsymbol{d}^j,\boldsymbol{d}^{j-1}}} \vartheta(\boldsymbol{d}^j, x; \boldsymbol{u}) d\boldsymbol{u} - \sum_{\boldsymbol{d}^{j+1}} \int_{\Delta_{\boldsymbol{d}^{j+1},\boldsymbol{d}^j}} \vartheta(\boldsymbol{d}^j, x; \boldsymbol{u}) d\boldsymbol{u},
\end{aligned} \tag{D.25}$$

where the second equality is by (D.22)–(D.23), and the third equality is by (D.24). Also, for $j = 0$,

$$\begin{aligned}
&\int_{\boldsymbol{R}^*_{\boldsymbol{d}^0}(\boldsymbol{z})} \vartheta(\boldsymbol{d}^0, x; \boldsymbol{u}) d\boldsymbol{u} - \int_{\boldsymbol{R}^*_{\boldsymbol{d}^0}(\boldsymbol{z}')} \vartheta(\boldsymbol{d}^0, x; \boldsymbol{u}) d\boldsymbol{u} \\
&= - \sum_{\boldsymbol{d}^1} \int_{\Delta_{\boldsymbol{d}^1, \boldsymbol{d}^0}} \vartheta(\boldsymbol{d}^0, x; \boldsymbol{u}) d\boldsymbol{u},
\end{aligned} \tag{D.26}$$

since $\Delta_{\boldsymbol{d}^0}(\boldsymbol{z}, \boldsymbol{z}') = \emptyset$ by the choice of $(\boldsymbol{z}, \boldsymbol{z}')$. And, for $j = S$,

$$\begin{aligned}
&\int_{\boldsymbol{R}^*_{\boldsymbol{d}^S}(\boldsymbol{z})} \vartheta(\boldsymbol{d}^S, x; \boldsymbol{u}) d\boldsymbol{u} - \int_{\boldsymbol{R}^*_{\boldsymbol{d}^S}(\boldsymbol{z}')} \vartheta(\boldsymbol{d}^S, x; \boldsymbol{u}) d\boldsymbol{u} \\
&= \sum_{\boldsymbol{d}^{S-1}} \int_{\Delta_{\boldsymbol{d}^S, \boldsymbol{d}^{S-1}}} \vartheta(\boldsymbol{d}^S, x; \boldsymbol{u}) d\boldsymbol{u},
\end{aligned} \tag{D.27}$$

since $-\Delta_S(\boldsymbol{z}, \boldsymbol{z}') = \emptyset$ by the choice of $(\boldsymbol{z}, \boldsymbol{z}')$. Therefore, by combining (D.25)–(D.27), we have

$$\begin{aligned}
h(\boldsymbol{z}, \boldsymbol{z}', x) &= \sum_{j=0}^{S} \sum_{\boldsymbol{d}^j} \left\{ \int_{\boldsymbol{R}^*_{\boldsymbol{d}^j}(\boldsymbol{z})} \vartheta(\boldsymbol{d}^j, x; \boldsymbol{u}) d\boldsymbol{u} - \int_{\boldsymbol{R}^*_{\boldsymbol{d}^j}(\boldsymbol{z}')} \vartheta(\boldsymbol{d}^j, x; \boldsymbol{u}) d\boldsymbol{u} \right\} \\
&= \sum_{j=0}^{S} \sum_{\boldsymbol{d}^j} \left\{ \sum_{\boldsymbol{d}^{j-1}} \int_{\Delta_{\boldsymbol{d}^j, \boldsymbol{d}^{j-1}}} \vartheta(\boldsymbol{d}^j, x; \boldsymbol{u}) d\boldsymbol{u} - \sum_{\boldsymbol{d}^{j+1}} \int_{\Delta_{\boldsymbol{d}^{j+1}, \boldsymbol{d}^j}} \vartheta(\boldsymbol{d}^j, x; \boldsymbol{u}) d\boldsymbol{u} \right\} \\
&= \sum_{j=1}^{S} \sum_{\boldsymbol{d}^j} \sum_{\boldsymbol{d}^{j-1}} \int_{\Delta_{\boldsymbol{d}^j, \boldsymbol{d}^{j-1}}} \{ \vartheta(\boldsymbol{d}^j, x; \boldsymbol{u}) - \vartheta(\boldsymbol{d}^{j-1}, x; \boldsymbol{u}) \} d\boldsymbol{u} \\
&= \sum_{\boldsymbol{d}_{-s}} \int_{\Delta_{(1, \boldsymbol{d}_{-s}), (0, \boldsymbol{d}_{-s})}} \{ \vartheta((1, \boldsymbol{d}_{-s}), x; \boldsymbol{u}) - \vartheta((0, \boldsymbol{d}_{-s}), x; \boldsymbol{u}) \} d\boldsymbol{u},
\end{aligned}$$

where the last equality is by the definition of $\Delta_{\boldsymbol{d}^j, \boldsymbol{d}^{j-1}}$. Also, by a similar argument, we can show that

$$\begin{aligned}
\tilde{h}(\boldsymbol{z}, \boldsymbol{z}', \tilde{\boldsymbol{x}}) &= \sum_{j=1}^{S} \sum_{\boldsymbol{d}^j} \sum_{\boldsymbol{d}^{j-1}} \int_{\Delta_{\boldsymbol{d}^j, \boldsymbol{d}^{j-1}}} \{ \vartheta(\boldsymbol{d}^j, x_j; \boldsymbol{u}) - \vartheta(\boldsymbol{d}^{j-1}, x_j; \boldsymbol{u}) \} d\boldsymbol{u} \\
&= \sum_{j=1}^{S} \sum_{(1, \boldsymbol{d}_{-s}) \in \mathcal{D}^j} \int_{\Delta_{(1, \boldsymbol{d}_{-s}), (0, \boldsymbol{d}_{-s})}} \{ \vartheta((1, \boldsymbol{d}_{-s}), x_j; \boldsymbol{u}) - \vartheta((0, \boldsymbol{d}_{-s}), x_{j-1}; \boldsymbol{u}) \} d\boldsymbol{u}.
\end{aligned} \tag{D.28}$$

This completes the proof of Lemma D.2.

Now we prove Lemma 3.2. For part (i), suppose that $\vartheta(1, \boldsymbol{d}_{-s}, x; \boldsymbol{u}) - \vartheta(0, \boldsymbol{d}_{-s}, x; \boldsymbol{u}) > 0$ a.e. $\boldsymbol{u}$ $\forall \boldsymbol{d}_{-s}, s$. Then by (D.11), $h > 0$. Conversely, if $h > 0$ then it should be that $\vartheta(1, \boldsymbol{d}_{-s}, x; \boldsymbol{u}) - \vartheta(0, \boldsymbol{d}_{-s}, x; \boldsymbol{u}) > 0$ a.e. $\boldsymbol{u}$ $\forall \boldsymbol{d}_{-s}, s$. Suppose not and suppose $\vartheta(1, \boldsymbol{d}_{-s}, x; \boldsymbol{u}) - \vartheta(0, \boldsymbol{d}_{-s}, x; \boldsymbol{u}) \leq 0$ with positive measure for some $\boldsymbol{d}_{-s}$ and s. Then by Assumption M, this implies that $\vartheta(1, \boldsymbol{d}_{-s}, x; \boldsymbol{u}) - \vartheta(0, \boldsymbol{d}_{-s}, x; \boldsymbol{u}) \leq 0 \ \forall \boldsymbol{d}_{-s}, s$ a.e. $\boldsymbol{u}$, and thus $h \leq 0$ which is contradiction. By applying similar arguments for other signs, we have the desired result. Now we prove part (ii). Note that (D.28) can be rewritten as

$$\tilde{h}(\boldsymbol{z}, \boldsymbol{z}', \tilde{\boldsymbol{x}}) - \sum_{k \neq j} \sum_{(1,\boldsymbol{d}_{-s}) \in \mathcal{D}^k} \int_{\Delta_{(1,\boldsymbol{d}_{-s}),(0,\boldsymbol{d}_{-s})}} \{\vartheta((1, \boldsymbol{d}_{-s}), x_k; \boldsymbol{u}) - \vartheta((0, \boldsymbol{d}_{-s}), x_{k-1}; \boldsymbol{u})\} \, d\boldsymbol{u}$$
$$= \sum_{(1,\boldsymbol{d}_{-s}) \in \mathcal{D}^j} \int_{\Delta_{(1,\boldsymbol{d}_{-s}),(0,\boldsymbol{d}_{-s})}} \{\vartheta((1, \boldsymbol{d}_{-s}), x_j; \boldsymbol{u}) - \vartheta((0, \boldsymbol{d}_{-s}), x_{j-1}; \boldsymbol{u})\} \, d\boldsymbol{u}. \tag{D.29}$$

We prove the case $\iota = 1$; the proof for the other cases follows symmetrically. For $k \neq j$, when $-\vartheta((1, \boldsymbol{d}_{-s}), x_k; \boldsymbol{u}) + \vartheta((0, \boldsymbol{d}_{-s}), x_{k-1}; \boldsymbol{u}) > 0$ a.e. $\boldsymbol{u}$ $\forall (1, \boldsymbol{d}_{-s}) \in \mathcal{D}^k$, it satisfies

$$- \sum_{(1,\boldsymbol{d}_{-s}) \in \mathcal{D}^k} \int_{\Delta_{(1,\boldsymbol{d}_{-s}),(0,\boldsymbol{d}_{-s})}} \{\vartheta((1, \boldsymbol{d}_{-s}), x_k; \boldsymbol{u}) - \vartheta((0, \boldsymbol{d}_{-s}), x_{k-1}; \boldsymbol{u})\} \, d\boldsymbol{u} > 0.$$

Combining with $\tilde{h}(\boldsymbol{z}, \boldsymbol{z}', \tilde{\boldsymbol{x}}) > 0$ implies that the l.h.s. of (D.29) is positive. This implies that $\vartheta((1, \boldsymbol{d}_{-s}), x_j; \boldsymbol{u}) - \vartheta((0, \boldsymbol{d}_{-s}), x_{j-1}; \boldsymbol{u}) > 0$ a.e. $\boldsymbol{u}$ $\forall (1, \boldsymbol{d}_{-s}) \in \mathcal{D}^j$. If not, then it results in a contradiction as in the previous argument.

D.7 Proof of Theorem 3.3

Consider

$$E[Y_{\boldsymbol{d}^j} | X = x] = E[Y | \boldsymbol{D} = \boldsymbol{d}^j, \boldsymbol{Z} = \boldsymbol{z}, X = x] \Pr[\boldsymbol{D} = \boldsymbol{d}^j | \boldsymbol{Z} = \boldsymbol{z}]$$
$$+ \sum_{\boldsymbol{d}' \neq \boldsymbol{d}^j} E[Y_{\boldsymbol{d}^j} | \boldsymbol{D} = \boldsymbol{d}', \boldsymbol{Z} = \boldsymbol{z}, X = x] \Pr[\boldsymbol{D} = \boldsymbol{d}' | \boldsymbol{Z} = \boldsymbol{z}]. \tag{D.30}$$

Consider $j' < j$ for $E[Y_{\boldsymbol{d}^j} | \boldsymbol{D} = \boldsymbol{d}^{j'}, \boldsymbol{Z}, X]$ in (D.30) with $\boldsymbol{d}^{j'} \in \mathcal{D}^{<}(\boldsymbol{d}^j)$. Then, for example, if $(x_k, x_{k-1}) \in \mathcal{X}_{k,k-1}(-1) \cup \mathcal{X}_{k,k-1}(0)$ for $j' + 1 \leq k \leq j$, then $\vartheta(\boldsymbol{d}^j, x; \boldsymbol{u}) \leq \vartheta(\boldsymbol{d}^{j'}, x'; \boldsymbol{u})$ where $x = x_j$

and $x' = x_{j'}$ by transitively applying (3.13). Therefore

$$
\begin{aligned}
E[Y_{\boldsymbol{d}^j}|\boldsymbol{D}=\boldsymbol{d}^{j'},\boldsymbol{Z}=\boldsymbol{z},X=x] &= E[\theta(\boldsymbol{d}^j,x,\epsilon)|\boldsymbol{U}\in R_{\boldsymbol{d}^{j'}}(\boldsymbol{z}),\boldsymbol{Z}=\boldsymbol{z},X=x] \\
&= \frac{1}{\Pr[\boldsymbol{U}\in R_{\boldsymbol{d}^{j'}}(\boldsymbol{z})]}\int_{R_{\boldsymbol{d}^{j'}}(\boldsymbol{z})}\vartheta(\boldsymbol{d}^j,x;\boldsymbol{u})d\boldsymbol{u} \\
&\le \frac{1}{\Pr[\boldsymbol{U}\in R_{\boldsymbol{d}^{j'}}(\boldsymbol{z})]}\int_{R_{\boldsymbol{d}^{j'}}(\boldsymbol{z})}\vartheta(\boldsymbol{d}^{j'},x';\boldsymbol{u})d\boldsymbol{u} \\
&= E[\theta(\boldsymbol{d}^{j'},x',\epsilon)|\boldsymbol{U}\in R_{\boldsymbol{d}^{j'}}(\boldsymbol{z}),\boldsymbol{Z}=\boldsymbol{z},X=x'] \\
&= E[Y|\boldsymbol{D}=\boldsymbol{d}^{j'},\boldsymbol{Z}=\boldsymbol{z},X=x']. \qquad \text{(D.31)}
\end{aligned}
$$

Symmetrically, for $j' > j$, if $(x_k, x_{k-1}) \in \mathcal{X}_{k,k-1}(1) \cup \mathcal{X}_{k,k-1}(0)$ for $j+1 \le k \le j'$, then $\vartheta(\boldsymbol{d}^j, x; \boldsymbol{u}) \le \vartheta(\boldsymbol{d}^{j'}, x'; \boldsymbol{u})$ where $\boldsymbol{d}^{j'} \in \mathcal{D}^{>}(\boldsymbol{d}^j)$, $x = x_j$ and $x' = x_{j'}$. Therefore the same bound as (D.31) is derived. Given these results, to collect all $x' \in \mathcal{X}$ that yield $\vartheta(\boldsymbol{d}^j, x; \boldsymbol{u}) \le \vartheta(\boldsymbol{d}^{j'}, x'; \boldsymbol{u})$ for $\boldsymbol{d}^{j'} \in \mathcal{D}^{<}(\boldsymbol{d}^j) \cup \mathcal{D}^{>}(\boldsymbol{d}^j)$, we can construct a set

$$
\begin{aligned}
x' \in &\left\{x_{j'} : (x_k, x_{k-1}) \in \mathcal{X}_{k,k-1}(-1) \cup \mathcal{X}_{k,k-1}(0) \text{ for } j'+1 \le k \le j, x_j = x\right\} \\
&\cup \left\{x_{j'} : (x_k, x_{k-1}) \in \mathcal{X}_{k,k-1}(1) \cup \mathcal{X}_{k,k-1}(0) \text{ for } j+1 \le k \le j', x_j = x\right\}.
\end{aligned}
$$

Then we can further shrink the bound in (D.31) by taking the infimum over all x' in this set. The lower bound on $E[Y_{\boldsymbol{d}^j}|\boldsymbol{D}=\boldsymbol{d}^{j'},\boldsymbol{Z}=\boldsymbol{z},X=x]$ can be constructed by simply choosing the opposite signs in the preceding argument. Since the other terms in (D.30) are observed, we have the desired bounds in the theorem.

E Airline and Pollution Data

We combine data spanning the period 2000–2015 from two sources: airline data from the U.S. Department of Transportation and pollution data from the Environmental Protection Agency (EPA).

Airline Data. Our first data source contains airline information and combines publicly available data from the Department of Transportation's Origin and Destination Survey (DB1B) and Domestic Segment (T-100) database. These datasets have been used extensively in the literature to analyze the airline industry (see, e.g., Borenstein (1989), Berry (1992), Ciliberto and Tamer (2009), and more recently, Li et al. (2018) and Ciliberto et al. (2018)). The DB1B database is a quarterly sample of all passenger domestic itineraries. The dataset contains coupon-specific information, including origin and destination airports, number of coupons, the corresponding operating carriers, number of passengers, prorated market fare, market miles flown, and distance. The T-100 dataset is a monthly census of all domestic flights broken down by airline, and origin and destination airports.

Our time-unit of analysis is a quarter and we define a market as the market for air connection between a pair of airports (regardless of intermediate stops) in a given quarter.[33] We restrict the sample to include the top 100 metropolitan statistical areas (MSA's), ranked by population at the beginning of our sample period. We follow Berry (1992) and Ciliberto and Tamer (2009) and define an airline as actively serving a market in a given quarter, if we observe at least 90 passengers in the DB1B survey flying with the airline in the corresponding quarter.[34] We exclude from our sample city pairs in which no airline operates in the whole sample period. Notice that we do include markets that are temporarily not served by any airline. This leaves us with 181,095 market-quarter observations.

In our analysis, we allow for airlines to have a heterogeneous effect on pollution, and to simplify computation, in each market we allow for six potential participants: American (AA), Delta (DL), United (UA), Southwest (WN), a medium-size airline, and a low-cost carrier.[35] The latter is not a bad approximation to the data in that we rarely observe more than one medium-size or low-cost in a market but it assumes that all low-cost airlines have the same strategic behavior, and so do the medium airlines. Table 1 shows the number of firms in each market broken down by size as measured by population. As the table shows, market size alone does not explain market structure, a point first made by Ciliberto and Tamer (2009).

In our application, we consider two instruments for the entry decisions. The first is the *airport presence* of an airline proposed by Berry (1992). For a given airline, this variable is constructed as the number of markets it serves out of an airport as a fraction of the total number of markets served by all airlines out of the airport. A hub-and-spoke network allows firms to exploit demand-side and cost-side economies, which should affect the firm's profitability. While Berry (1992) assumes that an airline's airport presence only affects its own profits (and hence, is excluded from rivals' profits),

[33]In cities that operate more than one airport, we assume that flights to different airports in the same metropolitan area are in separate markets.

[34]This corresponds to approximately the number of passengers that would be carried on a medium-size jet operating once a week.

[35]That is, to limit the number of potential market structures, we lump together all the low cost carriers into one category, and Northwest, Continental, America West, and USAir under the medium airline type.

Table 1: Distribution of the Number of Carriers by Market Size

	Market size			
# firms	Large	Medium	Small	Total
0	7.96	8.20	8.62	8.18
1	41.18	22.53	20.58	30.30
2	28.14	23.41	21.25	25.04
3	12.65	20.00	16.67	16.05
4	7.65	14.72	15.17	11.51
5	1.98	9.90	16.48	7.80
6+	0.52	1.23	2.21	1.12
# markets	79,326	64,191	37,578	181,095

Table 2: Airline Summary Statistics

		American	Delta	United	Southwest	medium	low-cost
Market presence (0/1)	mean	0.44	0.57	0.28	0.25	0.56	0.17
	sd	0.51	0.51	0.46	0.44	0.51	0.38
Airport presence (%)	mean	0.43	0.56	0.27	0.25	0.39	0.10
	sd	0.17	0.18	0.16	0.18	0.14	0.08
Cost (%)	mean	0.71	0.41	0.76	0.29	0.22	0.04
	sd	1.56	1.28	1.43	0.83	0.60	0.17

Ciliberto and Tamer (2009) argue that this may not be the case in practice, since airport presence might be a measure of product differentiation, rendering it likely to enter the profit function of all firms through demand. While an instrument that enters all of the profit functions is fine in our context (see Appendix C.4), we also consider the instrument proposed by Ciliberto and Tamer (2009), which captures shocks to the fixed cost of providing a service in a market. This variable, which they call *cost*, is constructed as the percentage of the nonstop distance that the airline must travel in excess of the nonstop distance, if the airline uses a connecting instead of a nonstop flight.[36] Arguably, this variable only affects its own profits and is excluded from rivals' profits.

Table 2 presents the summary statistics of the airline related variables. Of the leading airlines, we see that American and Delta are present in about half of the markets, while United and Southwest are only present in about a quarter of the markets. American and Delta tend to dominate the airports in which they operate more than United and Southwest. From the cost variable, we see that both American and United tend to operate a hub-and-spoke network, while Southwest (and to a lesser extent Delta) operates most markets nonstop.

[36]Mechanically, the variable is constructed as the difference between the sum of the distances of a market's endpoints and the closest hub of an airline, and the nonstop distance between the endpoints, divided by the nonstop distance.

Pollution Data. The second component of our dataset is the air pollution data. The EPA compiles a database of outdoor concentrations of pollutants measured at more than 4,000 monitoring stations throughout the U.S., owned and operated mainly by state environmental agencies. Each monitoring station is geocoded, and hence, we are able to merge these data with the airline dataset by matching all the monitoring stations that are located within a 10km radius of each airport in our first dataset.

The principal emissions of aircraft include the greenhouse gases carbon dioxide (CO_2) and water vapor (H_2O), which have a direct impact on climate change. Aircraft jet engines also produce nitric oxide (NO) and nitrogen dioxide (NO_2) (which together are termed nitrogen oxides (NO_x)), carbon monoxide (CO), oxides of sulphur (SO_x), unburned or partially combusted hydrocarbons (also known as volatile organic compounds or VOC's), particulates, and other trace compounds (see, Federal Aviation Administration (2015)). In addition, ozone (O_3) is formed by the reaction of VOC's and NO_x in the presence of heat and sunlight. The set of pollutants other than CO_2 are more pernicious in that they can harm human health directly and can result in respiratory, cardiovascular, and neurological conditions. Research to date indicates that fine particulate matter (PM) is responsible for the majority of the health risks from aviation emissions, although ozone has a substantial health impact too.[37] Therefore, as our measure of pollution, we will consider both.

Our measure of ozone is a quarterly mean of daily maximum levels in parts per million. In terms of PM, as a general rule, the smaller the particle the further it travels in the atmosphere, the longer it remains suspended in the atmosphere, and the more risk it poses to human health. PM that measure less than 2.5 micrometer can be readily inhaled, and thus, potentially pose increased health risks. The variable PM2.5 is a quarterly average of daily averages and is measured in micrograms/cubic meter. For each airport in our sample, we take an average (weighted by distance to the airport) of the data from all air monitoring stations within a 10km radius. The top panel of Table 3 shows the summary statistics of the pollution measures.

Other Market-Level Controls. We also include in our analysis market-level covariates that may affect both market structure and pollution levels. In particular, we construct a measure of market size by computing the (geometric) mean of the MSA populations at the market endpoints and a measure of economic activity by computing the average per capita income at the market

[37] See Federal Aviation Administration (2015).

Table 3: Market-level Summary Statistics

	Mean	Std. Dev.
Pollution		
Ozone (O_3)	.0477	.0056
Particulate matter (PM2.5)	8.3881	2.5287
Other controls		
Market size (pop.)	2307187.8	1925533.4
Income (per capita)	34281.6	4185.5
# of markets	181,095	

endpoints, using data from the Regional Economic Accounts of the Bureau of Economic Analysis.

Finally, as we mentioned in Section 3.4, having access to data on a variable that affects pollution but is excluded from the airline participation decisions can greatly help in calculating the bounds of the ATE. Therefore, we construct a variable that measures the economic activity of pollution related industries (manufacturing, construction, and transportation other than air transportation) in a given market (MSA) as a fraction of total economic activity in that market, again, using data from the Regional Economic Accounts of the Bureau of Economic Analysis.

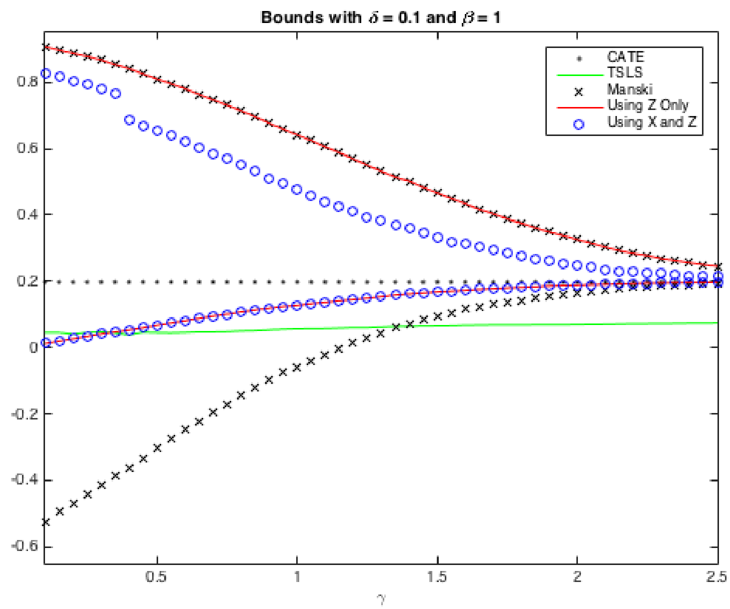


Figure 5: Bounds on the ATE with different strength of vector $\boldsymbol{Z} = (Z_1, Z_2)$ of binary instruments when X takes three different values ($|\mathcal{X}| = 3$).

This figure (and the next) depicts the simulated bounds for $E[Y_{11} - Y_{00}|X = 0] = 0.2$ (the straight dotted line). The horizontal axis is the value of the coefficients on the instruments ($\gamma_1 = \gamma_2 = \gamma$). The stronger the instruments, the narrower the bounds are. The cross lines are Manski (1990)'s bounds. The red solid lines are our bounds using only the variation of $\boldsymbol{Z}$, which identify the sign of the ATE. The blue circle lines are bounds where the variation of X, the exogenous variable excluded from the treatment selection process, is also used. Lastly, the green solid line is the simulated TSLS estimand assuming a linear simultaneous equations model.

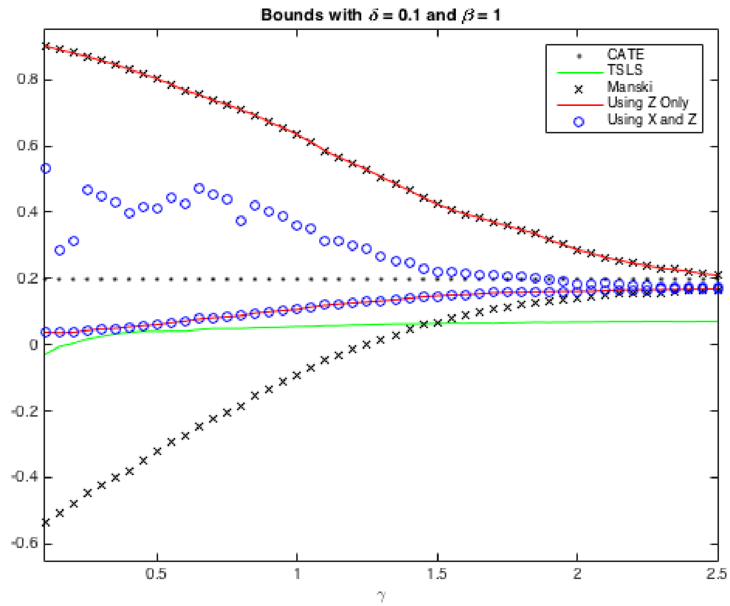


Figure 6: Bounds with different strength of vector $\boldsymbol{Z} = (Z_1, Z_2)$ of binary instrument when X takes fifteen different values ($|\mathcal{X}| = 15$).

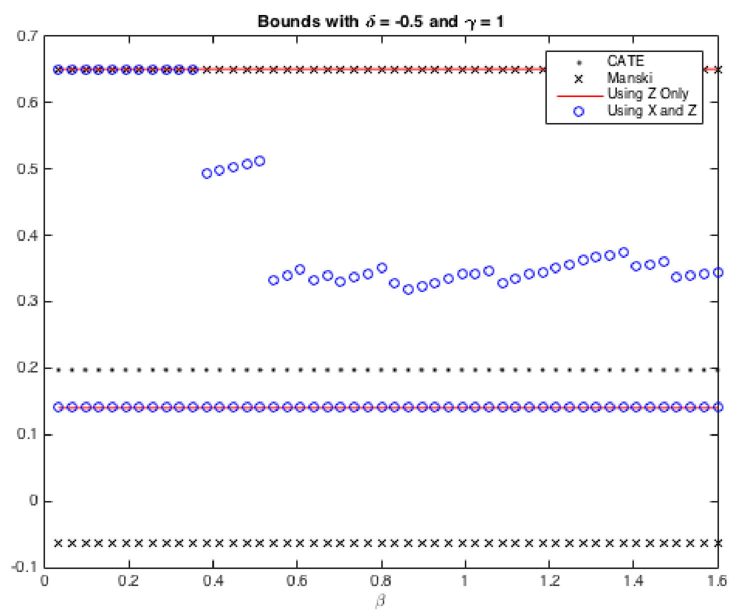


Figure 7: Bounds under Different Strength of X with $|\mathcal{X}| = 15$.

The horizontal axis is the value of the coefficient on the exogenous variable X excluded from the treatment selection process. The jumps in the bounds when both the variations of $\boldsymbol{Z}$ and X are used (the blue circle lines) are because different inequalities are involved for different values of the coefficient; see the text for details.

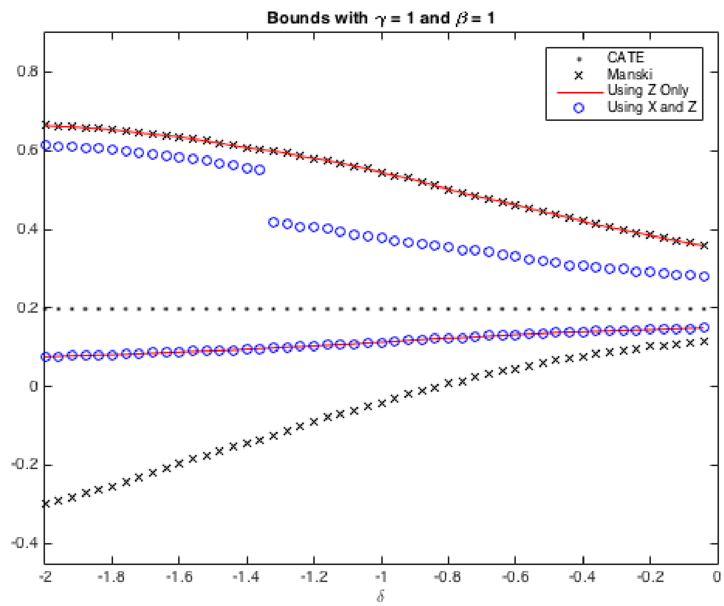
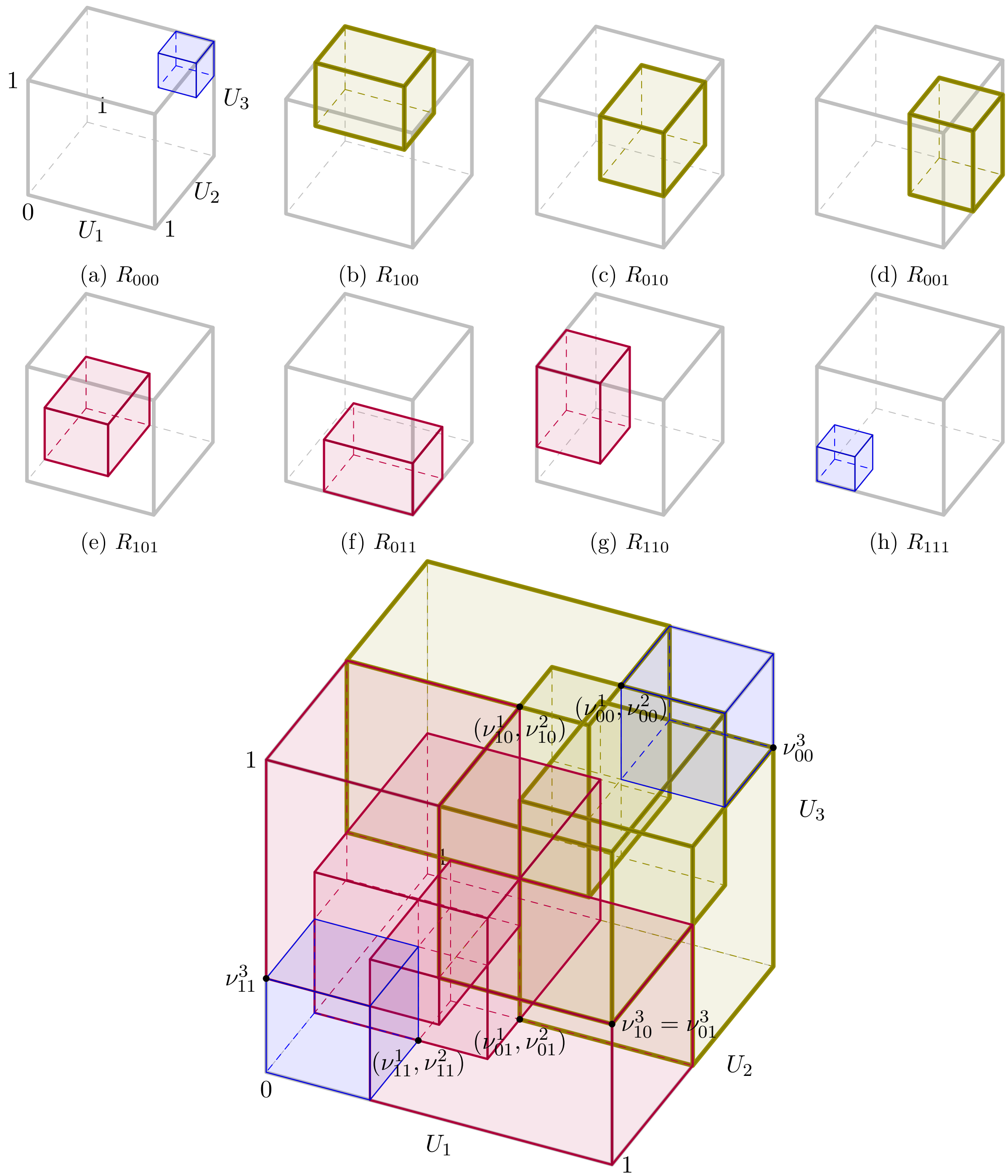


Figure 8: Bounds under Different Strength of Interaction with $|\mathcal{X}| = 3$.

The horizontal axis is the value of the coefficients on the opponents' decisions ($\delta_1 = \delta_2 = \delta$). The smaller the interaction effects, the narrower the bounds are. Again, the jumps in the bounds when both the variations of $\boldsymbol{Z}$ and X are used (the blue circle lines) are because different inequalities are involved for different values of the coefficient.



(a) R_{000} (b) R_{100} (c) R_{010} (d) R_{001}

(e) R_{101} (f) R_{011} (g) R_{110} (h) R_{111}

(i) $\bigcup_{0 \leq j \leq 3} \left\{ \bigcup_{d \in M_j} R_d \right\} = \mathcal{U} \equiv (0, 1]^3$

Figure 9: Regions of Equilibrium for $S = 3$.

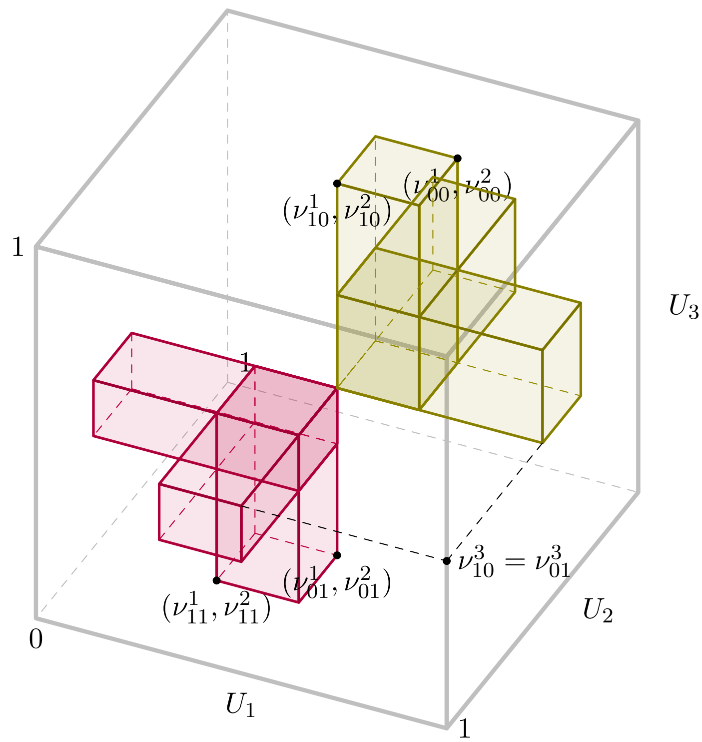


Figure 10: Depicting the Regions of Multiple Equilibria for $S = 3$.