Sharp Testable Implications of Encouragement Designs*

Yuehao Bai
Department of Economics
University of Southern California
yuehao.bai@usc.edu

Max Tabord-Meehan
Department of Economics
University of Chicago
maxtm@uchicago.edu

November 14, 2024

Abstract

This paper studies the sharp testable implications of an additive random utility model with a discrete multi-valued treatment and a discrete multi-valued instrument, in which each value of the instrument only weakly increases the utility of one choice. Borrowing the terminology used in randomized experiments, we call such a setting an encouragement design. We derive inequalities in terms of the conditional choice probabilities that characterize when the distribution of the observed data is consistent with such a model. Through a novel constructive argument, we further show these inequalities are sharp in the sense that any distribution of the observed data that satisfies these inequalities is generated by this additive random utility model.

KEYWORDS: Multi-valued treatment, instrumental variable, encouragement design, additive random utility model, moment inequalities

JEL classification codes: C14, C31, C35, C36

*We thank Kirill Ponomarev, Azeem Shaikh, and Alex Torgovitsky for helpful comments.

1 Introduction

This paper studies the sharp testable implications of an additive random utility model with a discrete multi-valued treatment and a discrete multi-valued instrument, in which each value of the instrument only weakly increases the (indirect) utility of one choice; such a model naturally arises in both experimental (see, for instance, Kline and Walters, 2016) and observational (see, for instance, Kirkeboen et al., 2016) settings. Borrowing the terminology used in randomized experiments, we call such a setting an encouragement design. When the treatment and the instrument are binary, this model is equivalent to the model studied in Imbens and Angrist (1994). We derive inequalities that characterize when the distribution of the observed data is consistent with such a model. These inequalities are in terms of the conditional choice probabilities, and are sharp in the sense that any distribution of the observed data that satisfies these inequalities is generated by this additive random utility model; in other words, they exhaust all information in the model. When the treatment and the instrument are both binary, we recover the inequalities in Balke and Pearl (1997a,b) and Kitagawa (2015). However, both the inequalities and our construction to establish their sharpness beyond this special case are, to our knowledge, novel to the literature.

We start by deriving the implications on the distribution of potential treatments that are equivalent to the additive random utility model. A key feature of these implications is that for each person, there exists a "default" choice (which possibly differs across people), which is the choice made if the instrument didn't exist; when the instrument equals j, then the person chooses either treatment j or the default choice. These implications on potential treatments immediately lead to a set of inequalities on conditional choice probabilities which are easy to interpret. To the best of our knowledge, these inequalities are new to the literature beyond the setting where both the treatment and the instrument are binary. Finally, we show through a novel constructive argument that these inequalities are sharp; that is, for each distribution of the observed data that satisfies these inequalities, we construct a distribution of the potential outcomes and potential treatments that is consistent with it.

To accommodate a larger class of empirical examples, we further allow that the utilities of certain choices are not affected by the instrument at all, and that there exists a "base state" of the instrument that does not change the utility of any choice. Examples of such settings include Kline and Walters (2016) and Kirkeboen et al. (2016). Both examples are also studied in Lee and Salanié (2023), who study identification of treatment effect parameters conditional on "response groups" defined as sets of possible values of potential treatments. When there is a "base state" of the instrument, the default choice further coincides with the choice made under this base state. Therefore, the inequalities simplify drastically, although the construction to show sharpness becomes slightly more complicated.

For the setting with a binary treatment, a binary instrument, and a binary outcome, Balke and Pearl (1997a,b) provide the first set of inequalities that sharply characterize when the distribution of the observed data is consistent with instrument exogeneity and monotonicity, using a linear programming formulation. Kitagawa (2015) studies these inequalities when the outcome is allowed to be continuous and, importantly, demonstrates that the inequalities are sharp constructively. He further proposes a corresponding test. Mourifié and Wan (2017) leverage the intersection bounds framework of Chernozhukov et al. (2013) to construct an alternative test based on the same testable implications. As shown by Vytlacil (2002), the model considered in the papers above is equivalent to the nonparametric selection model in Heckman and Vytlacil (2005), who also discuss testable implications in the binary setting with a possibly continuous instrument. Kédagni and Mourifié (2020) derive a set of sharp inequalities and a corresponding test for instrument exogeneity with a possibly multi-valued treatment and instrument. Sun (2023) derives a set of inequalities with a possibly multi-valued treatment and instrument, under instrument exogeneity and the "unordered monotonicity" assumption of Heckman and Pinto (2018), but does not establish that these are necessarily sharp. As we explain in Remark 3.2 below, however, the "unordered monotonicity" assumption is not implied by, nor does it imply, our assumption. Kwon and Roth (2024) characterize sharp testable implications in the setting with a multi-valued treatment and a binary instrument[1].

Our paper is also related to a vast literature that studies identification and inference for treatment effects using instrumental variables. See, for example, Bhattacharya et al. (2008, 2012), Machado et al. (2019), Słoczyński (2020), Mogstad et al. (2021), as well as the comprehensive review article on instrumental variables by Mogstad and Torgovitsky (2024). Of particular relevance for our setting are the papers which study multi-valued treatments and instruments: see, for instance, Lee and Salanié (2018), Kamat et al. (2023), Bai et al. (2024a), Bai et al. (2024b), and Bhuller and Sigstad (2024).

The remainder of the paper is organized as follows. In Section 2, we describe our setup and notation. In Section 3, we characterize the set of inequalities implied by our model and show that these are sharp. For simplicity, we first state versions of our results in a setting with only a treatment and an instrument. In Section 4, we extend the results to a setting

[1]However, we note that they frame their contribution in the context of developing tests for mediation analysis.

with an additional, possibly continuous, outcome variable.

2 Setup and Notation

Let $J \geq 2$ be an integer. Let $D \in \{0, \ldots, J-1\}$ denote a multi-valued treatment choice and $Z \in \mathcal{Z} \subseteq \{0, \ldots, J-1\}$ denote a multi-valued instrument (in Section 4, we additionally consider an outcome variable $Y \in \mathcal{Y}$, but we ignore this for the time being). In the models that we consider below, each value of the instrument encourages towards or "targets" a unique choice. To accommodate our empirical examples of interest, we do not require that the support of Z be the same as that of D; specifically, we consider two forms of $\mathcal{Z}$: (1) $\mathcal{Z} = \{0, \ldots, J-1\}$ and (2) $\mathcal{Z} = \{0, J_0, \ldots, J-1\}$, for some $1 \leq J_0 \leq J-2$. The first form corresponds to a setting where every choice has a corresponding instrument value which encourages towards it. The second form corresponds to a setting where the first J_0 choices are not affected by the instrument, and in this case we interpret $Z = 0$ as the "base state" of the instrument. Let D_z for $z \in \mathcal{Z}$ denote the potential treatment choice when assigned to the instrument value z. As usual, the observed choice is related to potential treatment choices and the instrument through

$$D = \sum_{z \in \mathcal{Z}} D_z I\{Z = z\} \ . \tag{1}$$

In what follows, let Q denote the distribution of $((D_z : z \in \mathcal{Z}), Z)$ and P denote the distribution of (D, Z). Note that given the mapping T such that $D = T((D_z : z \in \mathcal{Z}), Z)$ as implied by (1), we obtain by construction that $P = QT^{-1}$. Throughout, we will impose the assumption that the instrument is exogenous, formally:

Assumption 2.1. $(D_z : z \in \mathcal{Z}) \perp\!\!\!\perp Z$ under Q.

We will further rule out degenerate situations by requiring that the instrument takes on each value in its support with strictly positive probability:

Assumption 2.2. $Q\{Z = z\} > 0$ for $z \in \mathcal{Z}$.

In what follows, we will frequently use the following facts about the relationship between P and Q. Suppose $P = QT^{-1}$ for some Q that satisfies Assumptions 2.1–2.2. Then, for $z \in \mathcal{Z}$,

$$P\{Z = z\} = Q\{Z = z\} > 0 \ .$$

Therefore, the conditional choice probabilities can be defined and they satisfy

$$P\{D = j|Z = z\} = Q\{D_z = j|Z = z\} = Q\{D_z = j\} \ , \tag{2}$$

where the first equality follows because $D = D_z$ when $Z = z$ by (1), and the second equality follows from Assumption 2.1.

Our goal is to study the necessary and sufficient conditions for P to be consistent with a certain additive random utility model, which we now describe. Formally, let $U_0, \ldots, U_{J-1} : \mathcal{Z} \to \mathbf{R}$ be unknown functions and let $(\epsilon_0, \ldots, \epsilon_{J-1})$ be a random vector of unobservables, the distribution of which is absolutely continuous with respect to the Lebesgue measure on $\mathbf{R}^J$. Further suppose $(\epsilon_0, \ldots, \epsilon_{J-1}) \perp\!\!\!\perp Z$. The additive random utility model maintains that the potential treatments D_z for $z \in \mathcal{Z}$ are given by

$$D_z \in \operatorname*{argmax}_{0 \le j \le J-1} \ (U_j(z) + \epsilon_j) \ . \tag{3}$$

Because the distribution of $(\epsilon_0, \ldots, \epsilon_{J-1})$ is absolutely continuous with respect to the Lebesgue measure on $\mathbf{R}^J$, ties happen with probability zero, and therefore with probability one, D_z is unique and

$$I\{D_z = j\} = I\{U_j(z) + \epsilon_j > U_k(z) + \epsilon_k \text{ for all } k \neq j\} \ . \tag{4}$$

We further assume that

$$U_j(z) = \alpha_j + \beta_j I\{z = j\}$$

for some $\beta_0, \ldots, \beta_{J-1} \geq 0$. In other words, the model stipulates that each value of the instrument encourages towards or "targets" a unique choice. Because α_j can be absorbed into ϵ_j without loss of generality, we henceforth assume $\alpha_j = 0$, and accordingly

$$U_j(z) = \beta_j I\{z = j\} \ . \tag{5}$$

We are interested in the conditions under which an observed data distribution P is consistent with the additive random utility model above because, as we demonstrate in Examples 2.1 and 2.2 below, this model can effectively encode the types of behavioral restrictions commonly imposed on treatment take-up when analyzing encouragement designs. To do so, in some cases we will further want to restrict the model so that the utilities of some choices are not affected by the instrument. Formally, for $1 \leq J_0 \leq J - 2$, define

$\mathcal{Z} = \{0, J_0, \ldots, J-1\}$ and set $\beta_j = 0$ for all $0 \le j \le J_0 - 1$. In other words, the utilities of the first J_0 choices are not affected by the instrument, and we interpret $Z = 0$ as the "base state" for the instrument that does not change the utility of any choice. As noted in Remark 3.3, such a "normalization" is not without loss of generality, but in some settings is reasonable. Note that as a notational convention, in the case where $\mathcal{Z} = \{0, 1, \ldots, J-1\}$ and no additional restrictions are placed on the coefficients β_j, we will write that $J_0 = 0$, so it is clear that all of the choices are affected by the instrument.

We now present two empirical examples where the authors impose behavioral restrictions which are naturally encoded by this additive random utility model. These examples are also studied in Lee and Salanié (2023), who focus on the identification of certain conditional treatment effects.

Example 2.1. Kline and Walters (2016) considers an RCT with a "close substitute" to study the effects of preschooling on educational outcomes. In their setting, $D \in \{0, 1, 2\}$, where $D = 0$ denotes home care (no preschool), $D = 2$ denotes a preschool program called Head Start, and $D = 1$ denotes preschools other than Head Start, namely the close substitute. Here, $Z \in \{0, 2\}$, where $Z = 2$ denotes that the household receives an offer to attend Head Start, and $Z = 0$ denoted otherwise. Assumption 2.1 holds because Z is randomly assigned. Kline and Walters (2016) impose the restriction that

$$Q\{D_2 = 2 | D_0 \neq D_2\} = 1 \ . \tag{6}$$

The condition in (6) states that if a household switches their choice upon receiving a Head Start offer, then they must choose Head Start when receiving the offer. In other words, receiving an offer to Head Start does not change the comparison between no preschool and preschools other than Head Start. Restriction (6) can be represented in the utility model by imposing that $J_0 = 2$, so that $\beta_0 = \beta_1 = 0$. To see why, note that by the definition of the random utility model,

$$D_0 \in \operatorname{argmax}\{\epsilon_0, \epsilon_1, \epsilon_2\}$$
$$D_2 \in \operatorname{argmax}\{\epsilon_0, \epsilon_1, \beta_2 + \epsilon_2\} \ .$$

Suppose $D_0 \neq D_2$. We argue by contradiction that $D_2 \neq 0$. Indeed, if $D_2 = 0$, then $\epsilon_0 > \max\{\epsilon_1, \beta_2 + \epsilon_2\}$, so $\epsilon_0 > \max\{\epsilon_1, \epsilon_2\}$, and hence $D_0 = 0 = D_2$, a contradiction to $D_0 \neq D_2$. Similarly, we can also show $D_2 \neq 1$. ■

Example 2.2. Kirkeboen et al. (2016) study the effects of fields of study on earnings. In their setting, $D \in \{0, 1, 2\}$ represent three fields of study, ordered by their (soft) admission cutoffs from the lowest to the highest. The instrument is $Z \in \{0, 1, 2\}$, with $Z = 1$ when the student crosses the (soft) admission cutoff for field 1, $Z = 2$ when the student crosses the (soft) admission cutoff for field 2, and $Z = 0$ otherwise. The authors assume that Z is exogenous in the sense that Q satisfies Assumption 2.1 and impose the following monotonicity conditions:

$$Q\{D_1 = 1 | D_0 = 1\} = 1 \ , \tag{7}$$

$$Q\{D_2 = 2 | D_0 = 2\} = 1 \ . \tag{8}$$

The conditions in (7)–(8) require that crossing the cutoff for field 1 or 2 weakly encourages them towards that field. They further impose the following "irrelevance" conditions[2]:

$$Q\{I\{D_1 = 2\} = I\{D_0 = 2\} | D_0 \neq 1, D_1 \neq 1\} = 1 \ , \tag{9}$$

$$Q\{I\{D_2 = 1\} = I\{D_0 = 1\} | D_0 \neq 2, D_2 \neq 2\} = 1 \ . \tag{10}$$

The condition in (9) states that if crossing the cutoff for field 1 does not cause the student to switch to field 1, then it does not cause them to switch to or away from field 2. A similar interpretation applies to (10). Restrictions (7)–(10) can be represented in the utility model by imposing that $J_0 = 1$, so that $\beta_0 = 0$. Here we establish (7) and (9), while (8) and (10) follow from symmetric arguments. To see (7), note that if $D_0 = 1$, then $\epsilon_1 > \max\{\epsilon_0, \epsilon_2\}$, so $\beta_1 + \epsilon_1 > \max\{\epsilon_0, \epsilon_2\}$, and hence $D_1 = 1$. To see (9), suppose $D_0 \neq 1$ and $D_1 \neq 1$. Then,

$$\beta_1 + \epsilon_1 < \max\{\epsilon_0, \epsilon_2\} \ . \tag{11}$$

Note $D_1 = 2$ if and only if $\epsilon_2 > \max\{\epsilon_0, \beta_1 + \epsilon_1\}$ and $D_0 = 2$ if and only if $\epsilon_2 > \max\{\epsilon_0, \epsilon_1\}$. Suppose $\epsilon_0 \geq \epsilon_2$. Then, (11) implies $\beta_1 + \epsilon_1 < \epsilon_0$, so $D_1 = 2 = D_0$ if and only if $\epsilon_2 > \epsilon_0$. If instead $\epsilon_0 < \epsilon_2$, then (11) implies $\epsilon_2 > \max\{\epsilon_0, \beta_1 + \epsilon_1\}$ so $D_1 = 2 = D_0$. In either case, $I\{D_1 = 2\} = I\{D_0 = 2\}$. ■

Let $\mathbf{Q}_1$ denote the set of all distributions of $((D_z : z \in \mathcal{Z}), Z)$ which are consistent with our additive random utility model. That is, such that (a) D_z is determined by (3); (b) $U_0, \ldots, U_{J-1}$ is given by (5), where $\beta_j = 0$ for $0 \leq j \leq J_0 - 1$ and $\beta_j \geq 0$ for $J_0 \leq j \leq J - 1$; (c) the distribution of $(\epsilon_0, \ldots, \epsilon_{J-1})$ is absolutely continuous with respect to the Lebesgue

[2] Kirkeboen et al. (2016) also impose the restriction that $D_0 = 0$, which they call the "next-best" condition. We conjecture that similar inequalities to those derived in Theorem 3.1 could be obtained under this additional restriction.

measure on $\mathbf{R}^J$; and (d) $(\epsilon_0, \ldots, \epsilon_{J-1}) \perp\!\!\!\perp Z$.

3 Main Results

In this section, we present our main results on sharp testable implications of the additive random utility model $\mathbf{Q}_1$. In order to do so, in Section 3.1, we first derive a set of necessary and sufficient conditions on the distribution of potential treatments implied by $\mathbf{Q}_1$. Next, in Section 3.2, we use these implications on the potential treatments to derive inequalities in terms of the conditional choice probabilities. Finally, in Section 3.3, for each P that satisfies these inequalities, we explicitly construct a distribution $Q \in \mathbf{Q}_1$, thus showing the inequalities are sharp.

3.1 Implications of $\mathbf{Q}_1$ on Potential Treatments

We first derive some implications of the model $\mathbf{Q}_1$.

Lemma 3.1. *For each $Q \in \mathbf{Q}_1$, there exists a random variable j^* such that $0 \le j^* \le J-1$ and*

$$Q\{D_j \in \{j, j^*\}\} = 1 \text{ for all } 0 \le j \le J-1 \ . \tag{12}$$

Furthermore, when $J_0 > 0$, $Q\{j^* = D_0\} = 1$, so that (12) becomes

$$Q\{D_j \in \{j, D_0\}\} = 1 \text{ for all } J_0 \le j \le J-1 \ . \tag{13}$$

The implication in (12) states that, when a distribution satisfies the restrictions imposed by $\mathbf{Q}_1$, then there exists a random choice j^* which can be interpreted as the "default" choice that an individual would choose if they do not select the choice they are encouraged to take up. When $J_0 > 0$, the default choice further coincides with D_0, the potential choice when $Z = 0$.

We now sketch the proof of Lemma 3.1 because it reveals important features of the problem. Let Ω denote the underlying probability space. For each $\omega \in \Omega$, let

$$j^*(\omega) \in \operatorname*{argmax}_{0 \le j \le J-1} \ \epsilon_j(\omega) \ . \tag{14}$$

Because the distribution of $(\epsilon_0, \ldots, \epsilon_{J-1})$ is absolutely continuous with respect to the Lebesgue measure on $\mathbf{R}^J$, j^* is unique with probability one. Therefore, j^* is a random variable. Here, $j^*(\omega)$ is the "default" choice when the instrument does not encourage towards any choice, and the value of $j^*(\omega)$ could differ across $\omega \in \Omega$. Next, we show by contradiction that with probability one, for $0 \leq j \leq J-1$, $D_j(\omega) \in \{j, j^*(\omega)\}$. Indeed, suppose $D_j(\omega) = k \notin \{j, j^*(\omega)\}$ for a set of ω with strictly positive probability. Then,

$$U_k(j) + \epsilon_k(\omega) > U_{j^*(\omega)}(j) + \epsilon_{j^*(\omega)}(\omega) \ .$$

On the other hand, if $j \neq j^*(\omega)$, then (5) and $k \notin \{j, j^*(\omega)\}$ imply that

$$U_k(j) = 0$$
$$U_{j^*(\omega)}(j) = 0 \ ,$$

so that

$$\epsilon_k(\omega) > \epsilon_{j^*(\omega)}(\omega)$$

with strictly positive probability, a contradiction to the definition of $j^*(\omega)$. If $j = j^*(\omega)$, then

$$\epsilon_k(\omega) > \beta_{j^*(\omega)} + \epsilon_{j^*(\omega)}(\omega) \geq \epsilon_{j^*(\omega)}(\omega) \ ,$$

another contradiction. Therefore, we have shown with probability one, $D_j(\omega) \in \{j, j^*(\omega)\}$. In the special case where $J_0 > 0$, $U_j(0) = 0$ for $0 \leq j \leq J-1$, so

$$D_0(\omega) \in \operatorname*{argmax}_{0 \leq j \leq J-1} \epsilon_j(\omega) \ ,$$

and therefore (14) implies that $D_0(\omega) = j^*(\omega)$. ■

Note that (12) rules out a large number of vectors of potential treatments. Consider for instance the setting where $J = 3$ and $J_0 = 0$. The restriction in (12) implies

$$Q\{D_0 = 1, D_1 = 2\} = 0 \ .$$

To see why, note if $D_0 = 1$, then since $D_0 \in \{0, j^*\}$ with probability one, $j^* = 1$, so $D_1 \in \{1, j^*\} = \{1\}$, which implies $D_1 = 1$, so $D_1 \neq 2$. Following similar arguments, we can conclude that

$$Q\{(D_0, D_1, D_2) \in \{(0,0,0), (1,1,1), (2,2,2), (0,1,0), (0,0,2), (1,1,2), (0,1,1), (2,1,2), (0,2,2), (0,1,2)\}\} = 1 \ . \quad (15)$$

We have therefore ruled out a large number of vectors for potential treatments, from $3^3 = 27$ to 10.

Remark 3.1. The restriction in (12) is distinct from the restriction considered in Bai et al. (2024b), which in the current context states

$$Q\{D_j = j | D_k = j \text{ for some } k \neq j\} = 1 \ . \tag{16}$$

The restriction (16) can be shown to be equivalent to (12) when $J = 3$, but is weaker than (12) when $J \geq 4$. Indeed, $(D_0, D_1, D_2, D_3) = (1, 1, 2, 2)$ is not ruled out by (16), but is ruled out by (12), because $D_0(\omega) = 1 \neq 0$ implies $j^*(\omega) = 1$, whereas $D_3(\omega) = 2 \neq 3$ implies $j^*(\omega) = 2$, a contradiction. ■

Remark 3.2. In a setting with a multi-valued treatment and a multi-valued instrument, Heckman and Pinto (2018) proposes a notion called "unordered monotonicity," which requires that for each $0 \leq j \leq J - 1$ and $z, z' \in \mathcal{Z}$, either $Q\{I\{D_z = j\} \leq I\{D_{z'} = j\}\} = 1$ or $Q\{I\{D_{z'} = j\} \leq I\{D_z = j\}\} = 1$. Sun (2023) derives a set of inequalities that are implied by this assumption. We note this assumption is violated by $\mathbf{Q}_1$. Indeed, suppose $J = 3$ and $J_0 = 0$. Then, $D_0 = 1$ if and only if $\epsilon_1 > \max\{\beta_0 + \epsilon_0, \epsilon_2\}$, and $D_2 = 1$ if and only if $\epsilon_1 > \max\{\epsilon_0, \beta_2 + \epsilon_2\}$. In general, neither (12) nor unordered monotonicity imply each other. ■

Remark 3.3. When $J = 2$, it is genuinely without loss of generality to normalize $\beta_0 = 0$, so that $J_0 = 1$. Indeed, $\beta_0 I\{z = 0\} + \epsilon_0 < \beta_1 I\{z = 1\} + \epsilon_1$ if and only if $\epsilon_0 < (\beta_1 + \beta_0) I\{z = 1\} + \epsilon_1 - \beta_0$, so by redefining

$$\begin{aligned}
\tilde{\epsilon}_0 &= \epsilon_0 \\
\tilde{\beta}_0 &= 0 \\
\tilde{\epsilon}_1 &= \epsilon_1 - \beta_0 \\
\tilde{\beta}_1 &= \beta_1 + \beta_0 \ ,
\end{aligned}$$

we have $D_z = I\{\tilde{\epsilon}_0 < \tilde{\beta}_1 I\{z = 1\} + \tilde{\epsilon}_1\}$. When $J > 2$, however, setting $J_0 > 0$ is no longer without loss of generality. Indeed, consider $J = 3$ as an example. If $J_0 = 1$, then $Q\{(D_0, D_1, D_2) = (0, 1, 1)\} = 0$, because $D_2 \in \{D_0, 2\}$ with probability one. On the other hand, if $J_0 = 0$, then it is possible that $Q\{(D_0, D_1, D_2) = (0, 1, 1)\} > 0$. In light of this difference, in what follows, we will make a distinction between whether $J_0 = 0$ or $J_0 > 0$. ■

Let $\mathbf{Q}_2$ denote the set of all distributions of $((D_z : z \in \mathcal{Z}), Z)$ that satisfy Assumptions 2.1–2.2 and (12). The following straightforward result shows that $\mathbf{Q}_2 = \mathbf{Q}_1$, so that each Q that satisfies the restrictions on potential treatments can be generated by the additive random utility model under consideration. The lemma extends a small fraction of Vytlacil (2002) to settings with a multi-valued treatment and a multi-valued instrument, although a crucial distinction is that the instrument is discrete in our setting, and therefore the proof is elementary. See also Lee and Salanié (2023) for a discussion of this equivalence.

Lemma 3.2. $\mathbf{Q}_2 \subseteq \mathbf{Q}_1$ *and hence* $\mathbf{Q}_2 = \mathbf{Q}_1$.

We conclude this section by revisiting Examples 2.1–2.2 to show that they satisfy the restrictions in $\mathbf{Q}_2$.

Example 3.1. Recall that in Example 2.1, $D \in \{0, 1, 2\}$, $Z \in \{0, 2\}$, and $J_0 = 2$. The restriction in (6) is equivalent to $Q\{D_2 \in \{D_0, 2\}\} = 1$, which is exactly (13). ■

Example 3.2. Recall that in Example 2.2, $D \in \{0, 1, 2\}$, $Z \in \{0, 1, 2\}$, and $J_0 = 1$. We now show that (7)–(10) implies (13), and the other direction is straightforward. Suppose by contradiction that $D_1 \notin \{D_0, 1\}$. Under this assumption, if $D_1 = 2$, then (7) implies that $D_0 \neq 1$; (9) therefore implies that $D_0 = 2$, so $D_1 = D_0$, a contradiction to $D_1 \notin \{D_0, 1\}$. If instead $D_1 = 0$, then because $D_1 \notin \{D_0, 1\}$, we know $D_0 = 2$, which by (9) implies $D_1 = 2$, a contradiction to $D_1 = 0$. Therefore, (13) is satisfied for $j = 1$. Similar arguments show it is satisfied for $j = 2$ as well. ■

3.2 Testable implications

Lemma 3.2 shows the additive random utility model $\mathbf{Q}_1$ is equivalent to the model $\mathbf{Q}_2$ defined in terms of the restrictions on potential choices. In what follows, it therefore suffices to study the model $\mathbf{Q}_2$ directly. The following theorem characterizes a set of necessary conditions in order for P to be consistent with $\mathbf{Q}_2$ and hence the additive random utility model $\mathbf{Q}_1$. Define

$$\mathcal{Z}(j) = \begin{cases} \mathcal{Z}, & \text{for } 0 \leq j \leq J_0 - 1 \\ \mathcal{Z} \backslash \{j\}, & \text{for } J_0 \leq j \leq J - 1 \ . \end{cases}$$

Theorem 3.1. *Suppose* $P = QT^{-1}$ *for* $Q \in \mathbf{Q}_2$. *Then, for* $z(j) \in \mathcal{Z}(j)$, $0 \leq j \leq J - 1$,

$$\sum_{0 \leq j \leq J-1} P\{D = j | Z = z(j)\} \leq 1 \ . \tag{17}$$

The inequalities in (17) rely crucially on the restriction in (12). To see why, first suppose $J_0 = 0$, so that $z(j) \neq j$ for $0 \leq j \leq J-1$, and consider the events

$$\{D_{z(0)} = 0\}, \ldots, \{D_{z(J-1)} = J-1\} \ .$$

Fix $\omega \in \Omega$. Because $z(j) \neq j$ for all j, $D_{z(j)}(\omega) = j$ and (12) imply that the default choice $j^*(\omega) = j$. As a result, the events listed above are disjoint across $0 \leq j \leq J-1$, so their probabilities sum up to less than one, and (17) follows. When $J_0 > 0$, $D_0(\omega) = j$ for $0 \leq j \leq J-1$ implies that $j^*(\omega) = j$, and hence $\{D_0 = j\}$ is disjoint from all other events as well. Therefore, (17) holds in addition when $z(j) = 0$ for $0 \leq j \leq J_0 - 1$. ■

Note that $z(0), \ldots, z(J-1)$ do not have to be distinct. Therefore, for $j \neq k$, by setting $z(j) = k$ and $z(\ell) = j$ for all $\ell \neq j$, we have

$$P\{D = j | Z = k\} + \sum_{\ell \neq j} P\{D = \ell | Z = j\} \leq 1 \ ,$$

which implies

$$P\{D = j | Z = k\} \leq P\{D = j | Z = j\} \ . \tag{18}$$

The inequality in (18) states that the conditional probability of choosing j is maximized at $Z = j$, which aligns with the intuition that $Z = j$ "encourages" towards $D = j$.

Example 3.3. When $J = 2$ and $J_0 = 0$, Theorem 3.1 implies only one inequality:

$$P\{D = 1 | Z = 0\} \leq P\{D = 1 | Z = 1\} \ .$$

When $J = 3$ and $J_0 = 0$, (18) leads to six inequalities:

$$\begin{aligned}
P\{D = 0 | Z = 1\} &\leq P\{D = 0 | Z = 0\} \\
P\{D = 0 | Z = 2\} &\leq P\{D = 0 | Z = 0\} \\
P\{D = 1 | Z = 2\} &\leq P\{D = 1 | Z = 1\} \\
P\{D = 1 | Z = 0\} &\leq P\{D = 1 | Z = 1\} \\
P\{D = 2 | Z = 0\} &\leq P\{D = 2 | Z = 2\} \\
P\{D = 2 | Z = 1\} &\leq P\{D = 2 | Z = 2\} \ .
\end{aligned}$$

In addition, by considering the cases where $z(0), \ldots, z(J-1)$ are all distinct in (17), we end up with two additional inequalities:

$$P\{D=1|Z=0\}+P\{D=2|Z=1\}+P\{D=0|Z=2\}\leq 1$$
$$P\{D=2|Z=0\}+P\{D=0|Z=1\}+P\{D=1|Z=2\}\leq 1\ .$$

In total, we obtain eight inequalities. For a general J, when $J_0=0$, we obtain $(J-1)^J$ inequalities. ■

When $J_0>0$, we obtain the following simplification of the inequalities:

Corollary 3.1. *Suppose $P=QT^{-1}$ for $Q\in\mathbf{Q}_2$ and $J_0>0$. Then the inequalities described by (17) are equivalent to the statement that, for $0\leq j\leq J-1$, $J_0\leq k\leq J-1$, and $j\neq k$,*

$$P\{D=j|Z=k\}\leq P\{D=j|Z=0\}\ . \tag{19}$$

To see why Corollary 3.1 holds, first note that for $0\leq j\leq J-1$, $J_0\leq k\leq J-1$, and $j\neq k$, by setting $z(j)=k$ and $z(\ell)=0$ for $\ell\neq j$ (which is allowed for all ℓ by assumption) in (17), we get

$$\sum_{0\leq j\leq J-1:j\neq k}P\{D=j|Z=0\}+P\{D=j|Z=k\}\leq 1\ ,$$

which implies (19) immediately. On the other hand, suppose (19) holds for $0\leq j\leq J-1$ and $J_0\leq k\leq J-1$. Then, for $z(j)\in\mathcal{Z}(j)$, $0\leq j\leq J-1$,

$$\sum_{0\leq j\leq J-1}P\{D=j|Z=z(j)\}\leq\sum_{0\leq j\leq J-1}P\{D=j|Z=0\}\leq 1\ ,$$

so (17) follows.

3.3 Sharpness of the Implications

We now study the converse of Theorem 3.1, namely whether for each P that satisfies (17), there exists a distribution $Q\in\mathbf{Q}_2$ such that $P=QT^{-1}$. In the special case of $J=2$, Kitagawa (2015) shows the answer is yes. The proof in Kitagawa (2015), however, does not extend to the case when $J>2$. In particular, the construction in Kitagawa (2015) relies crucially on the fact that if $D=1$ when $Z=0$, then $(D_0,D_1)=(1,1)$; this subgroup of agents are referred to as the "always takers", and the group that takes $D=0$ when $Z=1$ are called the "never takers." The remaining probability mass is then assigned to the "compliers," for whom $(D_0, D_1) = (1,1)$. When $J > 2$, however, it is in general impossible to pin down the joint distribution of $(D_0, \dots, D_J)$ using this argument, and hence the proof requires an entirely new strategy. Importantly, the proof that we present is still constructive, in that we construct Q explicitly from the given distribution P.

Theorem 3.2. *Let P be a probability distribution on $\{0, \dots, J-1\} \times \mathcal{Z}$ such that $P\{Z = z\} > 0$ for every $z \in \mathcal{Z}$. Further suppose (17) holds for all $z(j) \in \mathcal{Z}(j)$, $0 \le j \le J-1$. Then, there exists a $Q \in \mathbf{Q}_2$ such that $P = QT^{-1}$.*

To shed some light on why (17) is sufficient for determining whether P is consistent with the model $\mathbf{Q}_2$, we sketch the construction when $J = 3$ and $J_0 = 0$. Let Q^* denote a candidate distribution for which we wish to show $P = Q^* T^{-1}$ and $Q^* \in \mathbf{Q}_2$. We separately consider four classes of conditional choice probabilities under P, and assign appropriate probabilities under Q^* to corresponding events in each class.

(a) Consider $P\{D = 0 | Z = z\}$. First note that if $P = Q^* T^{-1}$, then for $z \in \{0, 1, 2\}$,

$$P\{D = 0 | Z = z\} \ge Q^*\{(D_0, D_1, D_2) = (0,0,0)\} \ .$$

Respecting this constraint, we set

$$Q^*\{(D_0, D_1, D_2) = (0,0,0)\} = \min_{z \in \mathcal{Z}} P\{D = 0 | Z = z\} \ .$$

The inequalities in (18) imply the minimum on the right-hand side is attained at $z \in \{1, 2\}$, and without loss of generality suppose it is attained at $z = 1$. Next, note any candidate Q^* has to satisfy

$$\begin{aligned} P\{D = 0 | Z = 2\} &= Q^*\{(D_0, D_1, D_2) = (0,0,0)\} + Q^*\{(D_0, D_1, D_2) = (0,1,0)\} \\ &= P\{D = 0 | Z = 1\} + Q^*\{(D_0, D_1, D_2) = (0,1,0)\} \ , \end{aligned}$$

and hence we have to define

$$Q^*\{(D_0, D_1, D_2) = (0,1,0)\} = P\{D = 0 | Z = 2\} - P\{D = 0 | Z = 1\} \ .$$

Moreover,

$$Q^*\{(D_0, D_1, D_2) = (0,0,0)\} + Q^*\{(D_0, D_1, D_2) = (0,1,0)\} = \max_{z \neq 0} P\{D = 0 | Z = z\} \ .$$

Note in particular that we have assigned no mass to $(0,0,2)$, so that

$$Q^*\{(D_0, D_1, D_2) = (0,0,2)\} = 0 \ .$$

In all the events we have considered so far, 0 is the default choice, and is chosen for at least *two* values of the instrument.

(b) Similarly, carry out the construction for events corresponding to $P\{D = 1|Z = z\}$.

(c) Similarly, carry out the construction for events corresponding to $P\{D = 2|Z = z\}$.

All of the events we have considered so far are disjoint. To see it, note the default choice is different in each class (a), (b) and (c), and it is chosen for at least two values of the instrument. For all other values of the instrument, the choice has to coincide with the instrument. Therefore, these events cannot intersect across classes. They are furthermore all disjoint from the following event:

(d) Diagonal: Note the sum of the masses that we have assigned when considering $j = 0, 1, 2$ is

$$\sum_{0 \le j \le 2} \max_{z(j) \neq j} P\{D = j|Z = z(j)\} \le 1 \ ,$$

because of (17). We then assign all of the remaining mass to

$$Q^*\{(D_0, D_1, D_2) = (0,1,2)\} = 1 - \sum_{0 \le j \le 2} \max_{z(j) \neq j} P\{D = j|Z = z(j)\} \ge 0 \ .$$

Q^* is clearly a probability measure. We now show $P = Q^* T^{-1}$. It suffices to verify $Q^*\{D_z = j\} = P\{D = j|Z = z\}$ for $j \in \{0,1,2\}$ and $z \in \{0,1,2\}$. We start by verifying that $Q^*\{D_z = 0\} = P\{D = 0|Z = z\}$ for all z. Note $D_1 = 0$ and $D_2 = 0$ is only allowed in the events in (a) above, so that

$$\begin{aligned}
Q^*\{D_1 = 0\} &= Q^*\{(D_0, D_1, D_2) = (0,0,0)\} = P\{D = 0|Z = 1\} \\
Q^*\{D_2 = 0\} &= Q^*\{(D_0, D_1, D_2) = (0,0,0)\} + Q^*\{(D_0, D_1, D_2) = (0,1,0)\} \\
&= P\{D = 0|Z = 1\} + P\{D = 0|Z = 2\} - P\{D = 0|Z = 1\} \\
&= P\{D = 0|Z = 2\} \ .
\end{aligned}$$

Next, we show $Q^*\{D_0 = 0\} = P\{D = 0|Z = 0\}$. Consider the previous construction in (b), where $j = 1$. In constructing the probabilities as above, if $P\{D = 1|Z = 0\} \ge P\{D =$

$1|Z=2\}$, we would assign

$$
\begin{aligned}
&Q^*\{(D_0,D_1,D_2)=(1,1,1)\}=P\{D=1|Z=2\}\\
&Q^*\{(D_0,D_1,D_2)=(1,1,2)\}=P\{D=1|Z=0\}-P\{D=1|Z=2\}\ .
\end{aligned}
$$

while if $P\{D=1|Z=0\}<P\{D=1|Z=2\}$, we would assign

$$
\begin{aligned}
&Q^*\{(D_0,D_1,D_2)=(1,1,1)\}=P\{D=1|Z=0\}\\
&Q^*\{(D_0,D_1,D_2)=(0,1,1)\}=P\{D=1|Z=2\}-P\{D=1|Z=0\}\ .
\end{aligned}
$$

In either case, the probability of $\{D_0\neq 0\}$ considered above in (b) is $P\{D=1|Z=0\}$. By similar arguments, the probability of $\{D_0\neq 0\}$ is $P\{D=2|Z=0\}$ in (c). Again, these probabilities correspond to disjoint events. In (a) and (d), we assign no mass to $\{D_0\neq 0\}$. Therefore,

$$
\begin{aligned}
Q^*\{D_0=0\}&=1-Q^*\{D_0\neq 0\}\\
&=1-(P\{D=1|Z=0\}+P\{D=2|Z=0\})\\
&=P\{D=0|Z=0\}\ .
\end{aligned}
$$

Similar arguments apply to $j\in\{1,2\}$, and the desired result follows. ■

4 Results with An Outcome

In this section, we present the general results when we additionally consider an outcome variable. The discussion runs mostly in parallel with the results in Section 3. Let $Y\in\mathbf{R}$ denote an observed outcome and Y_d for $0\leq d\leq J-1$ denote the potential outcome under treatment decision d. We allow Y to be continuous or discrete and denote its support[3] by $\mathcal{Y}$. In addition to (1), the observed outcome and the potential outcomes are related through

$$
Y=\sum_{0\leq d\leq J-1}Y_d I\{D=d\}\ .
$$

With some abuse of notation, we continue to let $T(\cdot)$ denote the mapping defined by the equation above together with (1). Let P denote the distribution of (Y,D,Z) and Q denote the distribution of $(Y_0,\dots,Y_{J-1},(D_z:z\in\mathcal{Z}),Z)$. We modify Assumption 2.1 to include

[3]Following pp.73–74 of Lifshits (1995), we define the (topological) support of Y as the smallest closed set with probability one under P, i.e., $\mathcal{Y}:=\bigcap\{F\subseteq\mathbf{R}:F\text{ closed },P\{Y\in F\}=1\}$.

the potential outcomes:

Assumption 4.1. $(Y_0, \ldots, Y_{J-1}, (D_z : z \in \mathcal{Z}), Z) \perp\!\!\!\perp Z$ under Q.

Let $\mathbf{Q}_1^Y$ denote the set of all distributions Q for which Assumptions 4.1 and 2.2 as well as (12) hold. We first present the counterpart to Theorems 3.1 and 3.2. To do so,

Theorem 4.1. *Let P be a probability distribution on $\mathcal{Y} \times \{0, \ldots, J-1\} \times \mathcal{Z}$ such that $P\{Z = z\} > 0$ for every $z \in \mathcal{Z}$. Then, $P = QT^{-1}$ for $Q \in \mathbf{Q}_1^Y$ if and only if both of the following sets of conditions hold:*

(a) *If for each $0 \le j \le J-1$, the Borel sets $\{B_z(j) : z \in \mathcal{Z}(j)\}$ form a partition of $\mathcal{Y}$, then*

$$\sum_{0 \le j \le J-1} \sum_{z \in \mathcal{Z}(j)} P\{Y \in B_z(j), D = j | Z = z\} \le 1 \ . \tag{20}$$

(b) *For each Borel set $B \subseteq \mathcal{Y}$, $J_0 \le j \le J-1$, and $k \ne j$,*

$$P\{Y \in B, D = j | Z = k\} \le P\{Y \in B, D = j | Z = j\} \ . \tag{21}$$

For $z(j) \in \mathcal{Z}(j)$, $0 \le j \le J-1$, note that by taking $B_{z(j)}(j) = \mathcal{Y}$ and $B_z(j) = \emptyset$ for $0 \le j \le J-1$ and $z \ne z(j)$ in (20), we recover (17).

As in Section 3.2, when $J_0 > 0$, we obtain the following simplification of the inequalities:

Corollary 4.1. *$P \in \mathbf{Q}_1^Y T^{-1}$ if and only if Theorem 4.1(b) holds and for all Borel sets $B \subseteq \mathcal{Y}$, $0 \le j \le J-1$, $J_0 \le k \le J-1$, and $j \ne k$,*

$$P\{Y \in B, D = j | Z = k\} \le P\{Y \in B, D = j | Z = 0\} \ . \tag{22}$$

To see why Corollary 4.1 holds, first note for $0 \le j \le J-1$ and $J_0 \le k \le J-1$, by taking $B_k(j) = B$, $B_0(j) = \mathcal{Y} \backslash B$, and $B_0(\ell) = \mathcal{Y}$ and $B_z(\ell) = \emptyset$ for $\ell \ne j$ and $z \ne 0$, we get

$$P\{Y \in B, D = j | Z = k\} + P\{Y \notin B, D = j | Z = 0\} + \sum_{\ell \ne j} P\{Y \in \mathcal{Y}, D = \ell | Z = 0\} \le 1 \ ,$$

from which we immediately obtain (22). On the other hand, suppose (22) holds for all $0 \le j \le J-1$, $J_0 \le k \le J-1$, and $j \ne k$, and for $0 \le j \le J-1$, Borel sets $\{B_z(j) : z \in \mathcal{Z}(j)\}$ form a partition of $\mathcal{Y}$. Then, (20) holds because

$$\sum_{0 \le j \le J-1} \sum_{z \in \mathcal{Z}(j)} P\{Y \in B_z(j), D = j | Z = z\} \le \sum_{0 \le j \le J-1} \sum_{z \in \mathcal{Z}(j)} P\{Y \in B_z(j), D = j | Z = 0\} = 1 \ .$$

Remark 4.1. When $J = 2$, the only inequalities implied by Theorem 4.1 are

$$P\{Y \in B, D = 1|Z = 0\} \leq P\{Y \in B, D = 1|Z = 1\}$$
$$P\{Y \in B, D = 0|Z = 1\} \leq P\{Y \in B, D = 0|Z = 0\} \ .$$

Note that these inequalities coincide with the simplification described in Corollary 4.1 when $J_0 = 1$; this demonstrates that the normalization discussed in Remark 3.3 is innocuous in the case where $J = 2$. Furthermore, these inequalities are exactly those derived by Balke and Pearl (1997a,b) and Kitagawa (2015). Theorem 4.1 could therefore be thought of as a counterpart to their results when both the treatment and instrument can take more than two values and are unordered. ■

Remark 4.2. Using the characterizations in Theorem 4.1 or Corollary 4.1, formal tests could be constructed to assess the validity of the model. For instance, if the outcome variable Y is discrete (or is discretized ex-ante), then Theorem 4.1 and Corollary 4.1 generate a finite number of inequalities, which can be tested using any off-the-shelf inference method for a finite number of moment inequalities. See, for instance, Canay and Shaikh (2017) for an overview. Here we sketch how we can convert the problem into testing the feasibility of a linear program, so that we can directly apply recent results in, for instance, Fang et al. (2023). Suppose $\mathcal{Y}$ is discrete and let p denote the vector of $(P\{Y = y, D = j|Z = z\} : y \in \mathcal{Y}, 0 \leq j \leq J - 1, z \in \mathcal{Z})$. We can represent the inequalities in Theorem 4.1 and Corollary 4.1 as

$$\Gamma p - \gamma \leq 0 \ , \tag{23}$$

where Γ and γ have known entries which lie in $\{-1, 0, 1\}$. With a vector of slack variables x, (23) is equivalent to

$$Ax = \beta(P)$$
$$x \geq 0 \ ,$$

where A is the identity matrix and $\beta(P) = \gamma - \Gamma p$. This formulation maps exactly into the notation of Fang et al. (2023), and their tests apply immediately.

If instead the outcome variable is continuous, then we could develop a test based on the K-S statistic in Kitagawa (2015). Alternatively, the inequalities in Theorem 4.1 and Corollary 4.1 could be transformed, as in Mourifié and Wan (2017), into a finite number of conditional moment inequalities where the conditioning variable is Y instead of Z. The resulting inequalities could then be tested using any off-the-shelf inference method for a finite number of conditional moment inequalities. See, for instance, Andrews and Shi (2013), Chernozhukov et al. (2013), Armstrong (2014, 2015, 2018), and Chetverikov (2018). ■

A Proofs of Main Results

A.1 Proof of Lemma 3.1

The proof has been included in the main text and is omitted here. ■

A.2 Proof of Lemma 3.2

First suppose $J_0 = 0$. For $0 \leq j \leq J-1$, define

$$\beta_j = \begin{cases} 1 & \text{if } Q\{D_j = j, D_k \neq j \text{ for some } k \neq j\} > 0 \\ 0 & \text{otherwise .} \end{cases} \tag{24}$$

The restriction in (12) can equivalently be expressed as requiring the probabilities of some vectors of potential treatments are zero. In particular, define

$$\mathcal{S} = \{(d_0, \ldots, d_{J-1}) : d_j \neq j, d_k \neq k \implies d_j = d_k\} \ .$$

The restriction in (12) can then be expressed as $Q\{(D_0, \ldots, D_{J-1}) = (d_0, \ldots, d_{J-1})\} = 0$ if $(d_0, \ldots, d_{J-1}) \notin \mathcal{S}$. Because the default choice j^* only depends on the value of $(D_0, \ldots, D_{J-1})$, for $(d_0, \ldots, d_{J-1}) \in \mathcal{S} \backslash \{(0, \ldots, J-1)\}$, we define $j^*(d_0, \ldots, d_{J-1})$ as the associated default choice, which is the value of d_j as long as $d_j \neq j$. Further define

$$\Lambda(d_0, \ldots, d_{J-1}) = \{0 \leq j \leq J-1 : d_j = j \neq j^*(d_0, \ldots, d_{J-1})\} \ .$$

Let M be a constant to be chosen below. Define the corresponding region for $(\epsilon_0, \epsilon_1, \ldots, \epsilon_{J-1})$ as

$$\begin{aligned} R(d_0, \ldots, d_{J-1}) = \{(\epsilon_0, \ldots, \epsilon_{J-1}) : \beta_{j^*} + \epsilon_{j^*} > \epsilon_j \text{ for } j \neq j^*, \beta_j + \epsilon_j > \epsilon_k \text{ for } j \in \Lambda \\ \text{and } k \neq j, \beta_j + \epsilon_j < \epsilon_{j^*} \text{ for } j \notin \Lambda, |\epsilon_j| \leq M \text{ for all } j\} \subseteq \mathbf{R}^J \ . \end{aligned}$$

Finally, define

$$R(0, \ldots, J-1) = \{\beta_j + \epsilon_j > \epsilon_k \text{ for } 0 \leq j \leq J-1 \text{ and } k \neq j, |\epsilon_j| \leq M \text{ for all } j\} \ .$$

Such regions are obviously disjoint across all possible $(d_0, \ldots, d_{J-1}) \in \mathcal{S}$. By choosing M large enough, all of these regions are nonempty but bounded. Let $(\epsilon_0, \ldots, \epsilon_{J-1})$ be uniformly distributed in each region, with density

$$\frac{Q\{(D_0, \dots, D_{J-1}) = (d_0, \dots, d_{J-1})\}}{|R(d_0, \dots, d_{J-1})|}$$

where $|R(d_0, \dots, d_{J-1})|$ is the Lebesgue measure of $R(d_0, \dots, d_{J-1})$ in $\mathbf{R}^J$. The result now follows. When $J_0 > 0$, the proof is similar, with the only difference being we require $\beta_j = 0$ for $0 \leq j \leq J_0 - 1$. ■

A.3 Proof of Theorem 3.1

Proof of Theorem 3.1. First suppose $J_0 = 0$. To establish the inequalities, fix $z(0), \dots, z(J-1) \in \{0, \dots, J-1\}$ such that $z(j) \neq j$ for $0 \leq j \leq J-1$. By construction, $z(j) \in \mathcal{Z}(j)$ for $0 \leq j \leq J-1$. For a fixed j, if $D_{z(j)}(\omega) = j$, then because $j \neq z(j)$, we have that

$$j^*(\omega) = j \ .$$

As a result, $D_{z(k)} \in \{j, z(k)\}$ for $k \neq j$. Because $k \neq z(k)$ and $k \neq j$, it cannot be the case that $D_{z(k)} = k$. Therefore,

$$\{D_{z(0)} = 0\}, \dots, \{D_{z(J-1)} = J-1\}$$

are mutually exclusive events, so

$$\sum_{0 \leq j \leq J-1} Q\{D_{z(j)} = j\} \leq 1 \ .$$

The desired conclusion now follows from Assumption 2.1. When $J_0 > 0$, $\{D_0 = j\}$ for $0 \leq j \leq J_0 - 1$ is disjoint from all events above, and the result follows. ■

A.4 Proof of Theorem 3.2

A.4.1 Proof when $J_0 = 0$

We construct Q^* that satisfies Assumption 2.1 and (12). Fix $0 \leq j \leq J-1$. Let $\{z_1(j), \dots, z_J(j)\} = \{0, \dots, J-1\}$ be such that

$$P\{D = j | Z = z_1(j)\} \leq P\{D = j | Z = z_2(j)\} \leq \cdots \leq P\{D = j | Z = z_J(j)\} \ .$$

Note that (18) implies $P\{D = j|Z = z\}$ is maximized by $z = j$, so $z_J(j) = j$ (in the case of ties, simply define $z_J(j) = j$). Our construction for this fixed j consists of the following steps:

Step 1: define

$$Q^*\{D_z = j \text{ for } 0 \leq z \leq J-1\} = P\{D = j|Z = z_1(j)\} . \tag{25}$$

Step ℓ for $2 \leq \ell \leq J-1$ (note we stop at step $J-1$ instead of J): define

$$\begin{aligned} Q^*\{D_{z_1(j)} = z_1(j), \ldots, D_{z_{\ell-1}(j)} = z_{\ell-1}(j), D_z = j \text{ for } z \notin \{z_1(j) \ldots, z_{\ell-1}(j)\}\} \\ = P\{D = j|Z = z_\ell(j)\} - P\{D = j|Z = z_{\ell-1}(j)\} . \end{aligned} \tag{26}$$

After carrying out the construction for each $0 \leq j \leq J-1$, define the "diagonal" as

$$Q^*\{D_j = j \text{ for } 0 \leq j \leq J-1\} = 1 - \sum_{0 \leq j \leq J-1} P\{D = j|Z = z_{J-1}(j)\} , \tag{27}$$

which is nonnegative because of (17) and the fact that $z_{J-1}(j) \neq j$. (25)–(26) completely specify the probability of potential treatments where j appears at least twice, and (27) closes the gap by specifying the probability that one complies to all values of the instrument. For all other vectors $(d_0, \ldots, d_{J-1})$, define $Q^*\{(D_0, \ldots, D_{J-1}) = (d_0, \ldots, d_{J-1})\} = 0$

Finally, for each z and $(d_0, \ldots, d_{J-1})$, define

$$Q^*\{D_0 = d_0, \ldots, D_{J-1} = d_{J-1}, Z = z\} = Q^*\{D_0 = d_0, \ldots, D_{J-1} = d_{J-1}\}P\{Z = z\} . \tag{28}$$

We now verify that Q^* satisfies Assumption 2.1 and (12). First note Assumption 2.1 holds by (28). All probabilities are nonnegative by construction. To verify that Q^* is a probability measure, note for each j, the events in (25)–(26) are mutually exclusive from each other. In addition, they are mutually exclusive across $0 \leq j \leq J-1$ and are all exclusive from the event in (27) because for the events that appear in (25)–(26), j will be selected for at least two values of the instrument, and the treatment equals the instrument for all other values of the instrument. Furthermore, for each j, the sum of (25)–(26) from $1 \leq \ell \leq J-1$ is

$$P\{D = j|Z = z_{J-1}(j)\} .$$

Therefore, across $0 \le j \le J-1$, (25)–(27) sum up to

$$\sum_{0 \le j \le J-1} P\{D = j | Z = z_{J-1}(j)\} + 1 - \sum_{0 \le j \le J-1} P\{D = j | Z = z_{J-1}(j)\} = 1 \ .$$

As a result, Q^* is a probability measure. By construction, Q^* only assigns zero probability to all events that are ruled out by (12), so (12) holds for Q^*.

To conclude the proof, we verify that $P = Q^* T^{-1}$, i.e., (2) holds. In order to do so, note if $z \neq j$, then $z = z_\ell(j)$ for some $1 \le \ell \le J-1$, and

$$\begin{aligned}
&Q^*\{D_{z_\ell(j)} = j\} \\
&= Q^*\{D_z = j \text{ for } 0 \le k \le J-1\} + \dots \\
&\qquad + Q^*\{D_{z_1(j)} = z_1(j), \dots, D_{z_{\ell-1}(j)} = z_{\ell-1}(j), D_z = j \text{ for all } z \notin \{z_1(j) \dots, z_{\ell-1}(j)\}\} \\
&= P\{D = j | Z = z_1(j)\} + P\{D = j | Z = z_2(j)\} - P\{D = j | Z = z_1(j)\} + \dots \\
&\qquad + P\{D = j | Z = z_\ell(j)\} - P\{D = j | Z = z_{\ell-1}(j)\} \\
&= P\{D = j | Z = z_\ell(j)\} \ .
\end{aligned}$$

Therefore, we have verified $Q^*\{D_z = j\} = P\{D = j | Z = z\}$ when $z \neq j$. It therefore suffices to verify $Q^*\{D_k = k\} = P\{D = k | Z = k\}$ for $0 \le k \le J-1$. To do so, note for $0 \le k \le J-1$,

$$\begin{aligned}
Q^*\{D_k \neq k\} &= \sum_{0 \le j \le J-1 : j \neq k} Q^*\{D_k = j\} \\
&= \sum_{0 \le j \le J-1 : j \neq k} P\{D = j | Z = k\} = 1 - P\{D = k | Z = k\} \ . \qquad (29)
\end{aligned}$$

Because Q^* is a probability measure, $Q^*\{D_k = k\} = 1 - Q^*\{D_k \neq k\} = P\{D = k | Z = k\}$, and the proof is completed. ■

A.4.2 Proof when $J_0 > 0$

The construction is similar to that in the proof of Theorem 3.2, with a few important changes:

(a) For $J_0 \le j \le J-1$, (19) implies when ordering $P\{D = j | Z = z\}$ as in the proof of Theorem 3.2, $\{z_1(j), \dots, z_{J-J_0-1}(j)\} = \{J_0, \dots, J-1\} \backslash \{j\}$ and $z_{J-J_0-1}(j) = 0$. We can therefore carry out the construction as in there, but because z can only take $J - J_0$

values, the construction will stop at step $J - J_0 - 1$ instead of $J - 1$. In this part of the construction, because $z_{J-J_0-1}(j) = 0$, the total mass assigned by Q^* in this part is

$$P\{D = j|Z = z_{J-J_0-1}(j)\} = P\{D = j|Z = 0\} .$$

Note in this part, no mass is assigned to $\{D_0 = k\}$ for any $0 \leq k \leq J_0 - 1$.

(b) For $0 \leq j \leq J_0 - 1$, carry out the construction as above, and again stop at step $J - J_0 - 1$ instead of $J - 1$. Note $z_{J-J_0}(j) = 0$. The total mass assigned by Q^* in this part is

$$P\{D = j|Z = z_{J-J_0-1}(j)\} = \max_{z \neq 0} P\{D = j|Z = z\} .$$

(c) The diagonal now consists of a set of vectors of potential treatments instead of one. For $0 \leq j \leq J_0 - 1$, we simply define

$$Q^*\{(D_0, D_{J_0}, \ldots, D_{J-1}) = (j, J_0 \ldots, J-1)\} = P\{D = j|Z = 0\} - \max_{z \neq 0} P\{D = j|Z = z\} ,$$

which is positive by (19).

By construction, Q^* is obviously a probability measure and satisfies (13). Next, we verify $P = QT^{-1}$. First consider $0 \leq j \leq J_0 - 1$. Note (a) assigns no mass to $\{D_0 = j\}$ while (b) and (c) assign

$$\begin{aligned} Q^*\{D_0 = j\} &= \max_{z \neq 0} P\{D = j|Z = z\} + P\{D = j|Z = 0\} - \max_{z \neq 0} P\{D = j|Z = z\} \\ &= P\{D = j|Z = 0\} . \end{aligned}$$

On the other hand, for $z \neq 0$, $\{D_z = j\}$ is assigned mass only in (b), which by the same arguments as in the proof of Theorem 3.2 equals

$$Q^*\{D_z = j\} = P\{D = j|Z = z\} .$$

Next, consider $J_0 \leq j \leq J - 1$. Again for $z \notin \{0, j\}$, $\{D_z = j\}$ is assigned mass only in (b), and hence as in the proof of Theorem 3.2,

$$Q^*\{D_z = j\} = P\{D = j|Z = z\} .$$

On the other hand, for $z = 0$, $\{D_0 = j\}$ is only assigned mass in (a), and once again

$$Q^*\{D_0 = j\} = P\{D = j|Z = 0\} \ .$$

It suffices to verify $Q^*\{D_k = k\} = P\{D = k|Z = k\}$ for $J_0 \le k \le J - 1$. Note $\{D_k \ne k\}$ is assigned mass only in (a) and (b) for $j \ne k$, and

$$Q^*\{D_k \ne k\} = \sum_{0 \le j \le J-1 : j \ne k} P\{D = j|Z = k\} = P\{D \ne k|Z = k\} \ ,$$

so

$$Q^*\{D_k = k\} = 1 - Q^*\{D_k \ne k\} = 1 - P\{D \ne k|Z = k\} = P\{D = k|Z = k\} \ .$$

As a result, $P = Q^* T^{-1}$. ■

A.5 Proof of Theorem 4.1

When $J_0 = 0$, because for each $0 \le j \le J - 1$, $\{B_z(j) : z \ne j\}$ forms a partition of $\mathcal{Y}$,

$$(\{Y_j \in B_z(j)\} : 0 \le j \le J - 1, z \ne j)$$

are mutually exclusive. The inequalities in (20) now follow. (21) is also obvious because $D_k = j$ implies $D_j = j$ by (12). When $J_0 > 0$, the results also follow easily.

We now prove the converse when $J_0 = 0$ and note the result for $J_0 > 0$ can be proved by combining the proof of Theorem 3.2 and the arguments below (and defining $\lambda_j(y) = 1$ for $0 \le j \le J_0 - 1$ and following (30) for $J_0 \le j \le J$). To begin, note there exists a common dominating measure ν on $\mathbf{R}$ such that for $z \in \mathcal{Z}$, $P\{Y \in B|Z = z\}$ is absolutely continuous with respect to ν. Indeed, we can simply define ν as $\sum_{z \in \mathcal{Z}} P\{Y \in B|Z = z\}$. Let $p_j(\cdot|z)$ denote the Radon-Nikodym derivative of $P\{Y \in B, D = j|Z = z\}$ with respect to ν. For each $y \in \mathcal{Y}$, let $\{z_1(y, j), \ldots, z_J(y, j)\} = \{0, \ldots, J - 1\}$ be such that

$$p_j(y|z_1(y, j)) \le \cdots \le p_j(y|z_J(y, j)) \ .$$

Note that (21) implies $p_j(y|z)$ is maximized by $z = j$ almost everywhere, so we can take $z_J(y, j) = j$ everywhere (in the case of ties, simply define $z_J(y, j) = j$). For $0 \le j \le J-1$ and each $y \in \mathcal{Y}$, define $\lambda_j(y)$ to be an arbitrary nonnegative function such that $\int_{\mathcal{Y}} \lambda_j(y) d\nu(y) = 1$

if $\int_{\mathcal{Y}}(p_j(y|j) - p_j(y|z_{J-1}(y,j))d\nu(y) = 0$, and otherwise define

$$\lambda_j(y) = \frac{p_j(y|j) - p_j(y|z_{J-1}(y,j))}{\displaystyle\int_{\mathcal{Y}}(p_j(y|j) - p_j(y|z_{J-1}(y,j))d\nu(y)} \ . \tag{30}$$

Note by definition that $\int_{\mathcal{Y}} \lambda_j(y)d\nu(y) = 1$ for $0 \leq j \leq J-1$.

We seek construct a probability measure Q^* that satisfies Assumption 4.1 and 2.2 as well as (12), and such that $P = Q^* T^{-1}$. For each $(d_0, \ldots, d_{J-1}) \in \{0, \ldots, J-1\}^J$ and Borel sets $B_0, \ldots, B_{J-1} \subseteq \mathcal{Y}$, we will use $Q^*_{(d_0,\ldots,d_{J-1})}(B_0, \ldots, B_{J-1})$ to denote $Q^*\{Y_0 \in B_0, \ldots, Y_{J-1} \in B_{J-1}, D_0 = d_0, \ldots D_{J-1} = d_{J-1}\}$. Let $\nu^J = \nu \times \cdots \times \nu$ denote the product measure on $\mathbb{R}^J$ with each marginal equal to ν. In what follows, we will construct nonnegative functions $q^*_{(d_0,\ldots,d_{J-1})}$ and for Borel sets $B_0, \ldots, B_{J-1} \subseteq \mathcal{Y}$, we will define

$$Q^*_{(d_0,\ldots,d_{J-1})}(B_0 \times \cdots \times B_{J-1}) = \int_{B_0} \cdots \int_{B_{J-1}} q^*_{(d_0,\ldots,d_{J-1})}(y_0, \ldots, y_{J-1})d\nu(y_0)\ldots d\nu(y_{J-1}) \ .$$

Finally, as in the proof of Theorem 3.2, we will define the joint distribution Q^* by multiplying the marginal distribution of Z.

Our construction for the each $y \in \mathcal{Y}$ and $0 \leq j \leq J-1$ consists of the following steps:

Step 1: for each $(y_k \in \mathcal{Y} : k \neq j)$, define

$$q^*_{(j,\ldots,j)}(y_0, \ldots, y, \ldots, y_{J-1}) = \prod_{k \neq j} \lambda_k(y_k) p_j(y|z_1(y,j)) \ .$$

Step ℓ for $2 \leq \ell \leq J-1$: for each $(y_k \in \mathcal{Y} : k \neq j)$, and for $d_{z_1(y,j)} = z_1(y,j), \ldots, d_{z_{\ell-1}(y,j)} = z_{\ell-1}(y,j)$ and $d_j = j$ for $z \notin \{z_1(y,j) \ldots, z_{\ell-1}(y,j)\}$, define

$$q^*_{(d_0,\ldots,d_{J-1})}(y_0, \ldots, y, \ldots, y_{J-1}) = \prod_{k \neq j} \lambda_k(y_k)(p_j(y|z_\ell(y,j)) - p_j(y|z_{\ell-1}(y,j))) \ .$$

Carry out these steps separately for $0 \leq j \leq J-1$. Finally, for each $(y_j \in \mathcal{Y} : 0 \leq j \leq J-1)$, define

$$q^*_{(0,\ldots,J-1)}(y_0, \ldots, y_{J-1}) = \prod_{0 \leq j \leq J-1} \lambda_j(y_j) \left(1 - \sum_{0 \leq j \leq J-1} \int_{\mathcal{Y}} p_j(y|z_{J-1}(y,j))\right) \ ,$$

which is nonnegative because of (20). To see it, take $B_z(j) = \{y \in \mathcal{Y} : z_{J-1}(y,j) = z\}$ for $0 \leq j \leq J-1$ and $j \neq z$. As defined, for $0 \leq j \leq J-1$, $B_z(j)$ is measurable as a function of a finite number of functions $p_j(y|0), \ldots, p_j(y|J-1)$, and $\{B_z(j) : z \neq j\}$ form a partition of $\mathcal{Y}$. Therefore,

$$\sum_{0 \le j \le J-1} \int_{\mathcal{Y}} p_j(y|z_{J-1}(y,j)) = \sum_{0 \le j \le J-1} \sum_{0 \le z \le J-1} P\{Y \in B_z(j), D = j | Z = z\} \le 1 \ .$$

Following the arguments in the proof of Theorem 3.2, it is easy to verify Q^* is a probability measure. Next, we verify $P = Q^* T^{-1}$. Fix a Borel subset $B \subseteq \mathcal{Y}$. Integrating over all possible values $y_k \in \mathcal{Y}$ for $k \neq j$, we get

$$\begin{aligned} & Q^*\{Y_j \in B, D_z = j \text{ for } 0 \le z \le J-1\} \\ &= \int_{\mathcal{Y} \times \cdots \times B \times \cdots \times \mathcal{Y}} q^*_{(j,\ldots,j)}(y_0, \ldots, y_{J-1}) d\nu(y_0) \ldots d\nu(y_{J-1}) \\ &= \int_B p_j(y|z_1(y,j)) d\nu(y) \end{aligned}$$

where the second equality follow because $\int_{\mathcal{Y}} \lambda_j(y) d\nu(y) = 1$ for $0 \le j \le J-1$. Similarly,

$$\begin{aligned} & Q^*\{Y_j \in B, D_{z_1(y,j)} = z_1(y,j), \ldots, D_{z_{\ell-1}(y,j)} = z_{\ell-1}(y,j), \\ & \qquad D_z = j \text{ for } z \notin \{z_1(y,j) \ldots, z_{\ell-1}(y,j)\}\} \\ &= \int_B (p_j(y|z_\ell(y,j)) - p_j(y|z_{\ell-1}(y,j))) d\nu(y) \end{aligned}$$

It then follows from similar arguments as in the proof of Theorem 3.2 that if $z \neq j$, then

$$Q^*\{Y_j \in B, D_z = j\} = \int_B p_j(y|z) d\nu(y) = P\{Y \in B, D = j | Z = z\} \ .$$

It remains to verify for each Borel set $B \subseteq \mathcal{Y}$ and $0 \le k \le J-1$,

$$Q^*\{Y_k \in B, D_k = k\} = P\{Y \in B, D = k | Z = k\} \ . \tag{31}$$

In order to do so, integrating over all possible values of $y_{j'} \in \mathcal{Y}$ for $j' \notin \{j, k\}$ in each step, we have for $j \neq k$,

$$Q^*\{Y_k \in B, D_z = j \text{ for } 0 \le z \le J-1\} = \int_{y_k \in B} \int_{y \in \mathcal{Y}} \lambda_k(y_k) p_j(y|z_1(y,j)) d\nu(y) d\nu(y_k)$$

and

$$Q^*\{Y_j \in \mathcal{Y}, Y_k \in B, D_{z_1(y,j)} = z_1(y,j), \ldots, D_{z_{\ell-1}(y,j)} = z_{\ell-1}(y,j),$$

$$D_z = j \text{ for } z \notin \{z_1(y,j)\ldots,z_{\ell-1}(y,j)\}\}$$
$$= \int_{y_k \in B} \int_{y \in \mathcal{Y}} \lambda_k(y_k)(p_j(y|z_\ell(y,j)) - p_j(y|z_{\ell-1}(y,j))))d\nu(y)d\nu(y_k) \ .$$

Moreover,

$$Q^*\{Y_k \in B, D_z = k \text{ for } 0 \le z \le J-1\} = \int_{y_k \in B} p_j(y_k|z_1(y_k,k))d\nu(y_k)$$

and

$$Q^*\{Y_k \in B, D_{z_1(y_k,j)} = z_1(y_k,j), \ldots, D_{z_{\ell-1}(y_k,j)} = z_{\ell-1}(y_k,j),$$
$$D_z = k \text{ for } z \notin \{z_1(y_k,j)\ldots,z_{\ell-1}(y_k,j)\}\}$$
$$= \int_{y_k \in B} (p_k(y_k|z_\ell(y_k,j)) - p_k(y_k|z_{\ell-1}(y_k,j)))d\nu(y_k) \ .$$

Finally,

$$Q^*\{Y_k \in B, D_j = j \text{ for } 0 \le j \le J-1\}$$
$$= \int_{y_k \in B} \lambda_k(y_k) \left(1 - \sum_{0 \le j \le J-1} \int_{\mathcal{Y}} p_j(y|z_{J-1}(y,j))d\nu(y)\right) d\nu(y_k) \ ,$$

Summing the previous displays, we have

$$Q^*\{Y_k \in B, D_k = k\} = \int_{y_k \in B} q_k(y_k)d\nu(y_k) \ , \tag{32}$$

where

$$\begin{aligned}
&q_k(y_k) \\
&= \sum_{j \neq k} \int_{y \in \mathcal{Y}} \lambda_k(y_k)(p_j(y|z_{J-1}(y,j)) - p_j(y|k))d\nu(y) \\
&\qquad + \lambda_k(y_k)\left(1 - \sum_{0 \le j \le J-1} \int_{y \in \mathcal{Y}} p_j(y|z_{J-1}(y,j))d\nu(y)\right) \\
&\qquad + p_k(y_k|z_{J-1}(y_k,k)) \\
&= \lambda_k(y_k) \int_{y \in \mathcal{Y}} \left(1 - p_k(y|z_{J-1}(y,k)) - \sum_{j \neq k} p_j(y|k)\right) d\nu(y) + p_k(y_k|z_{J-1}(y_k,k)) \\
&= \lambda_k(y_k) \int_{y \in \mathcal{Y}} \left(p_k(y|k) - p_k(y|z_{J-1}(y,k))\right) d\nu(y) + p_k(y_k|z_{J-1}(y_k,k))
\end{aligned}$$

$$= p_k(y_k|k) \ ,$$

where the third equality follows because

$$\int_{\mathcal{Y}} \left(\sum_{j \neq k} p_j(y|k) + p_k(y|k) \right) d\nu(y) = 1 \ ,$$

and the last equality holds regardless of whether or not denominator in the definition of $\lambda_k(y_k)$ is zero. Indeed, if it is not zero, then the equality holds because of the definition of $\lambda_k(y_k)$ in (30); if is zero, then $p_k(y_k|k) = p_k(y_k|z_{J-1}(y_k, k))$, so the equality still holds. The desired result in (31) now follows from (32), and the proof is complete. ■

References

Andrews, D. W. K. and Shi, X. (2013). Inference Based on Conditional Moment Inequalities. *Econometrica*, **81** 609–666.

Armstrong, T. B. (2014). Weighted KS statistics for inference on conditional moment inequalities. *Journal of Econometrics*, **181** 92–116.

Armstrong, T. B. (2015). Asymptotically exact inference in conditional moment inequality models. *Journal of Econometrics*, **186** 51–65.

Armstrong, T. B. (2018). On the choice of test statistic for conditional moment inequalities. *Journal of Econometrics*, **203** 241–255.

Bai, Y., Huang, S., Moon, S., Santos, A., Shaikh, A. M. and Vytlacil, E. J. (2024a). Inference for Treatment Effects Conditional on Generalized Principal Strata using Instrumental Variables.

Bai, Y., Huang, S., Moon, S., Shaikh, A. M. and Vytlacil, E. J. (2024b). On the Identifying Power of Monotonicity for Average Treatment Effects. ArXiv:2405.14104 [econ, stat], URL http://arxiv.org/abs/2405.14104.

Balke, A. and Pearl, J. (1997a). Bounds on Treatment Effects from Studies with Imperfect Compliance. *Journal of the American Statistical Association*, **92** 1171–1176.

Balke, A. A. and Pearl, J. (1997b). Probabilistic Counterfactuals: Semantics, Computation, and Applications. Tech. rep., DTIC Document.

Bhattacharya, J., Shaikh, A. M. and Vytlacil, E. (2008). Treatment Effect Bounds under Monotonicity Assumptions: An Application to Swan-Ganz Catheterization. *American Economic Review*, **98** 351–356.

Bhattacharya, J., Shaikh, A. M. and Vytlacil, E. (2012). Treatment effect bounds: An application to Swan–Ganz catheterization. *Journal of Econometrics*, **168** 223–243.

Bhuller, M. and Sigstad, H. (2024). 2sls with multiple treatments. *Journal of Econometrics*, **242** 105785.

Canay, I. A. and Shaikh, A. M. (2017). Practical and Theoretical Advances in Inference for Partially Identified Models. In *Advances in Economics and Econometrics: Eleventh World Congress* (A. Pakes, B. Honoré, L. Samuelson and M. Piazzesi, eds.), vol. 2 of *Econometric Society Monographs*. Cambridge University Press, Cambridge, 271–306.

Chernozhukov, V., Lee, S. and Rosen, A. M. (2013). Intersection Bounds: Estimation and Inference. *Econometrica*, **81** 667–737.

Chetverikov, D. (2018). Adaptive tests of conditional moment inequalities. *Econometric Theory*, **34** 186–227.

Fang, Z., Santos, A., Shaikh, A. M. and Torgovitsky, A. (2023). Inference for Large-Scale Linear Systems With Known Coefficients. *Econometrica*, **91** 299–327.

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

Heckman, J. J. and Vytlacil, E. (2005). Structural Equations, Treatment Effects, and Econometric Policy Evaluation. *Econometrica*, **73** 669–738.

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

Kamat, V., Norris, S. and Pecenco, M. (2023). Identification in multiple treatment models under discrete variation. *arXiv preprint arXiv:2307.06174*.

Kirkeboen, L. J., Leuven, E. and Mogstad, M. (2016). Field of study, earnings, and self-selection. *The Quarterly Journal of Economics*, **131** 1057–1111.

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

Kline, P. and Walters, C. R. (2016). Evaluating public programs with close substitutes: The case of head start. *The Quarterly Journal of Economics*, **131** 1795–1848.

Kwon, S. and Roth, J. (2024). Testing mechanisms. 2404.11739, URL https://arxiv.org/abs/2404.11739.

Kédagni, D. and Mourifié, I. (2020). Generalized instrumental inequalities: testing the instrumental variable independence assumption. *Biometrika*, **107** 661–675.

Lee, S. and Salanié, B. (2018). Identifying Effects of Multivalued Treatments. *Econometrica*, **86** 1939–1963.

Lee, S. and Salanié, B. (2023). Treatment Effects with Targeting Instruments. ArXiv:2007.10432 [econ, stat], URL http://arxiv.org/abs/2007.10432.

Lifshits, M. A. (1995). *Gaussian Random Functions*. Mathematics and Its Applications, Springer Netherlands.

Machado, C., Shaikh, A. M. and Vytlacil, E. J. (2019). Instrumental variables and the sign of the average treatment effect. *Journal of Econometrics*, **212** 522–555.

Mogstad, M. and Torgovitsky, A. (2024). Instrumental Variables with Unobserved Heterogeneity in Treatment Effects. URL https://www.nber.org/papers/w32927.

Mogstad, M., Torgovitsky, A. and Walters, C. R. (2021). The Causal Interpretation of Two-Stage Least Squares with Multiple Instrumental Variables. *American Economic Review*.

Mourifié, I. and Wan, Y. (2017). Testing Local Average Treatment Effect Assumptions. *The Review of Economics and Statistics*, **99** 305–313.

Słoczyński, T. (2020). When should we (not) interpret linear iv estimands as late? *arXiv preprint arXiv:2011.06695*.

Sun, Z. (2023). Instrument validity for heterogeneous causal effects. *Journal of Econometrics*, **237** 105523.

Vytlacil, E. (2002). Independence, Monotonicity, and Latent Index Models: An Equivalence Result. *Econometrica*, **70** 331–341.