Ordered, Unordered and Minimal Monotonicity Criteria

Manu Navjeevan and Rodrigo Pinto[*]

Revised April 16, 2022

Abstract

This paper performs a comparative analysis between ordered and unordered choice models. We present non-trivial symmetries between ordered and unordered monotonicity conditions. We show that these seemingly unrelated models share a weaker and more general condition called *Minimal Monotonicity.* This novel condition captures an essential property for the identification of causal parameters while being necessary for ascribing causal interpretation to Two Stage Least Squares estimands. We show that minimal monotonicity can by justified by a notion of rationality that naturally arises from revealed preference analysis. The condition serves as a theoretical foundation for a wide range of more sophisticated monotonicity conditions and economic behaviors that do not conform with the narrative dictated by ordered or unordered choice models.

Keywords: Monotonicity, Instrumental Variables, Discrete Choice, Selection bias, Roy Model, Identification, Discrete Mixture Model.
JEL codes: I21, C93, J15, V16.

---

[*] Rodrigo Pinto is a professor at the University of California at Los Angeles (UCLA), and Manu Navjeevan is a PhD candidate at UCLA. We are most grateful for the interaction with Professors James Heckman, Steven Durlauf, Moshe Buchinsky and Rosa Matzkin. We also benefit from useful suggestions from Stephane Bonhomme, Azeem Shaikh, Adriana Lleras-Muney, Dora Costa, Maurizio Mazzocco, Lawrence Katz, and Parag Pathak. We also thank useful comments on presentations at University of California at Los Angeles, University of Chicago, Science Po (Paris), and University of Geneva, All errors are our own.

1 Introduction

In 1994, Guido Imbens and Joshua Angrist published a widely influential paper that investigates mild restrictions to secure the nonparametric identification of causal parameters in a binary choice model using instrumental variables (IV). They introduce the *monotonicity condition* which states that a change in the instrument induces agents to shift their choice towards the same direction. The monotonicity condition has several desirable properties. It is simple, intuitive, and renders the identification of the Local Average Treatment Effect (LATE). It is often considered the minimal criteria necessary to ascribe causal interpretation to Two-Stage Least Squares (2SLS) estimands under binary choices.[1]

Contrary to what one might expect, identification in IV models with multiple choices is not a direct extension of the binary choice case. Multiple choices allow for a range of distinct monotonicity conditions. Notably, Angrist and Imbens (1995) directly apply the monotonicity of Imbens and Angrist (1994) to the case of multiple treatments. Vytlacil (2006) shows that the condition is equivalent to assuming an ordered choice model. In contrast, the unordered monotonicity of Heckman and Pinto (2018) applies to settings with unordered treatment choices; it neither implies nor is implied by the monotonicity of Angrist and Imbens (1995). Despite their specific motivations, both conditions are equivalent to the Imbens and Angrist (1994) monotonicity in the case of binary choices and both allow for a causal interpretation for 2SLS estimands.

While the two ordered and unordered monotonicity criteria discussed above share similarities, little is known about the relationship between them. The IV literature seldom considers a meta-analysis across monotonicity conditions in settings with multiple treatments. This paper fills this gap by considering two independent but linked inquiries. The first is on the relationship between ordered and unordered monotonicity. The shared key properties that suggest a deeper connection between the two criteria. We update the equivalence results of Vytlacil (2006) and Heckman and Pinto (2018) to provide symmetric characterizations of ordered and unordered monotonicity. These characterizations enable us to note some useful common properties of ordered and unordered monotonicity and set the stage for joint analysis of the two conditions.

---

[1] Huber and Mellace (2012) is an example of a paper that considers identification of the LATE via 2SLS under alternate conditions

The second inquiry considered is whether these two monotonicity conditions can be subsumed by a broader criterion that would still enable useful causal analysis. By leveraging their symmetric characterizations, we show that both conditions share a common property which we term the *minimal monotonicity condition.* This minimal monotonicity condition is precisely what is required for the two stage least squares between any two instrument values to identify an interpretable causal parameter. Moreover, in general in no weaker condition would allow for such causal interpretability of two stage least squares estimands.[2] We provide a characterization for minimal monotonicity that allows the researcher to easily verify whether the condition holds.

In addition, we show that minimal monotonicity can be justified by a notion of choice rationality significantly weaker than those displayed by agents in ordered and unordered choice models. Settings where ordered and unordered monotonicity hold can thus be seen as particular instances of a broad class of choice models described by the minimal monotonicity condition. By analyzing the properties of minimal monotonicity, we thus hope to facilitate development of monotonicity conditions that may be suitable in a range of economic settings that are not neatly described by ordered or unordered choice models. We provide some natural economic examples where ordered and unordered monotonicity fail, but where minimal monotonicity may still allow researchers to conduct meaningful causal analysis.

This paper contributes to the theoretical literature on ordered and unordered choice models. It adds to the literature that extends the understanding and usage of monotonicity conditions (Kamat, 2021; Mogstad et al., 2018; Mogstad and Torgovitsky, 2018; Hull, 2018). Our analyses are informative to a growing literature on empirical economics that examines non-standard monotonicity conditions to aid the identification and evaluation of treatment effects (Pinto, 2021; Kline and Walters, 2016; Mountjoy, 2021; Feller et al., 2016; Brinch et al., 2017; Kirkeboen et al., 2016). We additionally contribute to the literature tying monotonicity criterion to particular structural models (Vytlacil, 2002, 2006; Heckman and Pinto, 2018) by showing the minimal monotonicity is implied by a basic model of rationality.

This paper proceeds as follows. Section 2 reviews the prior literature on monotonicity conditions.

---

[2] Indeed, in cases where treatment is binary, we show that minimal monotonicity reduces exactly to the monotonicity of Imbens and Angrist (1994).

Section 3 describes the IV model and introduces our notation. Section 4 discusses the content of ordered and unordered monotonicity conditions and revisits the equivalence results for ordered and unordered choice models. It explores the symmetry of equivalence results between these two models to motivate a novel monotonicity condition. Section 5 discusses the properties of the Minimal Monotonicity Condition. Section 6 discusses the economic content of the minimal monotonicity condition. Section 7 discuss some applications of monotonicity criteria that are economically justified. Section 8 concludes.

2 Short Literature Review

Economists have long used instrumental variables (IV) to identify the causal effect of an endogenous treatment choice on an outcome of interest. The traditional literature uses structural equations to model the role of IV in determining the agent's choice (Goldberger, 1972; Heckman, 1976, 1979). In their seminal work, Imbens and Angrist (1994) departed from the traditional IV literature based on structural equations. Using the language of potential outcomes (Rubin, 1974, 1978), they introduce the notion of *monotonicity*, which formalizes an intuitive assumption stating that an IV change induces all agents toward choosing the same treatment choice.[3]

Angrist and Imbens (1995) extend this monotonicity condition to the case of multiple choices. They show that their monotonicity provides a causal interpretation of the conventional Two-Stage Square Least Squares (2SLS) estimand in models with endogenous choices and heterogeneous treatment responses. Their work spiked a substantial literature on both empirical and theoretical aspects of monotonicity conditions (Angrist et al., 2000; Barua and Lang, 2016; Dahl et al., 2017; Huber and Mellace, 2012, 2015; Imbens and Rubin, 1997; Klein, 2010; Small and Tan, 2007; Aliprantis, 2012; de Chaisemartin, 2017).[4]

Vytlacil (2002, 2006) bridge the gap between IV models that rely on monotonicity conditions and the previous literature that invokes structural equations. Vytlacil (2002) shows that the monotonicity condition of Imbens and Angrist (1994) is equivalent to the random threshold crossing model of Heckman and Vytlacil (1999, 2005, 2007a). Vytlacil (2006) shows the monotonicity criterion of

---

[3] See also Angrist et al. (1996).

[4] Huber et al. (2017) consider weaker assumptions at the principal strata level, which are also employed by Frölich (2007).

Angrist and Imbens (1995) is equivalent to an ordered choice model with random thresholds. This model is examined by Cameron and Heckman (1998) and further studied by Carneiro et al. (2003); Cunha et al. (2007).

Unordered choice models have been studied mostly by literature on structural equations. A common approach is to assume that the equations that govern the treatment are generated by additively separable threshold-crossing models. Examples of this literature are Heckman and Vytlacil (2007b); Heckman et al. (2006, 2008). A substantial contribution to this literature is due to Lee and Salanié (2018), who studied the identification of causal effects for choice models defined by an arbitrary set of threshold-crossing rules. Heckman and Pinto (2018) connect the structural and monotonicity approaches. They present an economically motivated condition termed *unordered monotonicity* which applies to treatment values that do not have a natural order. Building upon Vytlacil (2002), they further show that unordered monotonicity can be equivalently expressed as a multivariate choice model with latent crossing thresholds.

Little is known about the shared features of ordered and unordered choice models. The rationale that generates an ordered choice model is considerably different from the motivation that justifies unordered choices. Not surprisingly, each model often carries distinct mathematical formalizations. A rare example of a comparative discussion between ordered and unordered choice models is Heckman et al. (2006). Their ordered choice model employs a partition of the real line by non-stochastic thresholds. The treatment choice indicates the interval that the latent stochastic index lies in this partition. In contrast, their unordered choice model employs a set of latent indexes that are additive in the observed and unobserved characteristics of the agent.

As mentioned, we perform a comparative analysis between ordered and unordered monotonicities. To do so, we revisit the monotonicity condition of Angrist and Imbens (1995) using new tools of analysis developed in Heckman and Pinto (2018).

3 Setup

Following Imbens and Angrist (1994), we examine causal relationships using the potential outcomes framework of Rubin (1974, 1978). Our model consists of three observed variables: the treatment status T taking on values in $\mathcal{T} = \{t_1, \ldots, t_{N_T}\}$, the instrument Z taking on values in $\mathcal{Z} = \{z_1, \ldots, z_{N_Z}\}$, and an outcome of interest $Y \in \mathbb{R}$. The observed variables are generated by the latent potential treatments $\{T(z),\, z \in \mathcal{Z}\}$ and potential outcomes $\{Y(t),\, t \in \mathcal{T}\}$ by the following relationships:

$$Y = \sum_{t \in \mathcal{T}} \mathbf{1}\{T = t\} Y(t)$$
$$T = \sum_{z \in \mathcal{Z}} \mathbf{1}\{Z = z\} T(z)$$

Notice that we are implicitly assumingu SUTVA and the exclusion restriction here; the potential outcomes do not depend on the treatment status of other members of the population nor do they depend directly on the instrument. We suppress pre-treatment variables from the model for the sake of notational simplicity; the analysis can be understood as conditioned on these variables.

The identification of causal effects is complicated by the existence of unobserved confounders that may effect both the treatment decision T and the potential outcomes $Y(z)$. To get around this, the standard IV research design assumes that the instrument is assigned at random, formally;

$$\text{Exogeneity Condition:} \quad Z \perp\!\!\!\perp (T(z), Y(t)) \text{ for all } (z,t) \in \mathcal{Z} \times \text{supp}(Y) \tag{1}$$

Our analysis will make use of the response vector $\boldsymbol{S}$ that stacks together all the potential treatments as the instrument ranges in $\mathcal{Z}$:

$$\boldsymbol{S} = [T(z_1), ..., T(z_{N_Z})]^\intercal, \quad \text{supp}(\boldsymbol{S}) \equiv \{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_{N_S}\} \tag{2}$$

Elements of the support of the response vector, $\boldsymbol{s} \in \text{supp}(\boldsymbol{S})$, are called *response-types.* Consider the LATE model of Imbens and Angrist (1994) where $\mathcal{Z} = \{z_0, z_1\}$ and $\mathcal{T} = \{t_0, t_1\}$. Without any restrictions, the response vector $\boldsymbol{S} = [T(z_0),\, T(z_1)]^\intercal$, admits four possible response-types, $\{\boldsymbol{s}_{\text{nt}}, \boldsymbol{s}_{\text{c}}, \boldsymbol{s}_{\text{at}}, \boldsymbol{s}_{\text{d}}\}$: never-takers $\boldsymbol{s}_{\text{nt}} = [t_0,\, t_0]^\intercal$, compliers $\boldsymbol{s}_{\text{c}} = [t_0,\, t_1]^\intercal$, always-takers $\boldsymbol{s}_{\text{at}} = [t_1,\, t_1]^\intercal$, and defiers $\boldsymbol{s}_{\text{d}} = [t_1,\, t_0]^\intercal$.

Importantly, the response vector $\boldsymbol{S}$ plays the role of a balancing score for the unobserved confounders, $Y(t) \perp\!\!\!\perp T|\boldsymbol{S}$. This property enables us to connect observed conditional expectations with the counterfactual outcomes of interest and response-type probabilities via the following equation:[5]

$$\begin{aligned}&\underbrace{E(Y|T=t,Z=z)P(T=t|Z=z)}_{\text{Observed}}\\&=\sum_{\boldsymbol{s}\in\text{supp}(\boldsymbol{S})}\underbrace{\mathbf{1}[T=t|\boldsymbol{S}=\boldsymbol{s},Z=z]}_{\text{Known}}\cdot\underbrace{E(Y(t)|\boldsymbol{S}=\boldsymbol{s})P(\boldsymbol{S}=\boldsymbol{s})}_{\text{Unobserved}}.\end{aligned} \tag{3}$$

The first term of the right-hand side of the equation is nonrandom since T is a deterministic function of the instrument Z and the response type $\boldsymbol{S}$. The second term on the right-hand side is unobserved. It comprises expected value of counterfactual outcomes conditioned on response-types $E(Y(t)|\boldsymbol{S}=\boldsymbol{s})$ and response-type probabilities $P(\boldsymbol{S}=\boldsymbol{s})$. Our goal is to use the observed conditional expectations on the left-hand side of (3), $E[Y|T=t,Z=z]$ for $(t,z)\in\mathcal{T}\times\mathcal{Z}$, to make inferences on the unobserved counterfactuals on the right-hand side.

3.1 The Response Matrix

Much of our analysis is presented as (simple) matrix algebra conditions on the response matrix. The response matrix organizes the N_S eligible response types in the support of the response vector into a $N_Z\times N_S$ array where each column displays a response type and each row corresponds to an instrument value:

$$\boldsymbol{R}\equiv[\boldsymbol{s}_1,\dots,\boldsymbol{s}_{N_S}]\in\mathcal{T}^{N_Z\times N_S}. \tag{4}$$

The entry in the z^{th} row and $\boldsymbol{s}^{\text{th}}$ column of the response matrix is given by $\boldsymbol{R}[z,\boldsymbol{s}]$. It denotes the treatment choice that an agent i of type $\boldsymbol{S}_i=\boldsymbol{s}$ would take when exposed to instrumental value z. Heckman and Pinto (2018) show that the response matrix $\boldsymbol{R}$ contains all required information to investigate the non-parametric identification of counterfactual outcomes and response type probabilities.

In the case of the binary LATE model, the ordered monotonicity condition is equivalent to the unordered monotonicity condition. These conditions both eliminate defiers from the support of the response vector and so the LATE response matrix is given:

---

[5] See Heckman and Pinto (2018) for a proof.

$$\boldsymbol{R}^{\text{bLATE}} = \begin{array}{cccc} \boldsymbol{s}_{\text{nt}} & \boldsymbol{s}_{\text{c}} & \boldsymbol{s}_{\text{at}} & \\ \left[\begin{array}{ccc} t_0 & t_0 & t_1 \\ t_0 & t_1 & t_1 \end{array}\right] & & & \begin{array}{c} z_0 \\ z_1 \end{array} \end{array} . \tag{5}$$

The columns of the response matrix (5) stacks the three remaining response-types in the support of the response vector: never-takers ($\boldsymbol{s}_{\text{nt}}$), the compliers ($\boldsymbol{s}_{\text{c}}$), and the always-takers ($\boldsymbol{s}_{\text{at}}$). The first row indicates the treatment decisions of agents with each response types when exposed to instrument value z_0 while the second row corresponds to their treatment decisions when exposed to instrument value z_1.

In our analysis, it will also be useful to work with the binary matrices $\boldsymbol{B}_t = \mathbf{1}[\boldsymbol{R} = t]$, which indicate whether each element of $\boldsymbol{R}$ is equal to a patricular treatment $t \in \mathcal{T}$. Equation (6) displays the binary matrices for the LATE model.

$$\boldsymbol{B}_{t_0} \equiv \mathbf{1}[\boldsymbol{R} = t_0] = \begin{array}{c} \begin{array}{ccc} \boldsymbol{s}_{nt} & \boldsymbol{s}_{c} & \boldsymbol{s}_{at} \end{array} \\ \left[\begin{array}{ccc} 1 & 1 & 0 \\ 1 & 0 & 0 \end{array}\right] \end{array} \begin{array}{c} \\ z_0 \\ z_1 \end{array}, \qquad \boldsymbol{B}_{t_1} \equiv \mathbf{1}[\boldsymbol{R} = t_1] = \begin{array}{c} \begin{array}{ccc} \boldsymbol{s}_{nt} & \boldsymbol{s}_{c} & \boldsymbol{s}_{at} \end{array} \\ \left[\begin{array}{ccc} 0 & 0 & 1 \\ 0 & 1 & 1 \end{array}\right] \end{array} \begin{array}{c} \\ z_0 \\ z_1 \end{array} \tag{6}$$

Each entry of the binary matrix $\boldsymbol{B}_t[z, \boldsymbol{s}]$ indicates whether response type $\boldsymbol{s}$ takes up treatment t when exposed to instrument z. For example, in (6), $\boldsymbol{B}_{t_0}[z_0, \boldsymbol{s}_{nt}] = \boldsymbol{B}_{t_0}[z_0, \boldsymbol{s}_c] = 1$, indicates that the never takers and compliers take up treatment value t_0 when exposed to instrument z_0. Conversely, $\boldsymbol{B}_{t_0}[z_0, \boldsymbol{s}_{at}] = 0$, which indicates that the always-takers do not take up treatment value t_0 when exposed to instrument z_0.

For analysis of ordered monotonicity it will additionally be useful to work with the binary matrices $\boldsymbol{B}^*_t = \mathbf{1}[\boldsymbol{R} \geq t]$, which implicitly require an ordering on $\mathcal{T}$. The matrix $\boldsymbol{B}^*_t$ can be generated from $\boldsymbol{B}_{t_1}, \ldots, \boldsymbol{B}_{t_{N_T}}$ by summing over all the treatment values greater than or equal to t, $\boldsymbol{B}^*_t = \sum_{t' \geq t} \boldsymbol{B}_t$. Equation (7) displays the matrices $\boldsymbol{B}^*_t$ for the binary LATE model, under the assumption that $t_1 > t_0$:

$$\boldsymbol{B}^*_{t_0} \equiv \mathbf{1}[\boldsymbol{R} \geq t_0] = \begin{array}{c} \begin{array}{ccc} \boldsymbol{s}_{nt} & \boldsymbol{s}_{c} & \boldsymbol{s}_{at} \end{array} \\ \left[\begin{array}{ccc} 1 & 1 & 1 \\ 1 & 1 & 1 \end{array}\right] \end{array} \begin{array}{c} \\ z_0 \\ z_1 \end{array}, \qquad \boldsymbol{B}^*_{t_1} \equiv \mathbf{1}[\boldsymbol{R} \geq t_1] = \begin{array}{c} \begin{array}{ccc} \boldsymbol{s}_{nt} & \boldsymbol{s}_{c} & \boldsymbol{s}_{at} \end{array} \\ \left[\begin{array}{ccc} 0 & 0 & 1 \\ 0 & 1 & 1 \end{array}\right] \end{array} \begin{array}{c} \\ z_0 \\ z_1 \end{array} \tag{7}$$

4 Ordered and Unordered Monotonicity

A fundamental identification problem in the IV model is that the number of unknown counterfactuals typically far exceeds the number of identifying restrictions retrieved from observed data. Stacking equation (3) over all possible values of Z and T generates a linear system with $N_Z \cdot N_T$ linear equations. The number of unknowns in this system is proportional to the number of response-types in the support of the response vector $\boldsymbol{S}$. The response vector, $\boldsymbol{S}$, on the other hand, is a N_Z-dimensional vector whose elements can take on any of the N_T treatment values. Left unrestricted, this leaves $N_T^{N_Z}$ response types in supp$(\boldsymbol{S})$.

The exponential growth in the number of response types complicates identification of counterfactuals in models with multiple treatments and multiple instruments. Identification is then contingent on carefully restricting the support of $\boldsymbol{S}$. Monotonicity conditions are systematic ways of making such restrictions, motivated by economic reasoning. As discussed previously, Angrist and Imbens (1995) and Heckman and Pinto (2018) provide monotonicity criterions for ordered and unordered choice models, respectively. We will refer to these conditions as ordered monotonicity (8) and unordered monotonicity (9) for the sake of clarity:

Ordered Monotonicity (OM): For any $z, z' \in \mathcal{Z}$ either,

$$\begin{aligned} & T_i(z) \geq T_i(z') \text{ for all } i \in \mathcal{I} \\ \text{or} \quad & T_i(z) \leq T_i(z') \text{ for all } i \in \mathcal{I}. \end{aligned} \tag{8}$$

Unordered Monotonicity (UM): For any $z, z' \in \mathcal{Z}$ and any $t \in \mathcal{T}$ either,

$$\begin{aligned} & \mathbf{1}[T_i(z) = t] \geq \mathbf{1}[T_i(z') = t] \text{ for all } i \in \mathcal{I} \\ \text{or} \quad & \mathbf{1}[T_i(z) = t] \leq \mathbf{1}[T_i(z) = t] \text{ for all } i \in \mathcal{I} \end{aligned} \tag{9}$$

OM (8) captures the notion that a change in instrumental values produces incentives that either move all agents towards weakly"higher" treament values or move all agents towards weakly "lower" treatment values. The condition can be understood as stating that an instrumental change that induces one agent to increase their treatment choice cannot cause another agent to decrease their treatment choice. The condition requires an ordinal treatment, such as years of schooling.

UM (9) states that for each treatment, each instrumental change must either move all agents weakly towards that treatment or weakly away from the treatment. This differs from OM (8) as it compares *the indicator function* of the treatment instead of the treatment value itself. Because of this, UM (9) does not require ordered treatments, making it relevant for settings where the treatment has no natural ordering such as analysis of college major choice or neighborhood effects.[6]

Importantly, both OM (8) and UM (9) enable the researcher to identify a mixture of Local Average Treatment Effects (LATEs) with identifiable weights and both conditions ascribe causal interpretations to the estimands of Two-Stage Least Squares (2SLS) regressions.

4.1 Expressing Monotonicities as Sequences of Counterfactual Choices

Because the definition of OM (8) compares treatment values, it requires that $\mathcal{T}$ be an ordered set. We propose a slightly more inclusive definition of ordered monotonicity that does not require an ordered treatment. The central property of ordered monotonicity is a mapping between a sequence of IV values and some sequence of treatment values in which higher rankings of Z correspond to higher rankings of Z. The following formula expresses this criterion:

OM Sequence: There exist a sequencing of $\mathcal{Z}$, $(z_1, \ldots, z_{N_Z})$, and a strict ordering on $\mathcal{T}$ such that:[7]

$$(T_i(z_1), \ldots, T_i(z_{N_Z})) \text{ is an increasing sequence in } \mathcal{T} \text{ for any } i \in \mathcal{I}. \tag{10}$$

The OM sequential criteria (10) generates the OM condition (8) whenever the ordering $\mathcal{T}$ is assumed, however it does not require a specific ordering on $\mathcal{T}$ a priori. In Section 7 we will demonstrate the usefulness of this more inclusive definition with a plausible research design that generates OM-Sequence (10) on a treatment space that has no natural ordering.

---

[6] Additionally, Heckman and Pinto (2018) show that UM occurs naturally in economic settings where choice incentives weakly increase among all treatment choices as the instrument varies.Buchinsky and Pinto (2021) use revealed preference analysis to show how choice incentives induced by the instrumental variable generate a range of monotonicity conditions.

[7] A strict ordering is one such that for any $t, t' \in \mathcal{T}$ with $t \neq t'$ exactly one of $(t' \geq t)$ or $(t \geq t')$ is true.

We can also characterize the UM condition in (9) in terms of a sequence of counterfactual choices:

UM Sequence: For each $t \in \mathcal{T}$ there exists a sequencing of $\mathcal{Z}$, $(z_1^{(t)}, \ldots, z_{N_Z}^{(t)})$ such that:

$$(\mathbf{1}[T_i(z_1^{(t)}) = t], \ldots, \mathbf{1}[T_i(z_{N_Z}^{(t)}) = t]) \text{ is weakly increasing for any } i \in \mathcal{I}. \tag{11}$$

UM Sequence (11) differs from OM Sequence (10) in two significant ways. First, the sequence of IV values in the unordered case can differs across treatment values while the IV sequence of ordered case remains the same for all $t \in \mathcal{T}$. Second, UM Sequence (11) utilizes treatment indicators, while the OM Sequence (10) employs the treatment values themselves.

It is easy to see that the OM and UM sequence characterizations in (10) and (11) are equivalent for a binary treatment. However, this equivalence between ordered and unordered monotonicity breaks down for choice models with three or more treatment choices. In general, ordered monotonicity does not imply unordered monotonicity nor vice versa. To partially demonstrate why this is the case, consider an ordering on the treatments where $t_1 \leq t_2 \leq t_3$ and the following two pairs of treatment response patterns:

$$\begin{array}{cc} \boldsymbol{s}_a & \boldsymbol{s}_b \\ \end{array} \qquad\qquad \begin{array}{cc} \boldsymbol{s}_c & \boldsymbol{s}_d \\ \end{array}$$

$$\begin{pmatrix} t_1 & t_2 \\ t_2 & t_3 \end{pmatrix} \begin{array}{c} z \\ z' \end{array} \qquad\qquad \begin{pmatrix} t_1 & t_3 \\ t_2 & t_2 \end{pmatrix} \begin{array}{c} z \\ z' \end{array} \tag{12}$$

The treatment response patterns displayed by agents $\boldsymbol{s}_a$ and $\boldsymbol{s}_b$ are natural under ordered monotonicity, they could can be rationalized by a research design where z' provides uniformly greater treatment incentives than z. However, they cannot both exist in a response matrix that satisfies unordered monotonicity. By examining UM Sequence (11) we can see that there would be no possible sequencing of the instruments that would allow the sequence of t_2 indicators to be weakly increasing for all agents.

Similarly, the treatment response patterns displayed by agents $\boldsymbol{s}_c$ and $\boldsymbol{s}_d$ in (12) are not prohibited by unordered monotonicity; they can both be rationalized by a research design where instrument z' explicitly incentivizes treatment t_2. Despite this, they cannot both be present in a response matrix that satisfies ordered monotonicity under the ordering on the treatments given above. Since the switch from instrument z to instrument z' induces agent $\boldsymbol{s}_c$ to switch to a "lower" treatment while inducing agent $\boldsymbol{s}_d$ to switch to a "higher" treatment, there can be no sequencing on the instruments satisfying OM Sequence (10).

The two conditions are also not mutually exclusive, it is possible for a response matrix to satsify both ordered and unordered monotonicity, even when the treatment is multi-valued. Moreover, the lack of nesting between ordered and unordered monotonicity does not change if we consider all possible orderings on the treatment space. For a complete example of this and more in depth discussion, refer to Appendix C.

4.2 Characterizations of Unordered and Ordered Monotonicity

We first present an updated version of the unordered equivalence result in Heckman and Pinto (2018):

Theorem 1 (Unordered Equivalence)**.** *The following statements are equivalent:*

(i). For each $t \in \mathcal{T}$ there is a sequence of instruments $(z_1^{(t)}, ..., z_{N_T}^{(t)})$ such that UM Sequence (11) holds.

(ii). Given any $t \in \mathcal{T}$ and any $k \in \{1, ..., N_Z - 1\}$, we have that

$$\mathbf{1}[T_i(z_{k+1}^{(t)}) = t] \geq \mathbf{1}[T_i(z_k^{(t)}) = t] \text{ for all } i \in \mathcal{I}..$$

(iii). For any $t \in \mathcal{T}$ and $t', t'' \neq t$ there are no 2×2 submatrices in $\boldsymbol{R}$ of the form:

$$\begin{pmatrix} t & t'' \\ t' & t \end{pmatrix} \text{ or } \begin{pmatrix} t' & t \\ t & t'' \end{pmatrix}. \tag{13}$$

(iv). For the unordered verification matrix $\Psi_{\boldsymbol{U}}$ defined below, $\|\Psi_{\boldsymbol{U}}\| = 0$;

$$\Psi_{\boldsymbol{U}} \equiv \big((\mathbb{1} - \boldsymbol{U})^{\intercal}\boldsymbol{U}\big) \odot \big((\mathbb{1} - \boldsymbol{U})^{\intercal}\boldsymbol{U}\big)^{\intercal}. \tag{14}$$

where $\mathbb{1}$ *denotes a matrix of all ones,* $\odot$ *denotes the Hadamard (element-wise) product and:*

$$\boldsymbol{U} \equiv \begin{bmatrix} \boldsymbol{B}_1 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{1} & \boldsymbol{B}_2 & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \mathbf{1} & \mathbf{1} & \mathbf{1} & \cdots & \boldsymbol{B}_{N_T} \end{bmatrix}. \tag{15}$$

(v). For each $t \in \mathcal{T}$*, there are real-valued functions* $\varphi(\cdot, t)$ *and* $\zeta(\cdot, t)$ *such that the treatment choice* T *can be rationalized by:*

$$\mathbf{1}[T = t] = \mathbf{1}\left[\zeta(Z, t) \geq \varphi(\boldsymbol{V}, t)\right],$$

where $\zeta(z_{k+1}^{(t)}, t) > \zeta(z_k^{(t)}, t)$ *for* $k = 1, ..., N_Z - 1$ *and any* t*.*

Proof. See Appendix A. □

The first two items of Theorem 1 reflect the discussion in Section 4 relating UM-Sequence (11) to the classical definition of unordered monotonicity introduced in Heckman and Pinto (2018) and restated in (8).

Item *(iii)* resembles the discussion above and states that unordered monotonicity can be verified by individually checking each 2×2 submatrix of $\boldsymbol{R}$. We can see that the appearance of one of the restricted submatrices in (13) prevents the existence of a sequence of instruments that would make the sequence of treatment t indicators increasing for all agents. Unfortunately, this requirement may be difficult to verify in practice, since the number of 2×2 submatices is growing exponentially with the dimensions of the response matrix. Item *(iv)* of the theorem provides a practical method of verifying the condition using matrix algebra. Item *(v)* is faimilar to the literature and provides an equivalence between unordered monotonicity and separability conditions such as in Vytlacil (2002). For further discussion, Heckman and Pinto (2018) describe other useful aspects of each of the equivalent statements, *(iii)* and *(v)* above.

We next present an equivalence result for Ordered Monotonicity.

Theorem 2 (Ordered Equivalence)**.** *The following statements are equivalent:*

(i). There is is a sequence on $\mathcal{Z}$, $(z_1, \ldots, z_{N_Z})$ and a strict ordering on $\mathcal{T}$ that satisfies the requirement of OM-Sequence (10).

(ii). There is a strict ordering on $\mathcal{T}$ such that for any $k \in \{1, \ldots, N_Z - 1\}$ and any t:

$$\mathbf{1}[T_i(z_{k+1}) \geq t] \geq \mathbf{1}[T_i(z_k) \geq t] \text{ for all } i \in \mathcal{I}.$$

(iii). There is a strict ordering on $\mathcal{T}$ such that for any $t < t''$ and $t' > t'''$ there are no 2×2 submatrices of $\boldsymbol{R}$ of the form either

$$\begin{pmatrix} t & t' \\ t'' & t''' \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} t' & t \\ t''' & t'' \end{pmatrix}; \tag{16}$$

(iv). There is a strict ordering on $\mathcal{T}$ such that for the ordered verification matrix $\Psi_{\boldsymbol{O}}$ defined below, $\|\Psi_{\boldsymbol{O}}\| = 0$;

$$\Psi_{\boldsymbol{O}} \equiv \left((\mathbb{1} - \boldsymbol{O})^\intercal \boldsymbol{O}\right) \odot \left((\mathbb{1} - \boldsymbol{O})^\intercal \boldsymbol{O}\right)^\intercal, \tag{17}$$

where $\mathbb{1}$ indicates a matrix of all ones, $\odot$ represents the Hadamard (element-wise) product, and:

$$\boldsymbol{O} \equiv \left[\boldsymbol{B}^*_{t_1}, \ldots, \boldsymbol{B}^*_{t_{N_T}}\right];$$

(v). There is a strict ordering on $\mathcal{T}$ such that for some real-valued functions $\varphi(\cdot, t)$ and $\zeta(\cdot, t)$ the treatment choice can be rationalized by

$$\mathbf{1}[T \geq t] = \mathbf{1}[\zeta(Z, t) \geq \varphi(\boldsymbol{V}, t)],$$

where $\zeta(z_{k+1}, t) > \zeta(z_k, t)$ for $k = 1, \ldots, N_Z - 1$ and any t.

Proof. See Appendix A □

Theorem 2 extends Vytlacil (2006) in a fashion that enables us to compare ordered and unordered monotonicity conditions. The first and second items of the ordered equivalence result reconcile the two notions of ordered monotonicity presented above. It shows that if OM-Sequence (10) holds, we can find an ordering on $\mathcal{T}$ that satisfies the typical definition of ordered monotonicity and vice versa; if there is an ordering on $\mathcal{T}$ that satisfies ordered monotonicity we can find a sequence on $\mathcal{Z}$ to satisfy OM-Sequence (10).[8] Item *(iii)* of Theorem 2 provides a similar insight to Item *(iii)* of Theorem 1, namely that ordered monotonicity can be verified simply by looking at the 2×2 submatrices of the response matrix $\boldsymbol{R}$. Item *(iv)* provides a tractable method for verifying this property.

The final item of the theorem restates the equivalence result of Vytlacil (2006) and shows that assuming ordered monotonicity is equivalent to taking an ordered choice behavioral model. While this result is familar to the literature, we provide an alternative proof in Appendix A using properties of *lonesum binary matrices*; a concept we borrow from the information theory literature (Ryser, 1957).

Symmetries between Ordered and Unordered Monotonicity

The characterizations of ordered monotonicity in Theorem 2 are symmetric to those of unordered monotonicity in Theorem 1. We have already discussed the usefulness of some of these specific symmetries above. For example, the symmetry between the sequential characterizations of ordered and unordered monotonicities provides an easy way of seeing that ordered and unordered monotonicity are equivalent in the case of a binary treatment. Other symmetries are new to our discussion and are worth briefly mentioning. The symmetric matrix verification characterizations provides an easy way to verify if a response matrix satisfies both the ordered and unordered monotonicity conditions by checking if $\|\Psi_{\boldsymbol{U}}\| + \|\Psi_{\boldsymbol{O}}\| = 0$. Verifying this allows researchers to take advantage of both sets of identification results.

The symmetry between the restricted 2×2 submatrices in ordered and unordered monotonicity provides insight on how a response matrix could satisfy ordered monotonicity but not unordered monotonicity and vice versa. More importantly, however, the similarity between the two restricted submatrices in Theorems 1 and 2 suggests a common condition shared by both criteria. In particular, note that both restricted patterns in (13) and (16) prevent any two agents from having exactly

---

[8] In particular we can take the sequence that orders z' after z if $T_i(z') \geq T_i(z)$ for all $i \in \mathcal{I}$.

opposite treatment responses for the same instrumental variable shift. We term this common restriction the *Minimal Monotonicity Condition* and analyze its properties in Section 5.

5 The Minimal Monotonicity Condition

The minimal monotonicity (MM) condition (18) is a weak criteria shared by both ordered and unordered conditions. It is determined by a symmetric restriction that is common to Theorems 1 and 2. Indeed, it turns out that minimal monotonicity is the core common property of ordered and unordered monotonicity that enables the 2SLS estimand to achieve causal interpretability.

Minimal Monotonicity (MM): For any pair of instruments $z, z' \in \mathcal{Z}$ and any pair of treatments $t, t' \in \mathcal{T}$ either

$$\begin{aligned} & \mathbf{1}[T_i(z)=t]\mathbf{1}[T_i(z')=t'] \geq \mathbf{1}[T_i(z)=t']\mathbf{1}[T_i(z')=t] \, \forall i \ \in \mathcal{I} \\ \text{or} \quad & \mathbf{1}[T_i(z)=t]\mathbf{1}[T_i(z')=t'] \leq \mathbf{1}[T_i(z)=t']\mathbf{1}[T_i(z')=t] \, \forall \, i \in \mathcal{I}. \end{aligned} \tag{18}$$

The first row in (18) states that an instrumental change from z to z' incentives *all* agents to shift their choice away from t and towards t'. The second row in (18) describes the opposite behavior. In summary, the MM condition states that an intrumental change that induces an agent to switch its choice from t to t' cannot induce another another agent to switch its choice from t' to t. Lemma 1 provides an equivalent characterization of the MM condition in terms of response-types.

Lemma 1. *Minimal monotonicity MM holds if and only if for all distinct instruments* $z, z' \in \mathcal{Z}$ *and all distinct treatments* $t, t' \in \mathcal{T}$*, there are* no *response-types* $\boldsymbol{s}, \boldsymbol{s}' \in \text{supp}(\boldsymbol{S})$ *such that*

$$\begin{pmatrix} \boldsymbol{s}[z] & \boldsymbol{s}'[z] \\ \boldsymbol{s}[z'] & \boldsymbol{s}'[z'] \end{pmatrix} = \begin{array}{cc} \begin{array}{cc} \boldsymbol{s} & \boldsymbol{s}' \end{array} & \\ \begin{pmatrix} t & t' \\ t' & t \end{pmatrix} & \begin{array}{c} z \\ z' \end{array} \end{array} \tag{19}$$

Proof. See Appendix A ☐

Lemma 1 presents the prohibited pattern of 2 × 2 submatrices of the response matrix $\boldsymbol{R}$ induced by MM. The pattern is the common intersection between the submatrix characterizations in item *(iii)* of Theorems 1 and 2. Lemma 2 establishes that MM is *strictly* weaker than MM and OM.

Lemma 2. *The following relationships are true of ordered, unordered, and minimal monotonicity:*

1. *UM* $\Rightarrow$ *MM, but MM* $\nRightarrow$ *UM*
2. *OM* $\Rightarrow$ *MM, but MM* $\nRightarrow$ *OM*

Proof. See Appendix A ☐

To see why minimal monotonicity is crucial for the interpretability of 2SLS, it is useful to quickly define and discuss interpretable causal parameters.

5.1 Interpretable Causal Parameters

We follow an established literature that defines a meaningful causal parameter τ as a weighted average of local average treatment effects with positive weights:[9]

$$\tau = \sum_{\{t,t'\},t\neq t'} \omega_{t,t'} E[Y(t) - Y(t') \mid \boldsymbol{S} \in \mathcal{S}_{t,t'}] \quad \text{with } \omega_{t,t'} = 0 \text{ or } \omega_{t',t} = 0. \tag{20}$$

Here $\mathcal{S}_{t,t'}$ denotes a set of response types may vary according to the treatments being compared and $\omega_{t,t'} \geq 0$ are positive weights. The defining idea is that each treatment pair is only represented once, so we cannot have a positive weight on both $E[Y(t)-Y(t') \mid \boldsymbol{S} \in \mathcal{S}_{t,t'}]$ and $E[Y(t')-Y(t) \mid \boldsymbol{S} \in \mathcal{S}_{t',t}]$. The absence of negative weights allows this causal parameter to give us meaningful insight into the direction of the treatment effects.

Angrist and Imbens (1995) demonstrate that, under ordered monotonicity, the 2SLS estimand identifies such a meaningful causal parameter using a binary instrument with multiple treatments. Heckman and Pinto (2018) show a similar result for unordered monotonicity using comparisons of the outcome Y for any two instruments $z, z' \in \mathcal{Z}$. The equivalence result for minimal monotonicity in Theorem 3 establishes that it is indeed MM that is the driving force behind both of these identification results.

---

[9] For examples of works that adopt this criteria, see Angrist and Imbens (1995); Heckman and Urzúa (2010); Kirkeboen et al. (2016); Mogstad et al. (2021).

Lemma 2 provides some intuition for why this is the case. Consider the difference in average outcome between two values of the instrument, as in the numerator of a 2SLS estimand. This difference always has a unique decomposition into a weighted average of treatment effects among all the (ordered) pairs of possible treatment values.[10] The restriction on the response matrix imposed by minimal monotonicity (19) means that if one pair of treatment values, (t, t'), is represented in this weighted sum, the opposite pair, (t', t), cannot also be represented. So, minimal monotonicity is sufficient for this difference to satisfy the condition for an interpretable causal parameter (20). Moreover, because this decomposition is unique, we can show that minimal monotonicity is necessary for interpretability as well.

5.2 Equivalence Results

We now provide a set of equivalent characterizations of the minimal monotonicity condition in the spirit of the results for unordered and ordered monotonicity in Theorems 1 and 2.

Theorem 3 (Minimal Monotonicity Equivalence). *The following statements are equivalent:*

(i). For any distinct pair of instruments $z, z' \in \mathcal{Z}$ and any pair of treatments, $t, t' \in \mathcal{T}$, we have either:

$$\begin{aligned} &\mathbf{1}[T_i(z) = t]\mathbf{1}[T_i(z') = t'] \geq \mathbf{1}[T_i(z) = t']\mathbf{1}[T_i(z') = t] \; \forall i \; \in \mathcal{I} \\ \textit{or} \quad &\mathbf{1}[T_i(z) = t]\mathbf{1}[T_i(z') = t'] \leq \mathbf{1}[T_i(z) = t']\mathbf{1}[T_i(z') = t] \; \forall \, i \in \mathcal{I}. \end{aligned} \tag{21}$$

(ii). There are no 2×2 submatrices of $\boldsymbol{R}$ of the form in (19).

(iii). For the matrix $\Psi_{\boldsymbol{M}}$ defined below, $\|\Psi_{\boldsymbol{M}}\| = 0$

$$\Psi_{\boldsymbol{M}} \equiv \sum_{t \neq t'} \left(\boldsymbol{B}_t^\intercal \boldsymbol{B}_{t'}\right) \odot \left(\boldsymbol{B}_t^\intercal \boldsymbol{B}_{t'}\right)^\intercal. \tag{22}$$

where $\odot$ represents the Hadamard (element-wise) multiplication.[11]

---

[10] See Appendix B for a discussion of the exact forms of this decomposition as well as the 2SLS estimands for ordered and unordered monotonicity mentioned above.

[11] We use the short-hand notation $\sum_{t \neq t'} \xi(t, t') \equiv \sum_{t \in \mathcal{T}} \sum_{t' \in \mathcal{T} \setminus \{t\}} \xi(t, t')$.

(iv). For any *pair of instruments z, z' the 2SLS type estimand*

$$\beta_{z,z'} = E[Y \mid Z = z] - E[Y \mid Z = z']$$

identifies an interpretable causal parameter as described in (20).

Proof. See Appendix A ☐

Many features of this equivalence result are symmetric to the unordered and ordered equivalence results of Theorems 1 and 2. Item *(i)* defines the complete version of the MM condition. Items *(ii)* and *(iii)* of Theorem 3 provide ways of verify the MM condition symmetric to counterparts for unordered and ordered monotonicity in Theorems 1 and 2. Item *(ii)* presents a general response matrix condition. It states that no 2×2 submatrix of the response-matrix $\boldsymbol{R}$ presents the prohibited pattern in (19). Item *(iii)* provides a tractable method of verifying the MM condition. The verification requires an order of $\mathcal{T}^2$ matrix operations.

The last item of Theorem 3 is the empirically relevant feature of the MM condition. It provides a solution to our initial inquiry on a weak mononocity criteria that ensures interpretable causal parameters for the widely used method of 2SLS. Indeed, there can be no weaker monotonicity criterion that guarantees such causal interpretability.

5.3 Relationship Between Monotonicity Criterion

The three monotonicity conditions are equivalent in the case of a binary treatment. In this special case the definition of MM (18) reduces to:[12]

$$\begin{aligned} \mathbf{1}[T_i(z) = t] \geq \mathbf{1}[T_i(z') = t] \quad & \text{for all } i \in \mathcal{I} \\ \text{or } \mathbf{1}[T_i(z) = t] \leq \mathbf{1}[T_i(z') = t] \quad & \text{for all } i \in \mathcal{I}, \end{aligned}$$

which is exactly the requirement imposed by both ordered and unordered monotonicity. However, as demonstrated by Lemma 2, when there are multiple treatments the three monotonicity criterion are distinct.

---

[12] This is done by replacing $\mathbf{1}[T_i(z') = t']$ with $(1 - \mathbf{1}[T_i(z) = t'])$ on the left hand side and $\mathbf{1}[T_i(z) = t']$ with $(1 - \mathbf{1}[T_i(z) = t])$ on the right hand side. Afterwards, distribute and simplify.

We can gain further interpretation of the monotonicity restrictions by examining the relation between the verification matrices $\Psi_{\boldsymbol{U}}, \Psi_{\boldsymbol{O}}$, and $\Psi_{\boldsymbol{M}}$ of Theorems 1, 2 and 3. We express the verification matrices in terms of a primitive component defined by:

$$\Psi(t, t', t'', t''') \equiv \left(\boldsymbol{B}_t^{\intercal} \boldsymbol{B}_{t'}\right) \odot \left(\boldsymbol{B}_{t''}^{\intercal} \boldsymbol{B}_{t'''}\right). \tag{23}$$

$\Psi(t, t', t'', t''')$ is a function of four binary matrices $(\boldsymbol{B}_t, \boldsymbol{B}_{t'}, \boldsymbol{B}_{t''}, \boldsymbol{B}_{t'''})$ that returns a primitive verification matrix of dimension $N_Z \times N_S$ whose elements are either zeros or natural numbers.[13] Under this notation, the verification matrix $\Psi_{\boldsymbol{M}}$ can be expressed as:

$$\Psi_{\boldsymbol{M}} = \sum_{t \neq t'} \Psi(t, t', t', t). \tag{24}$$

Equation (24) explains the content of the verification matrix $\Psi_{\boldsymbol{M}}$. Theorem 3 states that MM (18) holds if and only if $\|\Psi_{\boldsymbol{M}}\| = 0$. By definition the matrix $\Psi_{\boldsymbol{M}}$ is the sum of the primitive verification matrices $\Psi(t, t', t', t)$ across all $N_T \cdot (N_T - 1)$ binary combinations of two distinct treatment choices $t, t' \in \mathcal{T}$. The elements of the primitive verification matrices are weakly positive and so $\|\Psi_{\boldsymbol{M}}\| = 0$ if and only if $\|\Psi(t, t', t', t)\| = 0$ for all distinct treatment values t and t'. Thus a necessary and sufficient condition for MM to hold is that each primitive verification matrix $\Psi(t, t', t', t)$ contains only zero elements for all $t, t' \in \mathcal{T}$ such that $t \neq t'$. Indeed, $\|\Psi(t, t', t', t)\| = 0$ is a equivalent to the nonexistence of any 2×2 submatrix of the response matrix $\boldsymbol{R}$ is of the form $\left(\begin{smallmatrix} t & t' \\ t' & t \end{smallmatrix}\right)$.

Theorem 4 relates the UM verification matrix to the primitive verification matrices and the MM verification matrix.

Theorem 4 (Decomposing Unordered Verification)**.** *The following relation holds for any response matrix* $\boldsymbol{R} \in \mathcal{T}^{N_Z \times N_S}$:

$$\|\Psi_{\boldsymbol{U}}\| = 0 \quad \Leftrightarrow \quad \|\Psi_{\boldsymbol{M}}\| + \|\Psi_{\boldsymbol{U} \setminus \boldsymbol{M}}\| = 0, \tag{25}$$

$$\text{where} \quad \Psi_{\boldsymbol{U} \setminus \boldsymbol{M}} \equiv \sum_{t \neq t' \neq t''} \Psi(t, t', t'', t). \tag{26}$$

---

[13] The output of function Ψ remains the same if first two inputs can be switched by the last two inputs $\Psi(t, t', t'', t''') = \Psi(t'', t''', t, t')$. We also have that $\Psi(t, t', t'', t''')^{\intercal} = \Psi(t', t, t''', t'')$, namely, the transpose of $\Psi(t, t', t'', t''')$ is equal to the matrix $\Psi(t', t, t''', t'')$ in which we switch t by t' and t'' by t'''.

Proof. See Appendix A $\square$

Lemma 2 explains that UM imposes extra constraints in addition to those required for MM to hold. Theorem 4 clarifies these additional constraints. Equation (25) decomposes the UM verification ($\|\Psi_{\boldsymbol{U}}\| = 0$) into two verification requirements. The first verification criterion, $\|\Psi_{\boldsymbol{M}}\| = 0$, means that MM must hold. The additional constraint, $\|\Psi_{\boldsymbol{U}\setminus\boldsymbol{M}}\| = 0$, means that the elements of the matrix $\Psi(t, t', t'', t)$ must be zero for any selection of three distinct treatment choices $t, t', t'' \in \mathcal{T}$. This additional criterion rules out violations of the prohibited pattern in item *(iii)* of Theorem 1 that involve three distinct treatment values.

Theorem 4 offers a combinatorial interpretation of monotonicity conditions MM and UM. MM imposes $\binom{N_T}{2}$ constraints on the primitive verification matrix $\Psi(t, t', t', t)$ across the combination of treatment choices taken *two* at a time. UM imposes an additional $\binom{N_T}{3}$ constraints on $\Psi(t, t', t'', t)$ across the combination of treatment choices taken *three* at a time.

Theorem 5 decomposes the verification criterion of ordered monotonicity ($\|\Psi_{\boldsymbol{O}}\| = 0$) into the verification criterion of the MM condition, $\|\Psi_{\boldsymbol{M}}\| = 0$, and two other criteria corresponding to matrices $\Psi^{(1)}_{\boldsymbol{O}\setminus\boldsymbol{M}}$ and $\Psi^{(2)}_{\boldsymbol{O}\setminus\boldsymbol{M}}$. The requirement that $\|\Psi_{\boldsymbol{M}}\| = 0$ means that OM implies MM, as in Lemma 2. Matrix $\Psi^{(1)}_{\boldsymbol{O}\setminus\boldsymbol{M}}$ contains a subset of the unordered monotonicity constrains in matrix $\Psi_{\boldsymbol{U}\setminus\boldsymbol{M}}$ of Theorem 4. Matrix $\Psi^{(2)}_{\boldsymbol{O}\setminus\boldsymbol{M}}$ contains constrains that are not in $\Psi_{\boldsymbol{U}\setminus\boldsymbol{M}}$. This is as expected since OM does not imply nor is implied by UM.

Theorem 5 (Decomposing Ordered Verification)**.** *The following relation holds for any response matrix* $\boldsymbol{R} \in \mathcal{T}^{N_Z \times N_S}$:

$$\|\Psi_{\boldsymbol{O}}\| = 0 \quad \Leftrightarrow \quad \|\Psi_{\boldsymbol{M}}\| + \|\Psi^{(1)}_{\boldsymbol{O}\setminus\boldsymbol{M}}\| + \|\Psi^{(2)}_{\boldsymbol{O}\setminus\boldsymbol{M}}\| = 0, \tag{27}$$

where

$$\Psi^{(1)}_{\boldsymbol{O}\setminus\boldsymbol{M}} \equiv \sum_{t_1 < min(t_2, t_3)} \Psi(t_1, t_2, t_3, t_1),$$

$$\Psi^{(2)}_{\boldsymbol{O}\setminus\boldsymbol{M}} \equiv \sum_{t_1 < t_2 \leq t_3} \Psi(t_1, t_3, t_2, t_2) + \sum_{t_1 < t_2 < t_3} \Psi(t_1, t_3, t_3, t_2) + \sum_{t_4 < t_2, t_1 < t_3} \Psi(t_1, t_2, t_3, t_4).$$

Proof. See Appendix A $\square$

6 An Economic Interpretation for Monotonicity Conditions

This section explores the economic content of the monotonicity criteria. We show that the minimal monotonicity condition (5.3) can be linked to a broad notion of rationality regarding treatment choices. This contrasts to ordered and unordered monotonicity, which as seen in Theorems 1 and 2 are equivalent to assuming particular ordered and unordered choice models.

Our analysis is based on the method of Pinto (2021); Buchinsky and Pinto (2021) who use revealed preference analysis to ascribe economic interpretation to response matrices. The method uses the concept of an incentive matrix $\boldsymbol{L}$ that characterizes the choice incentives (columns) generated by the IV-values (rows). Each column $\boldsymbol{L}[\cdot, t]$ displays the relative ranking of incentives towards choice $t \in \text{supp}(t)$ across the IV-values $z \in \mathcal{Z}$. $\boldsymbol{L}[z', t] < \boldsymbol{L}[z, t]$ means that the IV-value z yields strictly greater incentives towards t than IV-value z'. The matrix is ordinal, monotonic transformations characterize equivalent choice incentives.

To fix ideas, consider the binary LATE model of Sections 4 where T denotes college enrollment; $T = t_1$ for college enrollment and $T_i = t_0$ otherwise. The instrument Z denotes a randomly assigned tuition discount such that $Z = z_1$ if the discount is granted and $Z = z_0$ otherwise. Incentive matrix (28) characterizes the choice incentives of the LATE model. $\boldsymbol{L}[z_0, t_0] = \boldsymbol{L}[z_1, t_0] = 0$ means that the voucher offers no incentives for choice t_0 (no college). $\boldsymbol{L}[z_0, t_1] < \boldsymbol{L}[z_1, t_1]$ indicates that the tuition discount z_1 incentives college enrollment t_1.

$$\text{LATE Incentive Matrix} \quad \boldsymbol{L} = \begin{array}{c} \begin{array}{cc} t_0 & t_1 \end{array} \\ \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \end{array} \begin{array}{c} \\ z_0 \\ z_1 \end{array} \tag{28}$$

Pinto (2021); Buchinsky and Pinto (2021) use revealed preference analysis to translate the incentive matrix into choice restrictions. They invoke the Weak Axiom of Revealed Preferences (WARP) and

Choice Normality to generate the following choice rule:[14]

$$\text{Choice Rule: } T_i(z) = t \ \text{ and } \ \underbrace{\boldsymbol{L}[z',t'] - \boldsymbol{L}[z,t'] \leq \boldsymbol{L}[z',t] - \boldsymbol{L}[z,t]}_{\substack{\text{Switch from } z \text{ to } z' \text{ provides greater} \\ \text{incentives for } t \text{ than } t'}} \implies T_i(z') \neq t' \qquad (29)$$

Choice Rule (29) formalizes an intuitive behavioral restriction. If an agent i chooses t when exposed to instrument z, and the IV-shift from z to z' yields greater incentives towards t than t', then agent i does not choose t' under z'. Otherwise stated, each instrumental value z is associated with an incentive gap between t and t'. If an agent decides for choice t given z, then t is revealed preferred to t'. The agent should t' only if the incentive gap between t and t' increases.

Applying choice rule (29) to the LATE incentive matrix (28) generates the following *choice restriction*:

$$T_i(z_0) = t_1 \ \text{ and } \ \underbrace{\boldsymbol{L}[z_1,t_0] - \boldsymbol{L}[z_0,t_0]}_{=0} \leq \underbrace{\boldsymbol{L}[z_1,t_1] - \boldsymbol{L}[z_0,t_1]}_{=1} \implies T_i(z_1) \neq t_0. \qquad (30)$$

Choice restriction (30) is summarized by $T_i(z_0) = t_1 \Rightarrow T_i(z_1) \neq t_0$. It states that if an agent chooses to attend college when not offered any incentives ($T_i(z_0) = t_1$) they must also chooses choose to attend college when offered the tuition discount ($T_i(z_1) = t_1$). The restriction is equivalent to the monotonicity condition of Imbens and Angrist (1994), discussed in Section 3, which eliminates the defiers and enables the identification of LATE.

Buchinsky and Pinto (2021) characterize the class of incentive matrices that produce OM and UM conditions. For instance, they demonstrate that choice incentives characterized by lonesum matrices produce unordered choice models while increasing incentives such as those described by a Vandermonde matrix produce ordered choice models.

Incentives and Minimal Monotonicity

It is natural to inquire about which types of incentive designs ensure the MM condition. It turns out that the Choice Rule (29) itself ensures the MM condition. Theorem 6 asserts that the MM condition always arises whenever we apply the revealed preference analysis encoded by the choice

---

[14] [Define both] The choice rule would have a strict inequality if we were to assume WARP only.

rule to any choice incentives.

Theorem 6. *The minimal monotonicity condition* (18) *holds for all choice models generated by applying Choice Rule* (29) *to an arbitrary Incentive Matrix* $\boldsymbol{L}$.

Theorem 6 draws a sharp distinction in the interpretation of the monotonicity conditions. In essence, OM and UM can be interpreted as monotonicity conditions that arise when agents that display a rational behavior face a particular a class of choice incentives. This paradigm does not apply to MM, since MM is not a property ascribed to any particular pattern of incentives. Instead, MM is a supra-condition that can be justified by a weak notion of rationality itself.

With this in mind, the MM condition can be seen not as a final goal, but rather a starting point for generating and interpreting monotonicity criteria. To be more precise, minimal monotonicity ensures that a range of monotonicity conditions obtained by combining a broad notion of choice rationality with specific choice incentives satisfy particular basic properties, including interpretability of 2SLS-type estimands. Section 7 uses this insight to illustrate the flexibility of the MM condition in empirical analysis.

7 Economic Examples of Monotonicity Conditions

We present several examples of response matrices generated by combining specific incentive structures with the choice rule (29) defined in Section 6. The first two examples demonstrate specific incentive designs that generate response matrices satisfying unordered and ordered monotonicity, respectively. Afterwards, we present some natrual research designs that generate response matrices that do not comply with either ordered or unordered monotonicity. However, by Theorem 6, these response matrices still satisfy the minimal monotonicity condition. We discuss how minimal monotonicity may still enable the researcher to gain insight into causal effects under these research designs. Our minimal monotonicity examples include the popular Extensive Margin Compliers (EMCO) of Angrist and Imbens (1995) and a double RCT design. In all examples we consider incentive designs for a three-valued treatment choice $\mathcal{T} = \{t_1, t_2, t_3\}$ and four instrumental values $\mathcal{Z} = \{z_1, z_2, z_3, z_4\}$.

7.1 A Case of Choice Incentives that Justify Unordered Monotonicity

In this example, let T denote the student's decision among college majors: t_1 for humanities, t_2 for social sciences, and t_3 for the STEM fields of science, technology, engineering, and math. The instrumental variable Z represents a randomly assigned vouchers that offers a tuition discount that may apply to one, several or none of the majors. For example, consider the social experiment that randomly assigns one of the four vouchers z_1, z_2, z_3, z_4 to college students:

1. Voucher z_1 offers no tuition discount.
2. Voucher z_2 applies only to STEM (t_3).
3. Voucher z_3 applies to either STEM (t_3) or social sciences (t_2).
4. Voucher z_4 applies to all majors.

Assuming that the voucher amount is always the same, this design is characterized by incentive matrix $\boldsymbol{L}$ in (31).[15]

$$\boldsymbol{L} = \begin{array}{c} \\ \\ \\ \\ \end{array}\begin{array}{ccc|c} t_1 & t_2 & t_3 & \\ 0 & 0 & 0 & z_1 \\ 0 & 0 & 1 & z_2 \\ 0 & 1 & 1 & z_3 \\ 1 & 1 & 1 & z_4 \end{array} \tag{31}$$

We use this first example to describe the machinery that translates choice incentives into monotonicity conditions and identification results. We adopt a more parsimonious approach in the subsequent examples. We place detailed derivations in Appendix D.

Choice rule (29) converts the Incentive Matrix (31) into choice restrictions that determine the model response matrix $\boldsymbol{R}$. The choice rule applies to any two instrument-treatment pairs; $((z,t),(z',t')) \in (\mathcal{Z}\times\mathcal{T})^2$. To exemplify how this is done, Table 1 displays all the restrictions generated by applying Choice Rule (29) to an agent who chooses a humanities major (t_1) when offered no tuition discount (z_1). The first row of Table 1 applies the choice rule to the two instrument-treatment pairs, (z_1, t_1) and (z_2, t_2). By applying the choice rule, we see that an agent i who chooses a humanties major (t_1) when offered no tuition discount (z_1) would not chose treatment a social sciences major (t_2)

---

[15] Elements one indicate the presence of incentive (the tuition discount) while elements zero indicate the lack of it.

when offered instrument a tuition discount only for STEM majors (z_2). The incentives for choosing either t_1 or t_2 remain the same when the IV switches from z_1 to z_2. The incentive inequality in (29) is satisfied and the choice restriction $T_i(z_1) \neq t_2$ holds.

The other lines of Table 1 apply this same logic to all other instrument-treatment pairs. We can see that, in total, under the incentive structure summarized in (31) the Choice Rule (29) places the following restrictions on an agent who chooses a humanities major when offered no tuition discount

$$T_i(z_1) = t_1 \implies T_i(z_2) \neq t_2 \text{ and } T_i(z_4) \notin \{t_2, t_3\}$$

This analysis can be repeated for all types of agents, the sum total of all choice restrictions generated by applying the choice rule in this research design are presented in Table 2. All choice restrictions of Table 2 eliminate a total of 74 out of 81 possible response-types. The seven response-types that survive this elimination procedure are presented as columns of the following response matrix, $\boldsymbol{R}$, in (32):

Table 1: Applying Choice Rule (29) to $T_i(z_1) = t_1$ and Incentive Matrix (31)

Counterfactual Choice	Incentive Condition		Choice Restriction
$T(z_1) = t_1$	$\boldsymbol{L}[z_2, t_2] - \boldsymbol{L}[z_1, t_2] = 0 \leq 0 = \boldsymbol{L}[z_2, t_1] - \boldsymbol{L}[z_1, t_1]$	$\Rightarrow$	$T(z_2) \neq t_2$
$T(z_1) = t_1$	$\boldsymbol{L}[z_2, t_3] - \boldsymbol{L}[z_1, t_3] = 1 \not\leq 0 = \boldsymbol{L}[z_2, t_1] - \boldsymbol{L}[z_1, t_1]$	$\Rightarrow$	No Restriction
$T(z_1) = t_1$	$\boldsymbol{L}[z_3, t_2] - \boldsymbol{L}[z_1, t_2] = 1 \not\leq 0 = \boldsymbol{L}[z_3, t_1] - \boldsymbol{L}[z_1, t_1]$	$\Rightarrow$	No Restriction
$T(z_1) = t_1$	$\boldsymbol{L}[z_3, t_3] - \boldsymbol{L}[z_1, t_3] = 1 \not\leq 0 = \boldsymbol{L}[z_3, t_1] - \boldsymbol{L}[z_1, t_1]$	$\Rightarrow$	No Restriction
$T(z_1) = t_1$	$\boldsymbol{L}[z_4, t_2] - \boldsymbol{L}[z_1, t_2] = 1 \leq 1 = \boldsymbol{L}[z_4, t_1] - \boldsymbol{L}[z_1, t_1]$	$\Rightarrow$	$T(z_4) \neq t_2$
$T(z_1) = t_1$	$\boldsymbol{L}[z_4, t_3] - \boldsymbol{L}[z_1, t_3] = 1 \leq 1 = \boldsymbol{L}[z_4, t_1] - \boldsymbol{L}[z_1, t_1]$	$\Rightarrow$	$T(z_4) \neq t_3$

This table presents all the choice restrictions generated by applying the choice rule (29) to each of the tuples $((z_1, t_1), (z', t'))$ where $z' in\{z_2, z_3, z_4\}$ and $t' in\{t_2, t_3, t_4\}$ according to the choice incentives displayed in the incentive matrix (31).

$$\boldsymbol{R} = \begin{array}{c} \\ \\ \\ \\ \end{array}\begin{array}{ccccccc|c} \boldsymbol{s}_1 & \boldsymbol{s}_2 & \boldsymbol{s}_3 & \boldsymbol{s}_4 & \boldsymbol{s}_5 & \boldsymbol{s}_6 & \boldsymbol{s}_7 & \\ t_1 & t_1 & t_1 & t_1 & t_2 & t_2 & t_3 & z_1 \\ t_1 & t_1 & t_3 & t_3 & t_2 & t_3 & t_3 & z_2 \\ t_1 & t_2 & t_2 & t_3 & t_2 & t_2 & t_3 & z_3 \\ t_1 & t_1 & t_1 & t_1 & t_2 & t_2 & t_3 & z_4 \end{array} \tag{32}$$

Using the characterizations in Theorem 1, we can verify that unordered monotonicity holds for the response matrix $\boldsymbol{R}$ presented in (32). Notice that there is no 2×2 matrix of the matrix is of the

Table 2: Choice Restrictions generated by Incentive Matrix (31)

1	$T_i(z_1) = t_1$	$\Rightarrow$	$T_i(z_2) \neq t_2$ and $T_i(z_4) \notin \{t_2, t_3\}$
2	$T_i(z_2) = t_1$	$\Rightarrow$	$T_i(z_1) \notin \{t_2, t_3\}$ and $T_i(z_3) \neq t_3$ and $T_i(z_4) \notin \{t_2, t_3\}$
3	$T_i(z_3) = t_1$	$\Rightarrow$	$T_i(z_1) \notin \{t_2, t_3\}$ and $T_i(z_2) \notin \{t_2, t_3\}$ and $T_i(z_4) \notin \{t_2, t_3\}$
4	$T_i(z_4) = t_1$	$\Rightarrow$	$T_i(z_1) \notin \{t_2, t_3\}$ and $T_i(z_2) \neq t_2$
5	$T_i(z_1) = t_2$	$\Rightarrow$	$T_i(z_2) \neq t_1$ and $T_i(z_3) \notin \{t_1, t_3\}$ and $T_i(z_4) \notin \{t_1, t_3\}$
6	$T_i(z_2) = t_2$	$\Rightarrow$	$T_i(z_1) \notin \{t_1, t_3\}$ and $T_i(z_3) \notin \{t_1, t_3\}$ and $T_i(z_4) \notin \{t_1, t_3\}$
7	$T_i(z_3) = t_2$	$\Rightarrow$	$T_i(z_1) \neq t_3$ and $T_i(z_4) \neq t_3$
8	$T_i(z_4) = t_2$	$\Rightarrow$	$T_i(z_1) \notin \{t_1, t_3\}$ and $T_i(z_2) \neq t_1$ and $T_i(z_3) \notin \{t_1, t_3\}$
9	$T_i(z_1) = t_3$	$\Rightarrow$	$T_i(z_2) \notin \{t_1, t_2\}$ and $T_i(z_3) \notin \{t_1, t_2\}$ and $T_i(z_4) \notin \{t_1, t_2\}$
10	$T_i(z_2) = t_3$	$\Rightarrow$	$T_i(z_3) \neq t_1$
11	$T_i(z_3) = t_3$	$\Rightarrow$	$T_i(z_1) \neq t_2$ and $T_i(z_2) \notin \{t_1, t_2\}$ and $T_i(z_4) \neq t_2$
12	$T_i(z_4) = t_3$	$\Rightarrow$	$T_i(z_1) \notin \{t_1, t_2\}$ and $T_i(z_2) \notin \{t_1, t_2\}$ and $T_i(z_3) \notin \{t_1, t_2\}$

This table presents all the choice restrictions generated by applying the choice rule (29) to each of the tuples $((z, t), (z', t')) \in (\{t_1, t_2, t_3\} \times \{z_1, z_2, z_3, z_4\})^2$. according to the incentive matrix (31).

type of the form $\begin{pmatrix} t & t' \\ t'' & t \end{pmatrix}$ where $t' \neq t$, or $t'' \neq t$. Appendix D.1 corroborates the UM property using the verification matrix from item (*iv*) of Theorem 1.

While the incentive design described in (31) is plausible and the application of the choice rule a minimal behavioral requirement on the agents, it is helpful to remark on how this specific incentive structure generates unordered monotonicity. Note that the incentive structure increases in the sense that each successive instrument provides weakly more incentives for all treatment choices. No change in the instrument from z to z' would strictly decrease incentives for one treatment while strictly increasing incentives for another.

This property is crucial for generating unordered monotonicity. Under WARP, if an agent would choose treatment t under instrument value z, they must also choose treatment t under instrument value z' whenever the switch from z to z' weakly increases incentives for t relative to all other incentives. Because of the increasing nature of the incentive structure, each switch in the instrument value either increases the incentives for a choice t relative to *all* other treatments or decreases the incentives for the choice t relative to all other treatments. This, along with the choice rule, prevents an instrumental switch from moving one agent strictly towards choosing t while moving another agent strictly away from choice t and towards another treatment t'.

7.2 A Case of Choice Incentives that Justify Ordered Monotonicity

In this example, suppose the CEO of a company is interested in whether higher health insurance premiums lead to moral hazard in employees' safety behavior. Employees decide among three health insurance policies t_1, t_2, t_3 that have increasing premiums. The co-pay of each policy off-sets the increasing premium such that all policies cost the same.

To study this, the CEO randomly assigns agents to one of four groups, z_1, z_2, z_3, z_4, that incentivize (say by offering an additional week of vacation) various insurance plan options. We consider the following scheme of choice incentives:

1. Group z_1 is incentivized to choose treatment t_1.
2. Group z_2 is offered no incentives.
3. Group z_3 is offered incentives for all choices.
4. Group z_4 is incentivized to choose treatment t_3.

Equation (33) presents the incentive matrix $\boldsymbol{L}$ that characterizes the design of choice incentives.[16] This incentive design is rather peculiar because it is tailored to generate the OM criteria. Equation (33) also presents the corresponding response matrix $\boldsymbol{R}$ generated by the method of revealed preference analysis described in Section 7.1. Detailed derivations are presented in Appendix D.2.

$$\boldsymbol{L} = \begin{array}{cccc} t_1 & t_2 & t_3 & \\ 1 & 0 & 0 & z_1 \\ 0 & 0 & 0 & z_2 \\ 1 & 1 & 1 & z_3 \\ 0 & 0 & 1 & z_4 \end{array} \Rightarrow \boldsymbol{R} = \begin{array}{ccccccccc} \boldsymbol{s}_1 & \boldsymbol{s}_2 & \boldsymbol{s}_3 & \boldsymbol{s}_4 & \boldsymbol{s}_5 & \boldsymbol{s}_6 & \boldsymbol{s}_7 & \boldsymbol{s}_8 & \\ t_1 & t_1 & t_1 & t_1 & t_1 & t_2 & t_2 & t_3 & z_1 \\ t_1 & t_1 & t_2 & t_2 & t_3 & t_2 & t_2 & t_3 & z_2 \\ t_1 & t_1 & t_2 & t_2 & t_3 & t_2 & t_2 & t_3 & z_3 \\ t_1 & t_3 & t_2 & t_3 & t_3 & t_2 & t_3 & t_3 & z_4 \end{array} \tag{33}$$

Using Theorem 2 it is easy to check that OM holds for response matrix (33). The indices of the treatment choices weakly increase as the instrument ranges along $z_1 \to z_2 \to z_3 \to z_4$. Ordered monotonicity is satisfied by assigning treatment values that satisfy $t_1 \leq t_2 \leq t_3$. By applying the identification results of Angrist and Imbens (1995), one can verify that the 2SLS has the causal interpretation of a weighted average of LATEs of the type $E(Y(t_{k+1}) - Y(t_k)|\boldsymbol{S});\; k \in \{1,2\}$.

---

[16] Elements one indicate the presence of incentive (an additional one week vacation) while elements zero indicate the lack of it.

Similarly via Theorem 1 it is also easy to verify that UM does not hold for response matrix (33). The 2×2 submatrix of rows (z_1, z_4) and columns $(\boldsymbol{s}_3, \boldsymbol{s}_7)$ displays the values $\left(\begin{smallmatrix} t_1 & t_2 \\ t_2 & t_3 \end{smallmatrix}\right)$ which violates item (iii) of Theorem 1; the shift of IV-values from z_1 to z_4 induces some agents towards choice t_2 while inducing others away from t_2.

Interestingly, the natural ranking on the treatment space is not important for generating ordered monotonicity in this way. We could just as easily have considered a treatment space with no natural ranking, such as the choice of neighborhood to live in. The incentive design summarized by $\boldsymbol{L}$ in (33) would have still generated the response matrix $\boldsymbol{R}$ that satisfies ordered monotonicity. This demonstrates the usefulness of considering the slightly broader characterization of ordered monotonicity presented in OM Sequence (10). Had ordered monotonicity been ruled out a priori, the researcher may not have been able to take advantage of the Angrist and Imbens (1995) identification results.

7.3 Beyond Ordered or Unordered Monotonicity

The MM condition provides a theoretical foundation for a wide range choice behaviors that do not exactly conform to the paradigm imposed by ordered or unordered choices. It offers the necessary flexibility to examine economic settings that are not neatly described by ordered or unordered monotonicity. We illustrate this fact in the following examples.

The Double Randomization Design

A basic inquiry in social science is to evaluate the causal effect of a treatment t_1 versus its absence. The standard IV experiment that would allow us to assess this effect is to randomly offer a voucher that incentives a set of agents to choose a treatment choice t_1. This experiment can be described by the binary LATE model discussed in Section 3.

A straightforward extension of this setup is to insert a second treatment t_2 and randomize a second voucher that incentives t_2 for the same set of agents. The combination of the two randomization runs generate four groups according to the voucher assignments. Notationally, our experiment consists of three choices $T \in \{t_0, t_1, t_2\}$, where t_0 represents not choosing either treatment t_1 or t_2, and four instrumental values $\{z_1, z_2, z_3, z_4\}$ that classify the voucher recipients into:

1. Group z_1 comprise agents that do not receive any voucher.
2. Group z_2 comprise agents that receive a voucher that incentivizes choice (t_2).
3. Group z_3 comprise agents that receive a voucher that incentivizes choice (t_1).
4. Group z_4 are those that receive two vouchers, one for t_1 and another for t_2.

Equation (34) presents an incentive matrix corresponding to this research design and the corresponding response matrix generated by the revealed preference analysis described in Section 7.1. See Appendix D.3 for detailed derivations.

$$\boldsymbol{L} = \begin{array}{cc} \begin{matrix} t_0 & t_1 & t_2 \end{matrix} & \\ \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} & \begin{matrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{matrix} \end{array} \Rightarrow \boldsymbol{R} = \begin{array}{cc} \begin{matrix} \boldsymbol{s}_1 & \boldsymbol{s}_2 & \boldsymbol{s}_3 & \boldsymbol{s}_4 & \boldsymbol{s}_5 & \boldsymbol{s}_6 & \boldsymbol{s}_7 & \boldsymbol{s}_8 & \boldsymbol{s}_9 \end{matrix} & \\ \begin{bmatrix} t_0 & t_0 & t_0 & t_0 & t_0 & t_1 & t_1 & t_2 & t_2 \\ t_0 & t_0 & t_2 & t_2 & t_2 & t_1 & t_2 & t_2 & t_2 \\ t_0 & t_1 & t_0 & t_1 & t_1 & t_1 & t_1 & t_1 & t_2 \\ t_0 & t_1 & t_2 & t_1 & t_2 & t_1 & t_1 & t_2 & t_2 \end{bmatrix} & \begin{matrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{matrix} \end{array} \tag{34}$$

Response matrix (34) satisfies neither UM nor OM. UM does not hold because the 2×2 submatrix generated by rows (z_2, z_3) and columns $(\boldsymbol{s}_2, \boldsymbol{s}_3)$ displays matrix $\left(\begin{smallmatrix} t_0 & t_2 \\ t_1 & t_0 \end{smallmatrix}\right)$, which violates item (iii) of Theorem 1. For the ordering $t_0 \leq t_1 \leq t_2$ on $\mathcal{T}$, columns $(\boldsymbol{s}_2, \boldsymbol{s}_3)$ also preclude OM. Under this ordering, no matter how we order the IV values, we cannot generate that the sequence of treatments is increasing for both $\boldsymbol{s}_2$ and $\boldsymbol{s}_3$. In fact, no matter which ordering we take on $\mathcal{T}$, we can always find a pair of response types whose treatment uptake patterns violate OM Sequence.[17]

However, as guaranteed by Theorem 6, the response matrix $\boldsymbol{R}$ in (34) satisfies minimal monotonicity (MM). This can be further verified via the minimal monotonicity characterization in Theorem 3 by checking that there is no 2×2 submatrix in (34) of the form $\left(\begin{smallmatrix} t & t' \\ t' & t \end{smallmatrix}\right)$. Appendix D.3 additionally corroborates this fact by evaluating the verification matrix $\Psi_{\boldsymbol{M}}$ from Theorem 3(iii).

Despite the fact that neither UM nor OM holds, we can still explore causal relationships in this choice model. In particular, the response matrix still enables us to identify causal parameters using 2SLS type estimands. For example, if the researcher was interested in the effect of treatment t_2 against alternate treatments, the 2SLS-type estimand

---

[17] Using the violation of unordered monotonicity with t_0 we have that t_0 cannot be ranked highest or lowest in any ordering that satisfies OM; this would mean that some agents increase their treatment as the instrument ranges from z_2 to z_3 while other agents decrease their treatment. We thus only have to consider the orderings $t_1 \leq t_0 \leq t_2$ and $t_2 \leq t_0 \leq t_1$, which can both be eliminated by considering the alternating patterns displayed by response types $\boldsymbol{s}_4$ and $\boldsymbol{s}_7$.

$$\beta_{z_4,z_3} \equiv E[Y \mid Z = z_4] - E[Y \mid Z = z_3].$$

recovers a weighted average of $E[Y(t_2) - Y(t_1) \mid \boldsymbol{S} \in \mathcal{S}_{2,1}]$ and $E[Y(t_2) - Y(t_0) \mid \boldsymbol{S} \in \mathcal{S}_{2,0}]$, with positive weights, for some sets of response types $\mathcal{S}_{2,1}$ and $\mathcal{S}_{2,0}$ that can be found using the decomposition in Appendix B. If one was alternatively interested in the effect of t_1 against alternate treatments, the 2SLS-type estimand

$$\beta_{z_3,z_1} \equiv E[Y \mid Z = z_3] - E[Y \mid Z = z_2]$$

recovers a weighted average of $E[Y(t_1) - Y(t_0) \mid \boldsymbol{S} \in \mathcal{S}_{1,0}]$ and $E[Y(t_1) - Y(t_2) \mid \boldsymbol{S} \in \mathcal{S}_{1,2}]$ for two alternate sets of response types $\mathcal{S}_{1,0}$ and $\mathcal{S}_{1,2}$.

Incentives that Justify Extensive Margin Complies Only

We next exemplify how the choice rationale can be used to justify monotonicity criteria that are more restrictive than UM and OM. Consider a group of students of a technical college that decide among three possible majors: computer science (t_1), electrical engineering (t_2), or mechanical engineering (t_3).

College administration perform a double randomization of two types of tuition vouchers. The first voucher offers a tuition discount for computer science (t_1) while the other for the engineering courses (t_2 or t_3). Students can be divided into four groups according to the voucher assignment:

1. Group z_1 receives no voucher;
2. Group z_2 receives the voucher for computer science (t_1) only;
3. Group z_3 receives the voucher that incentivizes electrical (t_2) or mechanical (t_3) engineering.
4. Group z_4 receives both vouchers which offers incentives to all three choices.

Equation (35) presents the incentive matrix associated with this experimental design and its corresponding response matrix. See Appendix D.4 for derivation details.

$$\boldsymbol{L} = \begin{array}{c} \\ \\ \\ \\ \end{array}\!\!\begin{array}{ccc|c} t_1 & t_2 & t_3 & \\ \hline 0 & 0 & 0 & z_1 \\ 1 & 0 & 0 & z_2 \\ 0 & 1 & 1 & z_3 \\ 1 & 1 & 1 & z_4 \end{array} \Rightarrow \boldsymbol{R} = \begin{array}{ccccccc|c} \boldsymbol{s}_1 & \boldsymbol{s}_2 & \boldsymbol{s}_3 & \boldsymbol{s}_4 & \boldsymbol{s}_5 & \boldsymbol{s}_6 & \boldsymbol{s}_7 & \\ \hline t_1 & t_1 & t_1 & t_2 & t_2 & t_3 & t_3 & z_1 \\ t_1 & t_1 & t_1 & t_1 & t_2 & t_1 & t_3 & z_2 \\ t_1 & t_2 & t_3 & t_2 & t_2 & t_3 & t_3 & z_3 \\ t_1 & t_1 & t_1 & t_2 & t_2 & t_3 & t_3 & z_4 \end{array} \tag{35}$$

Response matrix (35) is an example where both OM and UM are satisfied. We can check that OM holds by assigning values $(1,2,3)$ to (t_1,t_2,t_3) and reordering the IV-values from z_1,z_2,z_3,z_4 to z_2,z_1,z_4,z_3. The resulting response matrix is presented in (36) which shows that treatment values weakly increase as Z ranges along its values. It is easy to check that each of the binary matrices $\boldsymbol{B}_t = \mathbf{1}[\boldsymbol{R}=t]; t \in \{1,2,3\}$ that indicate the treatment choices of response matrix (36) is triangular (i.e. lonesum). This implies that UM holds by item (iv) of Theorem 1.

$$\text{Reordered } \boldsymbol{R} = \begin{array}{cccccccc} \boldsymbol{s}_1 & \boldsymbol{s}_2 & \boldsymbol{s}_3 & \boldsymbol{s}_4 & \boldsymbol{s}_5 & \boldsymbol{s}_6 & \boldsymbol{s}_7 & \\ 1 & 1 & 1 & 1 & 2 & 1 & 3 & z_2 \\ 1 & 1 & 1 & 2 & 2 & 3 & 3 & z_1 \\ 1 & 1 & 1 & 2 & 2 & 3 & 3 & z_4 \\ 1 & 2 & 3 & 2 & 2 & 3 & 3 & z_3 \end{array} \tag{36}$$

Response matrix (35) has a special property beyond UM and OM: each of its compliers takes only two treatment values, one of them being t_1. Specifically, the matrix has four response-types that display a variation of treatment choice, these are the compliers $(\boldsymbol{s}_2,\boldsymbol{s}_3,\boldsymbol{s}_4,\boldsymbol{s}_6)$. The choice values of response-types $\boldsymbol{s}_2,\boldsymbol{s}_4$ are t_1 or t_2 and the choice values of $\boldsymbol{s}_3,\boldsymbol{s}_6$ are t_1 or t_3 This special property is called Extensive Margin Compliers Only (EMCO) which is formalized in (37).[18]

EMCO: There exists a treatment choice $t_1 \in \mathcal{T}$ such that for any $z,z' \in \text{supp}(Z)$ we have that

$$T_i(z) \neq T_i(z') \Rightarrow T_i(z) = t_1 \text{ or } T_i(z') = t_1 \text{ for all } i \in \mathcal{I} \tag{37}$$

EMCO (37) simplifies the multiple-choice decision of compliers into a binary decision that debates between choosing t_1 or not. In our example, compliers $\boldsymbol{s}_2,\boldsymbol{s}_4$ debate between choosing computer science t_1 or electrical engineering t_2 while compliers $\boldsymbol{s}_3,\boldsymbol{s}_6$ debate between computer science t_1 or mechanical engineering t_3. None of the compliers debate between electrical or mechanical engendering. Instead, they decide between choosing computer science or not.

EMCO 37 enable us to recode the multiple choice $T_i \in \{t_1,t_2,t_3\}$ into a binary choice $D_i = \mathbf{1}[T_i = t_1]$ that indicates if the agent i chooses t_1. The 2SLS regression that uses the binary indicator as the endogenous treatment evaluates a weighted average of LATE-type effects between choosing t_1 and not across compliers.

---

[18] See Rose and Shem-Tov (2021); Angrist and Imbens (1995).

In particular, the comparison between two IV-values identifies the causal effect for of choosing t_1 versus not choosing t_1 for a sub-set of compliers. For instance, consider the IV-values z_1 and z_2. We can use equation (3) to obtain the following identification result:

$$\frac{E(Y|Z=z_2)-E(Y|Z=z_1)}{P(T=t_2|Z=z_2)-P(T=t_2|Z=z_1)} = \tag{38}$$

$$\frac{E(Y(t_1)-Y(t_2)|\boldsymbol{S}=\boldsymbol{s}_4)P(\boldsymbol{S}=\boldsymbol{s}_4)+E(Y(t_1)-Y(t_3)|\boldsymbol{S}=\boldsymbol{s}_6)P(\boldsymbol{S}=\boldsymbol{s}_6)}{P(\boldsymbol{S}=\boldsymbol{s}_4)+P(\boldsymbol{S}=\boldsymbol{s}_6)}. \tag{39}$$

Equations (38)–(39) show that the comparison between IV-values z_1 and z_2 identifies the causal effect of choosing t_1 versus not choosing t_1 conditional on response-types $\boldsymbol{s}_4, \boldsymbol{s}_6$. The equations are similar to the LATE identification equation of Imbens and Angrist (1994). They imply that we can evaluate the causal effect via the 2SLS regression that uses the sub-sample of agents assigned to z_1 and z_2.

Orthogonal Array Design

We additionally examine an IV choice model based on the orthogonal array experimental design. Orthogonal arrays are a widely popular experimental design developed by C.D. Rao (Rao, 1946a,b, 1947, 1949). Orthogonal arrays are widely used in Agricultural and Industrial sciences to determine the optimum mix of treatments that maximize production yield. The method is based on the random assignment of a combinatorial arrangements of treatments for each randomization arm. We adapt this setup to an instrumental variable setting by exogenously providing incentives for one or more treatments instead of directly assigning agents to treatment arms. Below, we will see that this incentive structure allows for a broad range of identification results.

Formally, a binary orthogonal array is a matrix of zeros and ones such that any two-column submatrix displays all possible combinations of zeros and ones. In other words, the tuples

$$\{(0,0),(0,1),(1,0),(1,1)\}$$

are all rows in any two-column submatrix. An orthogonal array incentive design if its associated incentive matrix is a binary orthogonal array. The incentive matrix in (40) displays an example of an orthogonal array incentive design and the corresponding response matrix generated by applying Choice Rule (29).

$$\boldsymbol{L} = \begin{array}{cc} & \begin{array}{ccc} t_1 & t_2 & t_3 \end{array} \\ & \left[\begin{array}{ccc} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{array}\right] \begin{array}{c} z_1 \\ z_2 \\ z_3 \\ z_4 \end{array} \end{array} \Rightarrow \boldsymbol{R} = \begin{array}{c} \begin{array}{ccccccccc} \boldsymbol{s}_1 & \boldsymbol{s}_2 & \boldsymbol{s}_3 & \boldsymbol{s}_4 & \boldsymbol{s}_5 & \boldsymbol{s}_6 & \boldsymbol{s}_7 & \boldsymbol{s}_8 & \boldsymbol{s}_9 \end{array} \\ \left[\begin{array}{ccccccccc} t_1 & t_2 & t_2 & t_2 & t_2 & t_3 & t_3 & t_3 & t_3 \\ t_1 & t_1 & t_2 & t_2 & t_2 & t_1 & t_3 & t_3 & t_3 \\ t_1 & t_1 & t_2 & t_2 & t_2 & t_1 & t_1 & t_2 & t_3 \\ t_1 & t_1 & t_1 & t_2 & t_3 & t_1 & t_3 & t_3 & t_3 \end{array}\right] \begin{array}{c} z_1 \\ z_2 \\ z_3 \\ z_4 \end{array} \end{array} \tag{40}$$

This response matrix satisfies neither unordered nor ordered monotonicity. When the instrument switches from z_1 to z_4, agents in response type $\boldsymbol{s}_5$ move from treatment t_2 to treatment t_3 while agents in response type $\boldsymbol{s}_6$ move away from t_3 and towards t_1. This represents a violation of ordered monotonicity and also prevents t_3 from being ordered the highest or lowest in any ordering on $\mathcal{T}$ that would satisfy ordered monotonicity.[19] Similarly we can see a switch from z_3 to z_4 induces agents in response type $\boldsymbol{s}_3$ to move from treatment t_2 to treatment t_1 while inducing agents in response type $\boldsymbol{s}_7$ to move away from treatment t_1 and towards treatment t_3. This again represents a violation of unordered monotonicity and prevents t_1 from being ordered either the highest or the lowest in any ordering $\mathcal{T}$ that would satisfy ordered monotonicity. Since all orderings on $\mathcal{T} = \{t_1, t_2, t_3\}$ must have either t_1 or t_3 as the largest or smallest element, this means there is no ordering on $\mathcal{T}$ that satisfies ordered monotonicity.

Again, however, Theorem 6 guarantees that the response matrix satisfies minimal monotonicity (MM).

8 Conclusion

Analysis of ordered and unordered IV choice models has largely been conducted in parallel, with little overlap between the two strands of the literature. Ordered choice models are commonly analyzed using ordered monotonicity (OM), introduced by Angrist and Imbens (1995), while unordered choice models are commonly analyzed using the unordered monotonicity (UM) of Heckman and Pinto (2018).

---

[19] If t_3 is ranked highest a movement away from t_3 represents moving towards a lower treatment while a towards t_3 represents moving towards a higher treatment. Vice versa, if t_3 is ranked lowest a movement towards t_3 represents moving towards a lower treatment while a movement away from t_3 represents moving towards a higher treatment.

This paper bridges the gap between analysis of ordered and unordered IV choice models. We note symmetric features of ordered and unordered monotonicity and use them to derive symmetric characterizations of the two. The symmetric characterizations offer deep insights into the relationship between the two monotonicity criterion. Moreover they provide computationally tractable ways to verify the two monotonicity criterion, which may be useful to researchers who wish to utilize both sets of identification results.

The symmetric characterizations illuminate a shared monotonicity property, which we term the minimal monotonicity (MM) condition. We characterize this novel criterion and show it is the minimal requirement needed to identify interpretable causal parameters using 2SLS type estimands. Moreover, minimal monotonicity can be associated with a notion of rationality that enables the investigation of a range of economic choice models that do not comply with ordered or unordered monotonicity.

References

Aliprantis, D. (2012). Compulsory schooling laws and educational attainment. *Journal of Educational and Behavioral Statistics 37*, 316–338.

Angrist, J. D., K. Graddy, and G. W. Imbens (2000, July). The interpretation of instrumental variables estimators in simultaneous equations models with an application to the demand for fish. *Review of Economic Studies 67*(3), 499–527.

Angrist, J. D. and G. W. Imbens (1995, June). Two-stage least squares estimation of average causal effects in models with variable treatment intensity. *Journal of the American Statistical Association 90*(430), 431–442.

Angrist, J. D., G. W. Imbens, and D. Rubin (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association 91*(434), 444–455.

Barua, R. and K. Lang (2016). School entry, educational attainment, and quarter of birth: A cautionary tale of a local average treatment effect. *Journal of Human Capital 10*(3), 347–376.

Brinch, C. N., M. Mogstad, and M. Wiswall (2017). Beyond late with a discrete instrument. *Journal of Political Economy 125*(4), 985–1039.

Buchinsky, M. and R. Pinto (2021). Economics of monotonicity conditions. *NBER Working Paper*.

Cameron, S. V. and J. J. Heckman (1998, April). Life cycle schooling and dynamic selection bias: Models and evidence for five cohorts of American males. *Journal of Political Economy 106*(2), 262–333.

Carneiro, P., K. Hansen, and J. J. Heckman (2003, May). Estimating distributions of treatment effects with an application to the returns to schooling and measurement of the effects of uncertainty on college choice. *International Economic Review 44*(2), 361–422.

Cunha, F., J. J. Heckman, and S. Navarro (2007, November). The identification and economic content of ordered choice models with stochastic cutoffs. *International Economic Review 48*(4), 1273–1309.

Dahl, C. M., M. Huber, and G. Mellace (2017). It's never too late: A new look at local average treatment effects with or without defiers. *Discussion Papers on Business and Economics*.

de Chaisemartin, C. (2017). Tolerating defiance? local average treatment effects without monotonicity. *Quantitative Economics 8*(2), 367–396.

Feller, A., T. Grindal, L. Miratrix, and L. C. Page (2016). Compared to what? variation in the impacts of early childhood education by alternative care type. *The Annals of Applied Statistics 10*(3), 1245–1285.

Frölich, M. (2007). Nonparametric iv estimation of local average treatment effects with covariates. *Journal of Econometrics 139*(1), 35–75. Endogeneity, instruments and identification.

Goldberger, A. S. (1972, November). Structural equation methods in the social sciences. *Econometrica 40*(6), 979–1001.

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

Heckman, J. J. (1976, December). The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. *Annals of Economic and Social Measurement 5*(4), 475–492.

Heckman, J. J. (1979, January). Sample selection bias as a specification error. *Econometrica 47*(1), 153–162.

Heckman, J. J. and S. Urzúa (2010, May). Comparing IV with structural models: What simple IV can and cannot identify. *Journal of Econometrics 156*(1), 27–37.

Heckman, J. J., S. Urzua, and E. Vytlacil (2006). Understanding instrumental variables in models with essential heterogeneity. *The Review of Economics and Statistics 88*(3), 389–432.

Heckman, J. J., S. Urzúa, and E. J. Vytlacil (2008). Instrumental variables in models with multiple outcomes: The general unordered case. *Annales d'Economie et de Statistique 91–92*, 151–174.

Heckman, J. J. and E. J. Vytlacil (1999, April). Local instrumental variables and latent variable models for identifying and bounding treatment effects. *Proceedings of the National Academy of Sciences 96*(8), 4730–4734.

Heckman, J. J. and E. J. Vytlacil (2005, May). Structural equations, treatment effects and econometric policy evaluation. *Econometrica 73*(3), 669–738.

Heckman, J. J. and E. J. Vytlacil (2007a). Econometric evaluation of social programs, part I: Causal models, structural models and econometric policy evaluation. In J. J. Heckman and E. E. Leamer (Eds.), *Handbook of Econometrics*, Volume 6B, Chapter 70, pp. 4779–4874. Amsterdam: Elsevier B. V.

Heckman, J. J. and E. J. Vytlacil (2007b). Econometric evaluation of social programs, part II: Using the marginal treatment effect to organize alternative economic estimators to evaluate social programs, and to forecast their effects in new environments. In J. J. Heckman and E. E. Leamer (Eds.), *Handbook of Econometrics*, Volume 6B, Chapter 71, pp. 4875–5143. Amsterdam: Elsevier B. V.

Huber, M., L. Laffers, and G. Mellace (2017). Sharp iv bounds on average treatment effects on the treated and other populations under endogeneity and noncompliance. *Journal of Applied Econometrics 32*(1), 56–79.

Huber, M. and G. Mellace (2012). Relaxing monotonicity in the identification of local average treatment effects. Unpublished manuscript, Department of Economics, University of St. Gallen.

Huber, M. and G. Mellace (2015). Testing instrument validity for late identification based on inequality moment constraints. *Review of Economics and Statistics 97*, 398–411.

Hull, P. (2018). Isolateing: Identifying counterfactual-specific treatment effects with cross-stratum comparisons. *Umpublished Manuscript*.

Imbens, G. W. and J. D. Angrist (1994, March). Identification and estimation of local average treatment effects. *Econometrica 62*(2), 467–475.

Imbens, G. W. and D. B. Rubin (1997, October). Estimating outcome distributions for compliers in instrumental variables models. *Review of Economic Studies 64*(4), 555–574.

Kamat, V. (2021). Identifying the effects of a program offer with an application to head start. *Umpublished Manuscrip*.

Kirkeboen, L. J., E. Leuven, and M. Mogstad (2016). Field of study, earnings, and self-selection. *The Quarterly Journal of Economics 131*(3), 1057–1111.

Klein, T. J. (2010). Heterogeneous treatment effects: Instrumental variables without monotonicity? *Journal of Econometrics 155*, 99–116.

Kline, P. and C. R. Walters (2016, 07). Evaluating public programs with close substitutes: The case of head start. *The Quarterly Journal of Economics 131*(4), 1795–1848.

Lee, S. and B. Salanié (2018). Identifying effects of multivalued treatments. *Econometrica 86*, 1939–1963.

Mogstad, M., A. Santos, and A. Torgovitsky (2018). Using instrumental variables for inference about policy relevant treatment parameters. *Econometrica 86*(5), 1589–1619.

Mogstad, M. and A. Torgovitsky (2018). Identification and extrapolation of causal effects with instrumental variables. *Annual Review of Economics 10*(1), 577–613.

Mogstad, M., A. Torgovitsky, and C. R. Walters (2021). The causal interpretation of two-stage least squares with multiple instrumental variables. *American Economic Review 111*(11), 3663–3698.

Mountjoy, J. (2021). Community colleges and upward mobility. *Umpublished Manuscript*.

Pinto, R. (2021). Beyond intention to treat: Using the incentives in moving to opportunity to identify neighborhood effects. *NBER Working Paper*.

Rao, C. R. (1946a). Difference sets and combinatorial arrangements derivable from finite geometries. *Proceedings of the Indian National Science Academy 12*(3), 123–135.

Rao, C. R. (1946b). Hypercubes of strength 'd' leading to confounded designs in factorial experiments. *Bulletin of the Calcutta Mathematical Society 38*, 67–78.

Rao, C. R. (1947). Factorial experiments derivable from combinatorial arrangements of arrays. *Supplement to the Journal of the Royal Statistical Society 9*(1), 128–139.

Rao, C. R. (1949). On a class of arrangements. *Proceedings of the Edinburgh Mathematical Society 8*(3), 119–125.

Rose, E. K. and Y. Shem-Tov (2021). On recoding ordered treatments as binary indicators. *Umpublished Manuscript*.

Rubin, D. B. (1974, October). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology 66*(5), 688–701.

Rubin, D. B. (1978). Bayesian inference for causal effects. *The Annals of Statistics 6*(1), 34–58.

Ryser, H. (1957). Combinatorial properties of matrices of zeros and ones. *Canadian Journal of Mathematics 9*, 371–377.

Small, D. S. and Z. Tan (2007). A stochastic monotonicity assumption for the instrumental variables method. *Technical report, Department of Statistics, Wharton School, University of Pennsylvania.*.

Vytlacil, E. J. (2002, January). Independence, monotonicity, and latent index models: An equivalence result. *Econometrica 70*(1), 331–341.

Vytlacil, E. J. (2006, August). Ordered discrete-choice selection models and local average treatment effect assumptions: Equivalence, nonequivalence, and representation results. *Review of Economics and Statistics 88*(3), 578–581.

A Proofs of Main Results

A.1 Lonesum Matrix Characterizations

Using the characterization of unordered monotonicity in UM-Sequence (11), unordered monotonicity is equivalent to there being a permutation of the rows of $\boldsymbol{B}_t$ such that each column of $\boldsymbol{B}_t$ is weakly increasing.[20] Existence of such a reordering characterizes a class of binary matrices known as lonesum matrices, which are a generalization of lower triangular binary matrices. The lonesum property, and various characterizations of it, end up forming the basis of much of our proof strategy. We discuss the property breifly below and note a useful characterization of the property.

Lonesum Matrices

Following Ryser (1957), a binary matrix $\boldsymbol{A}$ is lonesum if each of its entries is uniquely determined by its column and row sums. Matrix $\boldsymbol{A}$ below is an example of such a lonesum matrix:

$$
\begin{array}{c} \\ \boldsymbol{A} \\ \text{column} \\ \text{column-sum} \end{array}
\begin{array}{c} \\ = \\ \\ \\ \end{array}
\underbrace{\begin{array}{cc}
 & \begin{array}{cc} \text{row} & \text{row-sum} \end{array} \\
\left[\begin{matrix} 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 \end{matrix}\right] & \begin{array}{cc} r_1 & 2 \\ r_2 & 4 \\ r_3 & 1 \end{array} \\
\begin{matrix} c_1 & c_2 & c_3 & c_4 & c_5 \\ 0 & 3 & 1 & 2 & 1 \end{matrix} &
\end{array}}_{\text{Original Matrix}}
\quad \Rightarrow \quad
\underbrace{\begin{array}{cc}
 & \begin{array}{cc} \text{row} & \text{row-sum} \end{array} \\
\left[\begin{matrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \end{matrix}\right] & \begin{array}{cc} r_3 & 1 \\ r_1 & 2 \\ r_2 & 4 \end{array} \\
\begin{matrix} c_1 & c_3 & c_5 & c_4 & c_2 \\ 0 & 1 & 1 & 2 & 3 \end{matrix} &
\end{array}}_{\text{Reordered Matrix}}
$$

We can reorder the rows of the matrix $\boldsymbol{A}$ such that the elements of each column are weaklly increasing. We can also reorder the columns so that the matrix $\boldsymbol{A}$ is lower triangular, which is why the lonesum property is considered a generalization of binary lower triangular matrices. The lonesum matrix property can also be productively characterized in the following ways:

Lemma 3 (Lonesum Matrices)**.** *A binary matrix* $\boldsymbol{A} \in \{0,1\}^{m \times n}$ *is lonesum if and only if:*

(i). Matrix $\boldsymbol{A}$ *is lower-triangular under column and row permutations.*

(ii). There are no 2×2 *submatrix in* $\boldsymbol{A}$ *of the form either:*

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \; or \; \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}. \tag{A.1}$$

---

[20] This permutation can differ for each $\boldsymbol{B}_t$, but there must be such a permutation for each $t \in \mathcal{T}$.

(iii). $\boldsymbol{\iota}^\intercal\big((\mathbb{1}-\boldsymbol{A})^\intercal \boldsymbol{A}\big) \odot \big((\mathbb{1}-\boldsymbol{A})^\intercal \boldsymbol{A}\big)^\intercal \boldsymbol{\iota} = 0$, *where* $\boldsymbol{\iota}$ *is a* n*-dimensional vector of elements ones and* $\mathbb{1}$ *is a* $m \times n$ *matrix of element ones.*

(iv). Let $r_i(\boldsymbol{A})$ *and* $c_j(\boldsymbol{A})$ *represent the row sum of row* i *and the column sum of column* j*, respectively. Each entry* $\boldsymbol{A}[i,j]$*, for* $1 \leq i \leq m$ *and* $1 \leq i \leq n$*, can be expressed as:*

$$\boldsymbol{A}[i,j] = \mathbf{1}\left[r_i(\boldsymbol{A}) \;\geq\; \sum_{j'=1}^{n} \mathbf{1}\left[c_{j'}(\boldsymbol{A}) \geq c_j(\boldsymbol{A})\right]\right] \tag{A.2}$$

The sum $\sum_{j'=1}^{n} \mathbf{1}\left[c_{j'}(\boldsymbol{A}) \geq c_j(\boldsymbol{A})\right]$ *represents the number of columns of* $\boldsymbol{A}$ *with a weakly larger column sum than column* j

Proof. The second item comes from Ryser (1957). We next show a series of implications.

(iii) $\iff$ *(ii).* Notice that *(iii)* is equivalent to the matrix

$$(\mathbb{1}-\boldsymbol{A})^\intercal \boldsymbol{A} \odot \boldsymbol{A}^\intercal(\mathbb{1}-\boldsymbol{A})$$

being a matrix of all zeros. By direct calculation, the ij^{th} element of $(\mathbb{1}-\boldsymbol{A})^\intercal \boldsymbol{A}$ is given

$$\sum_{k=1}^{m} \boldsymbol{A}[k,i] - \boldsymbol{A}[k,j]\boldsymbol{A}[k,i] = \sum_{k=1}^{m} \boldsymbol{A}[k,i](1-\boldsymbol{A}[k,j]).$$

This is non-zero if and only if $\boldsymbol{A}[k,i] = 1$ and $\boldsymbol{A}[k,j] = 0$ for some k. Conversely, the ij^{th} element of $\boldsymbol{A}^\intercal(\mathbb{1}-\boldsymbol{A})$ can be expressed

$$\sum_{k=1}^{m} \boldsymbol{A}[k,j] - \boldsymbol{A}[k,i]\boldsymbol{A}[k,j] = \sum_{k=1}^{m} \boldsymbol{A}[k,j](1-\boldsymbol{A}[k,i]).$$

This is non-zero if and only if $\boldsymbol{A}[k,i] = 0$ and $\boldsymbol{A}[k,j] = 1$ for some k. The ij^{th} element of $(\mathbb{1}-\boldsymbol{A})^\intercal \boldsymbol{A} \odot \boldsymbol{A}^\intercal(\mathbb{1}-\boldsymbol{A})$ is non-zero if and only if each of the terms above are nonzero, which in turn is equivalent to their existing two rows k, k' such that

$$\begin{pmatrix} \boldsymbol{A}[k,i] & \boldsymbol{A}[k,j] \\ \boldsymbol{A}[k',i] & \boldsymbol{A}[k',j] \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

which is equivalent to one of the restricted submatrices in item (ii) up to a relabeling of k and k'.

(iv) $\Longleftrightarrow$ *(ii).* For the forward direction, notice that if *(iv)* holds, then the matrix $\boldsymbol{A}$ can clearly be reproduced uniquely from its row and column sums, so it is lonesum and thus *(ii)* holds. We then want to show that *(ii)* $\Longrightarrow$ *(iv).* Consider the contrapositive. First, suppose that there is some element $\boldsymbol{A}[i,j]$ such that $\boldsymbol{A}[i,j] = 1$ but

$$r_i(\boldsymbol{A}) < \sum_{j'=1}^{n} \mathbf{1}[c_{j'}(\boldsymbol{A}) \geq c_j(\boldsymbol{A})]. \tag{A.3}$$

That is, the row sum of row i is less than the number of columns j' with column sum larger than column j. For this to be the case, there must be some column j' such that $\boldsymbol{A}[i,j'] = 0$ but $c_{j'}(\boldsymbol{A}) \geq c_j(\boldsymbol{A})$. In other words, there must be some column j' that contributes to the right hand side of (A.3) but not to the row sum on the left hand side. Because $\boldsymbol{A}[i,j'] = 0$ we know that $j \neq j'$.

Because $\boldsymbol{A}[i,j] = 1$ but $\boldsymbol{A}[i,j'] = 0$, for the column sum of j' to be weakly larger than the column sum of j, there must be some other row $i' \neq i$ such that $\boldsymbol{A}[i',j] = 0$ but $\boldsymbol{A}[i',j'] = 1$. This generates the restricted pattern which violates *(ii)* (up to a relabeling of columns j and j').

$$\begin{pmatrix} \boldsymbol{A}[i,j] & \boldsymbol{A}[i,j'] \\ \boldsymbol{A}[i',j] & \boldsymbol{A}[i',j'] \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Conversely, suppose that $\boldsymbol{A}[i,j] = 0$ but

$$r_i(\boldsymbol{A}) \geq \sum_{j'=1}^{n} \mathbf{1}[c_{j'}(\boldsymbol{A}) \geq c_j(\boldsymbol{A})] \tag{A.4}$$

By similar logic as the previous case, this means that there must be some column j' such that $\boldsymbol{A}[i,j'] = 1$ but $c_{j'}(\boldsymbol{A}) < c_j(\boldsymbol{A})$. Since column j does not contribute to the row sum on the right hand side of (A.4) but does contribute to the column sum on the right hand side, there must be some other column that does the opposite in order for (A.4) to hold.

Because $\boldsymbol{A}[i,j'] = 1$ and $\boldsymbol{A}[i,j] = 0$ but $c_{j'}(\boldsymbol{A}) < c_j(\boldsymbol{A})$, there must be some other row $i' \neq i$ such that $\boldsymbol{A}[i,j'] = 1$ but $\boldsymbol{A}[i',j'] = 0$. As before, we can show that this generates the restricted pattern which violates *(ii)* (up to a relabeling of columns j and j').

(i) $\iff$ *(ii).* That *(i)* $\implies$ *(ii)* is clear, since the existence of the restricted pattern is stable under row and column permutations and the existence of such a restricted pattern prevents a matrix from being lower triangular. To see that *(ii)* $\implies$ *(i)*, suppose that $\boldsymbol{A}$ is lonesum and let $\tilde{\boldsymbol{A}}$ be the matrix generated by ordering the rows of $\boldsymbol{A}$ is increasing row-sum and the colums of $\boldsymbol{A}$ in decreasing column sum. Now, notice that $\tilde{\boldsymbol{A}}$ must also be lonesum, since if the restricted pattern appears in $\tilde{\boldsymbol{A}}$ it must also appear in $\boldsymbol{A}$; we cannot generate the restricted pattern using row and column permutations of a lonesum matrix. If $\tilde{\boldsymbol{A}}$ is lonesum, then *(iv)* must hold, that is

$$\tilde{\boldsymbol{A}}[i,j] = \mathbf{1}\left[r_i(\tilde{\boldsymbol{A}}) \;\geq\; \sum_{j'=1}^{n} \mathbf{1}\left[c_{j'}(\tilde{\boldsymbol{A}}) \geq c_j(\tilde{\boldsymbol{A}})\right]\right]. \tag{A.5}$$

Since the columns of $\tilde{\boldsymbol{A}}$ are ordered in increasing column sum, we must have by (A.5) that $\tilde{\boldsymbol{A}}[i,j] \leq \tilde{\boldsymbol{A}}[i,j']$ for $j' > j$. Similarly, since the rows are ordered in decreasing row sum, (A.5) implies that $\tilde{\boldsymbol{A}}[i,j] \leq \tilde{\boldsymbol{A}}[i',j]$ for $i' < i$. Together, these imply that $\tilde{\boldsymbol{A}}$ is lower triangular.

$\square$

A.2 Proofs of Results in Section 4

A.2.1 Proof of Theorem 1

We prove Theorem 1 via a series of implications:

(i) $\implies$ *(ii).* If there is a response type $\boldsymbol{s}$ and a treatment t for which *(ii)* is not true, then the sequence

$$\left(\mathbf{1}[\boldsymbol{s}[z_1^{(t)}] = t], \ldots, \mathbf{1}[\boldsymbol{s}[z_{N_Z}^{(t)}] = t]\right)$$

is not increasing.

(ii) $\implies$ *(iii)*. If there is a 2×2 matrix of $\boldsymbol{R}$ of the form

$$\begin{array}{ccc} \boldsymbol{s} & \boldsymbol{s}' & \\ \left(\begin{matrix} t & t' \\ t'' & t \end{matrix}\right. & \left.\right) & \begin{matrix} z \\ z' \end{matrix} \end{array}$$

then we have $\mathbf{1}[\boldsymbol{s}[z] = t] > \mathbf{1}[\boldsymbol{s}[z'] = t]$ while $\mathbf{1}[\boldsymbol{s}'[z] = t] < \mathbf{1}[\boldsymbol{s}'[z'] = t]$, a violation of *(ii)*. A symmetric argument holds for the other restricted submatrix of $\boldsymbol{R}$.

(iii) $\implies$ *(iv)*. First notice that *(iv)* is equivalent to the matrix $\boldsymbol{U}$ being lonesum by part three of Lemma 3. Further notice that the matrix $\boldsymbol{U}$ is lonesum if and only if each $\boldsymbol{B}_t$ is lonesum. Because there are no restricted submatrices of $\boldsymbol{R}$ of the form in *(iii)* there are no submatrices of any $\boldsymbol{B}_t$ of the form either

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ or } \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

By part *(ii)* of Lemma 3, this is equivalent to each $\boldsymbol{B}_t$ being lonesum.

(iv) $\implies$ *(v)*. Item *(iv)* is equivalent to each $\boldsymbol{B}_t$ being lonesum. We seek to use the fourth item of Lemma 3. With this in mind, define the functions

$$\zeta(z, t) \equiv \text{row sum of the } z^{\text{th}} \text{ row of the matrix } \boldsymbol{B}_t$$
$$\varphi(\boldsymbol{s}, t) \equiv \# \text{ of columns of } \boldsymbol{B}_t \text{ with a larger column sum than that of its } \boldsymbol{s}^{\text{th}} \text{ column.}$$

Because $\boldsymbol{S}$ is implicitly a function of unobserved confounders $\boldsymbol{V}$, we can also write $\varphi(\cdot, t)$ as a function of $\boldsymbol{V}$. By the definition of $\boldsymbol{B}_t = \mathbf{1}[\boldsymbol{R} = t]$ and the fourth item of Lemma 3, we can then write

$$\mathbf{1}[T = t] = \mathbf{1}[\zeta(Z, t) \geq \varphi(\boldsymbol{V}, t)].$$

By the first item of Lemma 3, there is a reordering of the rows and columns of $\boldsymbol{B}_t$ such that $\boldsymbol{B}_t$ is lower triangular. Let $z_1^{(t)}$ denote the instrument associated with the "top" row of this matrix, $z_2^{(t)}$ denote the second row of this matrix, and so on till $z_{N_Z}^{(t)}$.

Then, the function ζ satisfies $\zeta(z_{k+1}^{(t)}, t) \geq \zeta(z_k^{(t)}, t)$ for $k = 1, ..., N_Z - 1$, by definition of the sequence

$(z_1^{(t)}, \ldots, z_{N_Z}^{(t)})$; weakly more response types must be taking up treatment t for each successive value of this sequence.

(v) $\implies$ *(i)*. Since $\zeta(z_{k+1}^{(t)}, t) > \zeta(z_k^{(t)}, t)$ for $k = 1, ..., N_Z - 1$, if *(v)* holds we must have that

$$(\mathbf{1}[T_i(z_1^{(t)}) = t], \ldots, \mathbf{1}[T_i(z_{N_Z}^{(t)}) = t])$$

is a weakly increasing sequence in $\{0, 1\}$ for all $i \in \mathcal{I}$.

A.2.2 Proof of Theorem 2

We prove Theorem 2 via a series of implications:

(i) $\implies$ *(ii)*. Take any strict ordering on $\mathcal{T}$. Suppose there is a violation of *(ii)* for some $i \in \mathcal{I}$. Then for that particular $i \in \mathcal{I}$ the sequence

$$(T_i(z_1), \ldots, T_i(z_{N_Z}))$$

is not increasing with respect to the ordering on $\mathcal{T}$. If there is no such ordering satisfying *(ii)*, then *(i)* cannot be satisfied.

(ii) $\implies$ *(iii)*. Suppose there is a 2×2 submatrix of $\boldsymbol{R}$ of the form:

$$\begin{array}{cc} \begin{array}{cc} \boldsymbol{s} & \boldsymbol{s}' \end{array} & \\ \begin{pmatrix} t & t' \\ t'' & t''' \end{pmatrix} & \begin{array}{c} z \\ z' \end{array} \end{array}$$

for some $t'' > t$ and $t''' < t'$. This means that $\boldsymbol{s}[z'] > \boldsymbol{s}[z]$ while $\boldsymbol{s}'[z] > \boldsymbol{s}'[z']$. These two statememnts cannot both be true under *(ii)*. A symmetric argument applies for the other 2×2 restricted submatrix.

(iii) $\implies$ *(iv)*. Notice that by part 3 of Lemma 3 that *(iv)* is equivalent to $\boldsymbol{O}$ being lonesum. Suppose that $\boldsymbol{O} = [\boldsymbol{B}_{t_1}^*, \ldots, \boldsymbol{B}_{t_{N_T}}^*]$ is not lonesum. That is, by the lonesum characterization (A.1)

there is a 2×2 submatrix of $\boldsymbol{O}$ of the form

$$\begin{array}{ccc} \boldsymbol{s} & \tilde{\boldsymbol{s}} & \\ \left(\begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix}\right. & \left.\right) & \begin{matrix} z \\ z' \end{matrix} \end{array} \quad \text{or} \quad \begin{array}{ccc} \boldsymbol{s} & \tilde{\boldsymbol{s}} & \\ \left(\begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix}\right. & \left.\right) & \begin{matrix} z \\ z' \end{matrix} \end{array}.$$

WLOG suppose there is a 2×2 submatrix of $\boldsymbol{O}$ of the first form. This indicates that for some $t \in \mathcal{T}$, the instrumental switch from z to z' induces agents of response type $\boldsymbol{s}$ to switch from a treatment weakly below t to a treatment strictly greater than t. That is $\boldsymbol{s}[z] < \boldsymbol{s}[z']$. Conversely, for some treatment $t' \in \mathcal{T}$ the instrumental switch from z to z' induces agents of response $\tilde{\boldsymbol{s}}$ to switch from a treatment strictly greater than t' to a treatment weakly lower than t'. That is $\tilde{\boldsymbol{s}}[z] < \tilde{\boldsymbol{s}}[z']$.

These two statements are incompatible with each other, so we cannot have that $\boldsymbol{s} = \tilde{\boldsymbol{s}}$. Leting $t = \boldsymbol{s}[z], t' = \tilde{\boldsymbol{s}}[z], t'' = \boldsymbol{s}[z']$, and $t''' = \tilde{\boldsymbol{s}}[z']$, this implies a 2×2 submatrix of $\boldsymbol{R}$ of the form

$$\begin{array}{ccc} \boldsymbol{s} & \tilde{\boldsymbol{s}} & \\ t & t' & z \\ t'' & t''' & z' \end{array}$$

with $t < t''$ and $t' > t'''$. This is a violation of *(iii)*. Similarly, considering 2×2 submatrices of $\boldsymbol{O}$ of the second form, we can find a violation of the other pattern restricted by *(iii)*.

(iv) $\implies$ *(v)*. Item *(iv)* is equivalent to $\boldsymbol{O}$ being lonesum. This in turn implies that $\boldsymbol{B}_t^* = \mathbf{1}[\boldsymbol{R} \geq t]$ is lonesum for each $t \in \mathcal{T}$. With this in mind, define the functions

$$\zeta(z,t) \equiv \text{row sum of the } z^{\text{th}} \text{ row of the matrix } \boldsymbol{B}_t^*$$
$$\varphi(\boldsymbol{s},t) \equiv \# \text{ of columns of } \boldsymbol{B}_t^* \text{ with a larger column sum than that of its } \boldsymbol{s}^{\text{th}} \text{ column.}$$

By the first item of Lemma 3, there is a reordering of the rows and columns of $\boldsymbol{O}$ such that $\boldsymbol{O}$ is lower triangular. Let $z_1^{(t)}$ denote the instrument associated with the "top" row of this matrix, $z_2^{(t)}$ denote the second row of this matrix, and so on till $z_{N_Z}^{(t)}$.

Because $\boldsymbol{S}$ is implicity a function of unobserved confounders $\boldsymbol{V}$, we can also write $\varphi(\cdot,t)$ as a function of $\boldsymbol{V}$. By the definition of $\boldsymbol{B}_t^* = \mathbf{1}[\boldsymbol{R} \geq t]$ and the fourth item of Lemma 3, we can then write

$$\mathbf{1}[T \geq t] = \mathbf{1}[\zeta(Z, t) \geq \varphi(\boldsymbol{V}, t)]$$

for each t. Moreover, by definition of the sequence $(z_1, \ldots, z_{N_Z})$ we know that, for each treatment $t \in \mathcal{T}$, weakly more response types take up treatments larger than t as the instrument cycles through the sequence $(z_1, \ldots, z_{N_Z})$. By definition of the $\zeta(\cdot, \cdot)$ function then, $\zeta(z_{k+1}, t) > \zeta(z_k, t)$ for $k = 1, \ldots, N_Z - 1$ and all t.

(v) $\implies$ *(i)*. Since $\zeta(z_{k+1}, t) > \zeta(z_k, t)$ for $k = 1, \ldots, N_Z - 1$ and all t, if holds *(v)* we must have that, for all t

$$(\mathbf{1}[T_i(z_1) \geq t], \ldots, \mathbf{1}[T_i(z_{N_Z}) \geq t])$$

is a weakly increasing sequence in $\{0, 1\}$ for all $i \in \mathcal{I}$. This implies that

$$(T_i(z_1), \ldots, T_i(z_{N_Z}))$$

must be a weakly increasing sequence with respect to the ordering on $\mathcal{T}$ for all $i \in \mathcal{I}$.

A.3 Proofs of Results in Section 5

A.3.1 Proof of Theorem 3

We show a system of implications.

(i) $\iff$ *(ii)*. This is provided by Lemma 1.

(ii) $\iff$ *(iv)*. This follows from the discussion in Appendix B, namely the decomposition of $\beta_{z,z'}$ in equation (B.6), and the definition of an interpretable causal parameter in (20). The decomposition of $\beta_{z,z'}$ gives the forward direction. The definition of an interpretable causal parameter gives the backwards direction: if there is a negative weight there must be a pair of treatments t, t' such that some agents that are switching from t to t' as the instrument ranges from z to z' whereas that same instrumental switch moves other agents from t' to t.

(iii) $\implies$ *(ii)*. We consider the contrapositive. Suppose there is no binary matrix $\boldsymbol{B}$ element-wise

less than or equal to $\sum_{t''\neq t,t'} \boldsymbol{B}_{t''}$ such that $\boldsymbol{B}_t + \boldsymbol{B}$ is lonesum. This means that $\boldsymbol{B}_t$ is lonesum, so that by Theorem 1 there is a 2×2 submatrix $\boldsymbol{R}$ of either the form

$$\begin{pmatrix} t & t''' \\ t'' & t \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} t'' & t \\ t & t''' \end{pmatrix},$$

for some $t'', t''' \neq t$. If either $t'' \neq t'$ or $t''' \neq t'$, then we can find a binary matrix that is element wise less than $\sum_{t''\neq t,t'} \boldsymbol{B}_{t''}$ to "fill in the gap" and get rid of the restricted pattern. In particular we can take $\tilde{\boldsymbol{B}}$ to be the matrix that is equal to one at the position of either t'' or t''' and zero everywhere else. If there is no such matrix that eliminates the restricted pattern then both $t'' = t'$ and $t''' = t'$. So we have the restricted pattern (19) in $\boldsymbol{R}$.

(ii) $\iff$ *(iii)*. The ij^{th} element of $\boldsymbol{B}_t^\intercal \boldsymbol{B}_{t'}$ is given

$$\sum_{z=z_1}^{N_Z} \mathbf{1}[\boldsymbol{s}_i[z] = t]\mathbf{1}[\boldsymbol{s}_j[z] = t'],$$

this is nonzero if and only if we have $\boldsymbol{s}_i[z] = t$ and $\boldsymbol{s}_j[z] = t'$ for some instrument value z. Similarly, the ij^{th} element of $(\boldsymbol{B}_t^\intercal \boldsymbol{B}_{t'})^\intercal = \boldsymbol{B}_{t'}^\intercal \boldsymbol{B}_t$ is given

$$\sum_{z=z_1}^{N_Z} \mathbf{1}[\boldsymbol{s}_i[z] = t']\mathbf{1}[\boldsymbol{s}_j[z] = t].$$

This is non-zero if and only if we have $\boldsymbol{s}_i[z'] = t'$ and $\boldsymbol{s}_j[z'] = t$ for some insrument value z'.

Then, the ij^{th} element of the Hadamard product $(\boldsymbol{B}_t^\intercal \boldsymbol{B}_{t'}) \odot (\boldsymbol{B}_t^\intercal \boldsymbol{B}_{t'})^\intercal$ is non-zero if and only if the ij^{th} elements of both $(\boldsymbol{B}_t^\intercal \boldsymbol{B}_{t'})$ and $(\boldsymbol{B}_t^\intercal \boldsymbol{B}_{t'})^\intercal$ are non-zero. By the characterizations above and because each response type is a well defined function, this is equivalent to $\boldsymbol{s}_i[z] = t, \boldsymbol{s}_i[z'] = t'$ but $\boldsymbol{s}_j[z] = t', \boldsymbol{s}_j[z'] = t$ for some $z' \neq z$. This is equivalent to the restricted pattern (19) existing between $\boldsymbol{s}_i$ and $\boldsymbol{s}_j$ for instrument values z, z' and the specific treatment values t and t'.

All elements of $(\boldsymbol{B}_t^\intercal \boldsymbol{B}_{t'}) \odot (\boldsymbol{B}_t^\intercal \boldsymbol{B}_{t'})^\intercal$ being equal to zero is then equivalent to their being no restricted patterns (19) between the specific treatment values t and t' in the response matrix $\boldsymbol{R}$.

Because each $(\boldsymbol{B}_t^\intercal \boldsymbol{B}_{t'}) \odot (\boldsymbol{B}_t^\intercal \boldsymbol{B}_{t'})^\intercal$ has weakly positive entries, checking whether

$$\iota^\intercal \left(\sum_{(t,t') \in \mathcal{C}_2(\mathcal{T})} (\boldsymbol{B}_t^\intercal \boldsymbol{B}_{t'}) \odot (\boldsymbol{B}_t^\intercal \boldsymbol{B}_{t'})^\intercal \right) \iota = 0$$

is equivalent to checking whether each $(\boldsymbol{B}_t^\intercal \boldsymbol{B}_{t'}) \odot (\boldsymbol{B}_t^\intercal \boldsymbol{B}_{t'})^\intercal$ is equal to the zero matrix. By the discussion above, this is equivalent to checking whether there are no matrices of the form (19) for any $t, t' \in \mathcal{T}$.

A.3.2 Proof of Theorem 4

Let $\Psi_{\boldsymbol{U}}(t) = \big((\tilde{\mathbb{1}} - \boldsymbol{B}_t)^\intercal \boldsymbol{B}_t\big) \odot \big((\mathbb{1} - \boldsymbol{B}_t)^\intercal \boldsymbol{B}_t\big)^\intercal$, where $\tilde{\mathbb{1}}$ is a $N_Z \times N_S$ matrix of element ones. Using this notation, we can rewrite the UM verification in Item (iv) of Theorem 1 as the following sum:

$$\begin{aligned}
\|\Psi_{\boldsymbol{U}}\| &= \|\big((\mathbb{1} - \boldsymbol{U})^\intercal \boldsymbol{U}\big) \;\odot\; \big((\mathbb{1} - \boldsymbol{U})^\intercal \boldsymbol{U}\big)^\intercal\| \\
&= \|\sum_{t \in \mathcal{T}} \big((\tilde{\mathbb{1}} - \boldsymbol{B}_t)^\intercal \boldsymbol{B}_t\big) \;\odot\; \big((\tilde{\mathbb{1}} - \boldsymbol{B}_t)^\intercal \boldsymbol{B}_t\big)^\intercal\| \\
&= \sum_{t \in \mathcal{T}} \|\big((\tilde{\mathbb{1}} - \boldsymbol{B}_t)^\intercal \boldsymbol{B}_t\big) \;\odot\; \big((\tilde{\mathbb{1}} - \boldsymbol{B}_t)^\intercal \boldsymbol{B}_t\big)^\intercal\| \\
&= \sum_{t \in \mathcal{T}} \|\Psi_{\boldsymbol{U}}(t)\|
\end{aligned}$$

The first equality is from the definition of the verification in Theorem 1. The second equality arises from the construction of matrix $\boldsymbol{U}$. The third equality is due to the fact that all elements of $\Psi_{\boldsymbol{U}}(t)$ are either zero or a natural number.

We can use the fact that $\sum_{t \in \mathcal{T}} \boldsymbol{B}_t = \tilde{\mathbb{1}}$ to express matrix $\Psi_{\boldsymbol{U}}(t)$ as:

$$\begin{aligned}
\Psi_{\boldsymbol{U}}(t) &= \Big((\tilde{\mathbb{1}} - \boldsymbol{B}_t)^\intercal \boldsymbol{B}_t\Big) \;\odot\; \Big((\tilde{\mathbb{1}} - \boldsymbol{B}_t)^\intercal \boldsymbol{B}_t\Big)^\intercal \\
&= \left(\left(\left(\sum_{t' \in \mathcal{T}} \boldsymbol{B}_{t'} \right) - \boldsymbol{B}_t \right)^\intercal \boldsymbol{B}_t \right) \;\odot\; \left(\left(\left(\sum_{t' \in \mathcal{T}} \boldsymbol{B}_{t'} \right) - \boldsymbol{B}_t \right)^\intercal \boldsymbol{B}_t \right)^\intercal \\
&= \left(\left(\sum_{t' \in \mathcal{T} \setminus \{t\}} \boldsymbol{B}_{t'} \right)^\intercal \boldsymbol{B}_t \right) \;\odot\; \left(\left(\sum_{t' \in \mathcal{T} \setminus \{t\}} \boldsymbol{B}_{t'} \right)^\intercal \boldsymbol{B}_t \right)^\intercal
\end{aligned}$$

$$
\begin{aligned}
&= \left(\sum_{t' \in \mathcal{T}\setminus\{t\}} \boldsymbol{B}_{t'}^{\intercal} \boldsymbol{B}_t\right) \odot \left(\sum_{t' \in \mathcal{T}\setminus\{t\}} \boldsymbol{B}_{t'}^{\intercal} \boldsymbol{B}_t\right)^{\intercal} \\
&= \left(\sum_{t' \in \mathcal{T}\setminus\{t\}} \boldsymbol{B}_{t'}^{\intercal} \boldsymbol{B}_t\right) \odot \left(\sum_{t' \in \mathcal{T}\setminus\{t\}} \boldsymbol{B}_{t}^{\intercal} \boldsymbol{B}_{t'}\right) \\
&= \sum_{t' \in \mathcal{T}\setminus\{t\}} \left(\boldsymbol{B}_{t'}^{\intercal} \boldsymbol{B}_t\right) \odot \left(\boldsymbol{B}_{t'}^{\intercal} \boldsymbol{B}_t\right)^{\intercal} + 2 \sum_{t',t'' \in \mathcal{T}\setminus\{t\}} \left(\boldsymbol{B}_{t'}^{\intercal} \boldsymbol{B}_t\right) \odot \left(\boldsymbol{B}_{t}^{\intercal} \boldsymbol{B}_{t''}\right)^{\intercal} \\
&= \sum_{t' \in \mathcal{T}\setminus\{t\}} \Psi(t', t, t, t') + 2 \sum_{t',t'' \in \mathcal{T}\setminus\{t\}} \Psi(t', t, t, t'')
\end{aligned}
$$

The derivation above use simple rules of matrix algebra and the formula for the product of sums. We use the fact that $\Psi(t', t, t, t'')^{\intercal} = \Psi(t, t', t'', t)$ and express the transpose of $\Psi_{\boldsymbol{U}}(t)$ as:

$$
\Psi_{\boldsymbol{U}}(t)^{\intercal} = \sum_{t' \in \mathcal{T}\setminus\{t\}} \Psi(t, t', t', t) + 2 \sum_{(t',t'') \in \mathcal{C}_2(\mathcal{T}\setminus\{t\})} \Psi(t, t', t'', t) \tag{A.6}
$$

Recall that the elements of matrix $\Psi(t, t', t'', t''')$ are either zero or natural numbers. Thus, equation (A.6) implies that $\|\Psi_{\boldsymbol{U}}(t)\| = 0$ (or equivalently $\|\Psi_{\boldsymbol{U}}(t)^{\intercal}\| = 0$) if and only if:

$$
\|\Psi(t, t', t', t)\| = 0 \text{ for all } t' \in \mathcal{T} \setminus \{t\} \tag{A.7}
$$

$$
\text{and } \|\Psi(t, t', t'', t)\| = 0 \text{ for all combinations of } t', t'' \in \mathcal{T} \setminus \{t\}. \tag{A.8}
$$

Now $\|\Psi_{\boldsymbol{U}}\| = 0$ only and only if $\|\Psi_{\boldsymbol{U}}(t)\| = 0$ for all $t \in \mathcal{T}$, which completes the proof.

A.4 Proof of Theorem 5

A.4.1 Lemmas

Proof of Lemma 1. Suppose there is a violation of MM (18). This is equivalent to there being pair of response types $\boldsymbol{s}, \boldsymbol{s}'$, a pair of treatments z, z', and a pair of treatments t, t' such that

$$
\begin{aligned}
&\mathbf{1}[\boldsymbol{s}[z] = t]\mathbf{1}[\boldsymbol{s}'[z'] = t'] > \mathbf{1}[\boldsymbol{s}[z] = t']\mathbf{1}[\boldsymbol{s}'[z'] = t] \\
\text{and } \ &\mathbf{1}[\boldsymbol{s}'[z] = t]\mathbf{1}[\boldsymbol{s}[z'] = t'] < \mathbf{1}[\boldsymbol{s}'[z] = t']\mathbf{1}[\boldsymbol{s}[z'] = t].
\end{aligned}
$$

This is in turn equivalent to $\boldsymbol{s}[z] = t$, $\boldsymbol{s}[z'] = t'$ and $\boldsymbol{s}'[z] = t'$, $\boldsymbol{s}'[z'] = t$, which is equivalent (up to a relabeling of $\boldsymbol{s}$ and $\boldsymbol{s}'$) to the restricted pattern (19) appearing in the response matrix $\boldsymbol{R}$.

Proof of Lemma 2. Follows from Theorems 4 and 5. That MM holds when OM and UM fail can be seen via examples in Section 7.

A.5 Proof of Results in Section 6

A.5.1 Proof of Theorem 6

Consider any pair of instrument values $z, z' \in \mathcal{Z}$ and any pair of treatments $t, t' \in \mathcal{T}$. Without loss of generality, it is enough show there are no 2×2 submatrices of the form

$$\begin{array}{cc} \boldsymbol{s} & \boldsymbol{s}' \\ \end{array}$$
$$\begin{pmatrix} t & t' \\ t' & t \end{pmatrix} \begin{array}{c} z \\ z' \end{array} .$$

Define $\Delta_t \equiv \boldsymbol{L}[z', t] - \boldsymbol{L}[z, t]$ and $\Delta_{t'} \equiv \boldsymbol{L}[z', t'] - \boldsymbol{L}[z, t']$. There are two scenarios. Either $\Delta_t \leq \Delta_{t'}$ or $\Delta_t \geq \Delta_{t'}$. In each case, we have the following behavioral restrictions from the Choice Rule (29).

$$\begin{aligned} &\text{If} \quad \Delta_t \leq \Delta_{t'} \quad \text{then} \quad T_i(z) = t' \implies T_i(z') \neq t \\ &\text{If} \quad \Delta_{t'} \leq \Delta_t \quad \text{then} \quad T_i(z) = t \implies T_i(z') \neq t' \end{aligned}$$

The first restriction would eliminate the response type $\boldsymbol{s}'$ from the matrix $\boldsymbol{R}$ while the second restriction would eliminate the response type $\boldsymbol{s}$ from the response matrix $\boldsymbol{R}$. In either case, we cannot have the restricted 2×2 submatrix displayed at the top of the proof.

B 2SLS Analysis

B.1 Interpretation of 2SLS under Ordered and Unordered Monotonicity

Under ordered monotonicity and a binary instruments, Angrist and Imbens (1995) show that the 2SLS estimand identifies the following:

$$\beta_{\text{2SLS}} = \frac{E[Y \mid Z = z_1] - E[Y \mid Z = z_0]}{E[T \mid Z = z_1] - E[T \mid Z = z_0]} \tag{B.1}$$

$$= \sum_{j=1}^{N_T} \omega_{t_j,t_{j-1}} E[Y(t_j) - Y(t_{j-1}) \mid \boldsymbol{S} \in \mathcal{S}_{t_j,t_{j-1}}]$$

where $\mathcal{S}_{t_j,t_{j-1}} \equiv \{\boldsymbol{s} \in \mathcal{S}; \boldsymbol{s}[z_1] \geq t_j > \boldsymbol{s}[z_0]\}$, that is the sets of response types for whom a change in instrument receipt from z_0 to z_1 induces a change in treatment from strictly "below" t_j to weakly "above" t_j. The weights $\omega_{t_j,t_{j-1}}$ are positive and given:

$$\omega_{t_j,t_{j-1}} = \frac{\Pr(\boldsymbol{S} \in \mathcal{S}_{t_j,t_{j-1}})}{\sum_{j=1}^{N_T} \Pr(\boldsymbol{S} \in \mathcal{S}_{t_j,t_{j-1}})}. \tag{B.2}$$

Unordered monotonicity also allows 2SLS type estimands to be expressed in terms of a weighted average of LATE parameters with positive weights.[21] In this setting the 2SLS numerator can be decomposed

$$\begin{aligned} \beta^u_{2\text{SLS}}(z,z') &= E[Y \mid Z = z] - E[Y \mid Z = z'] \\ &= \sum_{\{t,t'\},t\neq t'} \omega^u_{t,t'} E[Y(t) - Y(t') \mid \boldsymbol{S} \in \mathcal{S}_{t,t'}(z,z')] \end{aligned} \tag{B.3}$$

where $\mathcal{S}_{t,t'}(z,z') \equiv \{\boldsymbol{s} : \boldsymbol{s}[z] = t, \boldsymbol{s}[z'] = t'\}$ is the set of response types that switch treatments from t to t' as the instrument varies from z to z'. Under unordered monotonicity, the same instrument switch cannot induce some agents to switch towards choice t while inducing others to switch away from choice t. So, we must have either $\mathcal{S}_{t,t'}(z,z') = \emptyset$ or $\mathcal{S}_{t',t}(z,z') = \emptyset$ (or both). The weights $\omega^u_{t,t'}$ are weakly positive and given $\omega^u_{t,t'} = \Pr(\boldsymbol{S} \in \mathcal{S}_{t,t'})$.

B.2 General Unique Decomposition

Using the identification equality in (3) we can rewrite

$$E[Y \mid Z = z] = \sum_{\boldsymbol{s}\in \text{supp}(\boldsymbol{S})} \sum_{t\in\mathcal{T}} \mathbf{1}[\boldsymbol{s}[z] = t] E[Y(t) \mid \boldsymbol{S} = \boldsymbol{s}] \Pr(\boldsymbol{S} = \boldsymbol{s})$$

---

[21] In addition, Buchinsky and Pinto (2021) show that any variation in the instrumental variable can be used to identify a meaningful counterfactual outcome mean. For instance, the 2SLS estimate that uses a choice indicator for t and any IV-values $z, z' \in \mathcal{Z}$, such that $P(T = t|Z = z) > (T = t'|Z = z)$, identifies the following parameter:

$$\beta^t_{2\text{SLS}}(z,z') = \frac{E[Y\mathbf{1}[T = t] \mid Z = z] - E[Y\mathbf{1}[T = t] \mid Z = z']}{\Pr(T = t \mid Z = z) - \Pr(T = t \mid Z = z')} = E(Y(t)|\boldsymbol{S} \in \mathcal{S}^t_{z,z'}),$$

where $\mathcal{S}^t_{z,z'} = \{\boldsymbol{s} : \boldsymbol{s}[z] = t, \boldsymbol{s}[z'] \neq t\}$ is the set of response types that switch from treatment choice t to any other treatment choice as the instrument varies from z to z'.

$$E[Y \mid Z = z'] = \sum_{\boldsymbol{s} \in \operatorname{supp}(\boldsymbol{S})} \sum_{t \in \mathcal{T}} \mathbf{1}[\boldsymbol{s}[z'] = t] E[Y(t) \mid \boldsymbol{S} = \boldsymbol{s}] \Pr(\boldsymbol{S} = \boldsymbol{s})$$

Using these, we can express the quasi-2SLS estimand as the following

$$\beta_{z,z'} \equiv E[Y \mid Z = z] - E[Y \mid Z = z'] \tag{B.4}$$

$$= \sum_{\boldsymbol{s} \in \operatorname{supp}(\boldsymbol{S})} \sum_{t \in \mathcal{T}} \big(\mathbf{1}[\boldsymbol{s}[z] = t] - \mathbf{1}[\boldsymbol{s}[z'] = t]\big) E[Y(t) \mid \boldsymbol{S} = \boldsymbol{s}] \Pr(\boldsymbol{S} = \boldsymbol{s}) \tag{B.5}$$

$$= \sum_{\{t,t'\}, t \neq t'} E[Y(t) - Y(t') \mid \boldsymbol{S} \in \mathcal{S}_{t,t'}(z,z')] \Pr(\boldsymbol{S} \in \mathcal{S}_{t,t'}(z,z')), \tag{B.6}$$

where the last equality is due to the fact that sets $\mathcal{S}_{t,t'}(z,z')$, defined below (B.3), form a partition of $\operatorname{supp}(\boldsymbol{S})$ as t, t' ranges in $\mathcal{T}$. Equation (B.6) holds regardless of any monotonicity assumption. That is, no matter what the restriction is on the support of $\boldsymbol{S}$, we will always be able to rewrite the 2SLS numerator as in (B.6).

In view of Lemma 1, a violation of MM is equivalent to there being a pair of treatments t, t' such that the sets $\mathcal{S}_{t,t'}(z,z')$ and $\mathcal{S}_{t',t}(z,z')$ are *both* nonempty. This induces negative weights in the 2SLS estimand; both $E[Y(t) - Y(t') \mid \mathcal{S}_{t,t'}(z,z')]$ and $E[Y(t') - Y(t) \mid \mathcal{S}_{t',t}(z,z')]$ are represented in the decomposition (B.6). This in turn limits our ability to use the 2SLS estimand to gain useful insight into the direction of causal effects. The partial minimal monotonicity criterion is then crucial for interpreting $\beta_{z,z'}$ as a type of interpretable causal parameter defined in (20).

C Ordered vs. Unordered Example

We consider a setup with three treatments, $\mathcal{T} = \{t_1, t_2, t_3\}$, and three instruments, $\mathcal{Z} = \{z_1, z_2, z_3\}$. Response matrices (C.1)–(C.2) below are useful to understand the difference between ordered and unordered monotonicity conditions:

$$\boldsymbol{R}_1 = \overbrace{\left[\underbrace{\begin{array}{cccccc} \boldsymbol{s}_1 & \boldsymbol{s}_2 & \boldsymbol{s}_3 & \boldsymbol{s}_4 & \boldsymbol{s}_5 & \boldsymbol{s}_6 \\ t_1 & t_2 & t_3 & t_1 & t_1 & t_2 \\ t_1 & t_2 & t_3 & t_2 & t_3 & \colorbox{red}{t_2} \\ t_1 & t_2 & t_3 & t_3 & t_3 & \colorbox{yellow}{t_3} \end{array}}_{\text{Ordered AND Unordered}} \middle| \begin{array}{c} \boldsymbol{s}_7 \\ t_1 \\ \colorbox{yellow}{t_1} \\ \colorbox{red}{t_2} \end{array}\right]}^{\text{Ordered but NOT Unordered}} \begin{array}{c} \\ z_1 \\ z_2 \\ z_3 \end{array} \tag{C.1}$$

$$\boldsymbol{R}_2 = \underbrace{\begin{array}{c} \overbrace{\begin{array}{cccccc} \boldsymbol{s}_1 & \boldsymbol{s}_2 & \boldsymbol{s}_3 & \boldsymbol{s}_4 & \boldsymbol{s}_5 & \boldsymbol{s}_6 \end{array}} \quad \boldsymbol{s}_7^* \\ \left[\begin{array}{cccccc|c} t_1 & t_2 & t_3 & t_1 & t_1 & t_2 & t_1 \\ t_1 & t_2 & t_3 & t_2 & t_3 & t_2 & t_2 \\ t_1 & t_2 & t_3 & t_3 & t_3 & t_3 & t_1 \end{array}\right] \end{array}}_{\text{Unordered but NOT Ordered}} \begin{array}{c} \\ z_1 \\ z_2 \\ z_3 \end{array} \tag{C.2}$$

Columns $\boldsymbol{s}_1, ..., \boldsymbol{s}_7$ of response matrix $\boldsymbol{R}_1$ (C.1) denote response-types. Each column describes the sequence of counterfactual choices, $(T_i(z_1), T_i(z_2), T_i(z_3))$, for an agent in that column's response type. The counterfactual treatment in each of these sequences is weakly increasing with respect to the ordering $t_1 \le t_2 \le t_3$; for any agent $i \in \mathcal{I}$, $T_i(z_1) \le T_i(z_2) \le T_i(z_3)$. Thus OM-Sequence (10) holds.

However, response types $\boldsymbol{s}_6, \boldsymbol{s}_7$ in $\boldsymbol{R}_1$ violate the sequential representation of unordered monotonicity in (11) for choice t_2. Consider two agents $i, i' \in \mathcal{I}$ such that $\boldsymbol{S}_i = \boldsymbol{s}_6$ and $\boldsymbol{S}_{i'} = \boldsymbol{s}_7$. The sequence of t_2-indicators for agent i is weakly decreasing while the sequence for agent i' is weakly increasing

$$\big(\mathbf{1}[T_i(z_1) = t_2], \mathbf{1}[T_i(z_2) = t_2], \mathbf{1}[T_i(z_3) = t_2]\big) = (1, 1, 0)$$
$$\big(\mathbf{1}[T_{i'}(z_1) = t_2], \mathbf{1}[T_{i'}(z_2) = t_2], \mathbf{1}[T_{i'}(z_3) = t]\big) = (0, 0, 1).$$

This represents a violation of UM-Sequence (11) for the sequencing of $\mathcal{Z}$, (z_1, z_2, z_3). Moreover, because the switch from z_2 to z_3 induces agent i to move strictly away from treatment choice t_2 while moving agent i' strictly towards treatment choice t_2, there is no other sequencing of $\mathcal{Z}$ that would satisfy the requirement of UM Sequence (11). We can conclude that the response matrix $\boldsymbol{R}_1$ does not satisfy unordered monotonicity.

Response matrix $\boldsymbol{R}_2$ (C.2) replaces $\boldsymbol{s}_7$ in $\boldsymbol{R}_1$ with $\boldsymbol{s}_7^*$. The treatment indexes in $\boldsymbol{s}_7^*$ are not weakly increasing with respect to the ordering $t_1 \le t_2 \le t_3$. Thus the ordered monotonicity that held for $\boldsymbol{R}_1$ does not hold for $\boldsymbol{R}_2$. Indeed, the reader can confirm that there is no ordering $\mathcal{T}$ that satisfies OM-Sequence (10). In this case, though, the response matrix $\boldsymbol{R}_2$ satisfies unordered monotonicity. Equations (C.3)–(C.5) are instructive in establishing this fact.

$$\text{Reordered rows of } \boldsymbol{R}_2 \text{ for } t_1 = \begin{array}{cc} \begin{array}{ccccccc} \boldsymbol{s}_1 & \boldsymbol{s}_2 & \boldsymbol{s}_3 & \boldsymbol{s}_4 & \boldsymbol{s}_5 & \boldsymbol{s}_6 & \boldsymbol{s}_7^* \end{array} & \\ \left[\begin{array}{ccccccc} \colorbox{green}{t_1} & t_2 & t_3 & t_2 & t_3 & t_2 & t_2 \\ \colorbox{green}{t_1} & t_2 & t_3 & t_3 & t_3 & t_3 & \colorbox{green}{t_1} \\ \colorbox{green}{t_1} & t_2 & t_3 & \colorbox{green}{t_1} & \colorbox{green}{t_1} & t_2 & \colorbox{green}{t_1} \end{array}\right] & \begin{array}{c} z_2 \\ z_3 \\ z_1 \end{array} \end{array}, \tag{C.3}$$

$$\text{Reordered rows of } \boldsymbol{R}_2 \text{ for } t_2 = \begin{array}{cc} \begin{array}{ccccccc} \boldsymbol{s}_1 & \boldsymbol{s}_2 & \boldsymbol{s}_3 & \boldsymbol{s}_4 & \boldsymbol{s}_5 & \boldsymbol{s}_6 & \boldsymbol{s}_7^* \end{array} & \\ \left[\begin{array}{ccccccc} t_1 & \colorbox{green}{t_2} & t_3 & t_3 & t_3 & t_3 & t_1 \\ t_1 & \colorbox{green}{t_2} & t_3 & t_1 & t_1 & \colorbox{green}{t_2} & t_1 \\ t_1 & \colorbox{green}{t_2} & t_3 & \colorbox{green}{t_2} & t_3 & \colorbox{green}{t_2} & \colorbox{green}{t_2} \end{array}\right] & \begin{array}{c} z_3 \\ z_1 \\ z_2 \end{array} \end{array}, \tag{C.4}$$

$$\text{Reordered rows of } \boldsymbol{R}_2 \text{ for } t_3 = \begin{array}{cc} \begin{array}{ccccccc} \boldsymbol{s}_1 & \boldsymbol{s}_2 & \boldsymbol{s}_3 & \boldsymbol{s}_4 & \boldsymbol{s}_5 & \boldsymbol{s}_6 & \boldsymbol{s}_7^* \end{array} & \\ \left[\begin{array}{ccccccc} t_1 & t_2 & \colorbox{green}{t_3} & t_1 & t_1 & t_2 & t_1 \\ t_1 & t_2 & \colorbox{green}{t_3} & t_2 & \colorbox{green}{t_3} & t_2 & t_2 \\ t_1 & t_2 & \colorbox{green}{t_3} & \colorbox{green}{t_3} & \colorbox{green}{t_3} & \colorbox{green}{t_3} & t_1 \end{array}\right] & \begin{array}{c} z_1 \\ z_2 \\ z_3 \end{array} \end{array}. \tag{C.5}$$

Equation (C.3) reorders the rows of $\boldsymbol{R}_2$ in (C.2) from (z_1, z_2, z_3) to (z_2, z_3, z_1). At each move along the sequence (z_2, z_3, z_1), additional response types switch to treatment t_1 and no response types switch away from t_1. This means that

$$\mathbf{1}[T_i(z_2) = t_1] \leq \mathbf{1}[T_i(z_3) = t_1] \leq \mathbf{1}[T_i(z_1) = t_1]$$

holds for all agents $i \in \mathcal{I}$, regardless of response type. Thus UM Sequence (11) holds for t_1. By symmetric logic, equation (C.4) demonstrates that UM Sequence (11) holds for t_2 using the IV sequence (z_3, z_1, z_2) and equation (C.5) shows that UM Sequence holds for t_3 using sequence (z_1, z_2, z_3). We conclude that unordered monotonicity holds.

Response matrices $\boldsymbol{R}_1$ and $\boldsymbol{R}_2$ in (C.1)–(C.2) show that ordered monotonicity does not imply unordered monotonicity nor vice-versa. Ordered monotoncity holds for $\boldsymbol{R}_1$ but not for $\boldsymbol{R}_2$ while unordered monotonicity holds for $\boldsymbol{R}_2$ but not for $\boldsymbol{R}_1$. The two monotonicity conditions can intersect, both ordered and unordered monotonicity hold for the submatrix generated by response-types $\boldsymbol{s}_1$ to $\boldsymbol{s}_6$.

D Additional Information Regarding the Examples of Section 7

D.1 Verifying Unordered Monotonicity

We seek to show that the response matrix 32 is a case of UM (11) using the verification matrix of item (iv) of Theorem 1. The matrix is presented below for convenience.

$$\boldsymbol{R} = \begin{array}{cccccccc} \boldsymbol{s}_1 & \boldsymbol{s}_2 & \boldsymbol{s}_3 & \boldsymbol{s}_4 & \boldsymbol{s}_5 & \boldsymbol{s}_6 & \boldsymbol{s}_7 & \\ t_1 & t_1 & t_1 & t_1 & t_2 & t_2 & t_3 & z_1 \\ t_1 & t_1 & t_3 & t_3 & t_2 & t_3 & t_3 & z_2 \\ t_1 & t_2 & t_2 & t_3 & t_2 & t_2 & t_3 & z_3 \\ t_1 & t_1 & t_1 & t_1 & t_2 & t_2 & t_3 & z_4 \end{array}$$

Let $\boldsymbol{B}_t = \mathbf{1}[\boldsymbol{R} = t]; t \in \{t_1, t_2, t_3\}$ denote the binary matrices corresponding to response matrix (32). Those are displayed below:

$$\boldsymbol{B}_{t_1} = \begin{array}{cccccccc} \boldsymbol{s}_1 & \boldsymbol{s}_2 & \boldsymbol{s}_3 & \boldsymbol{s}_4 & \boldsymbol{s}_5 & \boldsymbol{s}_6 & \boldsymbol{s}_7 & \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & z_1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & z_2 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & z_3 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & z_4 \end{array}$$

$$\boldsymbol{B}_{t_2} = \begin{array}{cccccccc} \boldsymbol{s}_1 & \boldsymbol{s}_2 & \boldsymbol{s}_3 & \boldsymbol{s}_4 & \boldsymbol{s}_5 & \boldsymbol{s}_6 & \boldsymbol{s}_7 & \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & z_1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & z_2 \\ 0 & 1 & 1 & 0 & 1 & 1 & 0 & z_3 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & z_4 \end{array}$$

$$\boldsymbol{B}_{t_3} = \begin{array}{cccccccc} \boldsymbol{s}_1 & \boldsymbol{s}_2 & \boldsymbol{s}_3 & \boldsymbol{s}_4 & \boldsymbol{s}_5 & \boldsymbol{s}_6 & \boldsymbol{s}_7 & \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & z_1 \\ 0 & 0 & 1 & 1 & 0 & 1 & 1 & z_2 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & z_3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & z_4 \end{array}$$

Unordered monotonicity holds if and only if the binary matrices $\boldsymbol{B}_{t_1}, \boldsymbol{B}_{t_2}, \boldsymbol{B}_{t_3}$ are lonesum. For item (iv) of Theorem 1 to hold, it suffices to show that $\|\Psi_{\boldsymbol{U}}(t)\| = 0$ for all $t \in \{t_1, t_2, t_3\}$ where $\Psi_{\boldsymbol{U}}(t)$ is given by $\Psi_{\boldsymbol{U}}(t) \equiv \big((\mathbb{1} - \boldsymbol{B}_t)^\intercal \boldsymbol{B}_t\big) \odot \big((\mathbb{1} - \boldsymbol{B}_t)^\intercal \boldsymbol{B}_t\big)^\intercal$. It is useful to express $\Psi_{\boldsymbol{U}}(t_1)$ as $\Psi_{\boldsymbol{U}}(t) = \tilde{\Psi}_{\boldsymbol{U}}(t) \odot \tilde{\Psi}_{\boldsymbol{U}}(t)^\intercal$ where $\tilde{\Psi}_{\boldsymbol{U}}(t) = \big((\mathbb{1} - \boldsymbol{B}_t)^\intercal \boldsymbol{B}_t\big)$.

The matrices $\tilde{\Psi}_{\boldsymbol{U}}(t_1), \tilde{\Psi}_{\boldsymbol{U}}(t_2), \tilde{\Psi}_{\boldsymbol{U}}(t_3)$ are computed below:

Note that $\|\Psi_{\boldsymbol{U}}(t)\| = 0$ if $\tilde{\Psi}_{\boldsymbol{U}}(t)$ is a triangular matrix with a zero diagonal. Thus it suffices to evaluate matrix $\tilde{\Psi}_{\boldsymbol{U}}(t)$ for $t \in \{t_1, t_2, t_3\}$.

$$\tilde{\Psi}_{\boldsymbol{U}}(t_1) = \underbrace{\begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}^{\intercal}}_{(\mathbb{1}-\boldsymbol{B}_{t_1})^{\intercal}} \underbrace{\begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 \end{bmatrix}}_{\boldsymbol{B}_{t_1}} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 4 & 3 & 2 & 2 & 0 & 0 & 0 \\ 4 & 3 & 2 & 2 & 0 & 0 & 0 \\ 4 & 3 & 2 & 2 & 0 & 0 & 0 \end{bmatrix}$$

$$\tilde{\Psi}_{\boldsymbol{U}}(t_2) = \underbrace{\begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 1 \end{bmatrix}^{\intercal}}_{(\mathbb{1}-\boldsymbol{B}_{t_2})^{\intercal}} \underbrace{\begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 \end{bmatrix}}_{\boldsymbol{B}_{t_2}} = \begin{bmatrix} 0 & 1 & 1 & 0 & 4 & 3 & 0 \\ 0 & 0 & 0 & 0 & 3 & 2 & 0 \\ 0 & 0 & 0 & 0 & 3 & 2 & 0 \\ 0 & 1 & 1 & 0 & 4 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 4 & 3 & 0 \end{bmatrix}$$

$$\tilde{\Psi}_{\boldsymbol{U}}(t_3) = \underbrace{\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 \end{bmatrix}^{\intercal}}_{(\mathbb{1}-\boldsymbol{B}_{t_3})^{\intercal}} \underbrace{\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}}_{\boldsymbol{B}_{t_3}} = \begin{bmatrix} 0 & 0 & 1 & 2 & 0 & 1 & 4 \\ 0 & 0 & 1 & 2 & 0 & 1 & 4 \\ 0 & 0 & 0 & 1 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 1 & 2 & 0 & 1 & 4 \\ 0 & 0 & 0 & 1 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

It is easy to see that in $\tilde{\Psi}_{\boldsymbol{U}}(t_1) \odot \tilde{\Psi}_{\boldsymbol{U}}(t_1)^{\intercal}$ is equal to a matrix of zeros. Indeed, the matrix $\tilde{\Psi}_{\boldsymbol{U}}(t_1)$ is triangular with a zero diagonal. Thus, when we perform the element wise multiplication of $\tilde{\Psi}_{\boldsymbol{U}}(t_1)$ and its transpose, at least one of the elements of the multiplication will be zero. The same occurs for matrices $\tilde{\Psi}_{\boldsymbol{U}}(t_2)$ and $\tilde{\Psi}_{\boldsymbol{U}}(t_3)$.

D.2 A Case of Choice Incentives for Ordered Monotonicity

The can summarize the above incentive structure the binary incentive matrix given below:

$$\boldsymbol{L} = \begin{array}{c c} \begin{array}{ccc} t_1 & t_2 & t_3 \end{array} & \\ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} & \begin{array}{c} z_1 \\ z_2 \\ z_3 \\ z_4 \end{array} \end{array} \tag{D.1}$$

We use this first example to describe the machinery that translates choice incentives into monotonicity conditions and identification results. We adopt a more parsimonious approach in the subsequent examples.

Table 3: Choice Restrictions generated by Incentive Matrix (D.1)

1	$T_i(z_1) = t_1$	$\Rightarrow$	$\varnothing$
2	$T_i(z_2) = t_1$	$\Rightarrow$	$T_i(z_1) \notin \{t_2, t_3\}$ and $T_i(z_3) \notin \{t_2, t_3\}$ and $T_i(z_4) \neq t_2$
3	$T_i(z_3) = t_1$	$\Rightarrow$	$T_i(z_1) \notin \{t_2, t_3\}$ and $T_i(z_2) \notin \{t_2, t_3\}$ and $T_i(z_4) \neq t_2$
4	$T_i(z_4) = t_1$	$\Rightarrow$	$T_i(z_1) \notin \{t_2, t_3\}$ and $T_i(z_2) \notin \{t_2, t_3\}$ and $T_i(z_3) \notin \{t_2, t_3\}$
5	$T_i(z_1) = t_2$	$\Rightarrow$	$T_i(z_2) \notin \{t_1, t_3\}$ and $T_i(z_3) \notin \{t_1, t_3\}$ and $T_i(z_4) \neq t_1$
6	$T_i(z_2) = t_2$	$\Rightarrow$	$T_i(z_1) \neq t_3$ and $T_i(z_3) \notin \{t_1, t_3\}$ and $T_i(z_4) \neq t_1$
7	$T_i(z_3) = t_2$	$\Rightarrow$	$T_i(z_1) \neq t_3$ and $T_i(z_2) \notin \{t_1, t_3\}$ and $T_i(z_4) \neq t_1$
8	$T_i(z_4) = t_2$	$\Rightarrow$	$T_i(z_1) \neq t_3$ and $T_i(z_2) \notin \{t_1, t_3\}$ and $T_i(z_3) \notin \{t_1, t_3\}$
9	$T_i(z_1) = t_3$	$\Rightarrow$	$T_i(z_2) \notin \{t_1, t_2\}$ and $T_i(z_3) \notin \{t_1, t_2\}$ and $T_i(z_4) \notin \{t_1, t_2\}$
10	$T_i(z_2) = t_3$	$\Rightarrow$	$T_i(z_1) \neq t_2$ and $T_i(z_3) \notin \{t_1, t_2\}$ and $T_i(z_4) \notin \{t_1, t_2\}$
11	$T_i(z_3) = t_3$	$\Rightarrow$	$T_i(z_1) \neq t_2$ and $T_i(z_2) \notin \{t_1, t_2\}$ and $T_i(z_4) \notin \{t_1, t_2\}$
12	$T_i(z_4) = t_3$	$\Rightarrow$	$\varnothing$

This table presents all the choice restrictions generated by applying the choice rule (29) to each of the combination of choices $(t, t') \in \{t_1, t_2, t_3\}$ and instrumental values $(z, z') \in \{z_1, z_2, z_3, z_4\}$ of the incentive matrix (D.1).

Choice rule (29) converts the Incentive Matrix (D.1) into choice restrictions that determine the model response matrix $\boldsymbol{R}$. These choice restrictions are displayed in Table 3. Choice restrictions in Table 3 are in turn used to eliminate the response-types that are not economically justifiable.

Each counterfactual choice $T(z)$ of the response vector $\boldsymbol{S} = [T(z_1), T(z_2), T(z_3), T(z_4)]'$ takes up to three values in $\{t_1, t_2, t_3\}$. Thus, there are $3^4 = 81$ potential response types. The combination of all choice restrictions of Table 3 eliminate a total of 74 out of the 81 potential response-types. Response matrix $\boldsymbol{R}$ in (D.2) displays the resulting seven response-types that survive the elimination process.

$$\boldsymbol{R} = \begin{array}{ccccccccc} \boldsymbol{s}_1 & \boldsymbol{s}_2 & \boldsymbol{s}_3 & \boldsymbol{s}_4 & \boldsymbol{s}_5 & \boldsymbol{s}_6 & \boldsymbol{s}_7 & \boldsymbol{s}_8 & \\ t_1 & t_1 & t_1 & t_1 & t_1 & t_2 & t_2 & t_3 & z_1 \\ t_1 & t_1 & t_2 & t_2 & t_3 & t_2 & t_2 & t_3 & z_2 \\ t_1 & t_1 & t_2 & t_2 & t_3 & t_2 & t_2 & t_3 & z_3 \\ t_1 & t_3 & t_2 & t_3 & t_3 & t_2 & t_1 & t_3 & z_4 \end{array} \tag{D.2}$$

We use equations (**??**)–(**??**) to evaluate the causal parameters identified by response matrix (33) in the same fashion that Lemma **??** in Section 7.1 does. The response-matrix (33) enables the identification of eight response-type probabilities:

Point Identified	$P(\boldsymbol{S} = \boldsymbol{s}_1), P(\boldsymbol{S} = \boldsymbol{s}_2), P(\boldsymbol{S} = \boldsymbol{s}_5), P(\boldsymbol{S} = \boldsymbol{s}_8)$.
Partially Identified	$P(\boldsymbol{S} \in \{\boldsymbol{s}_3, \boldsymbol{s}_4\}), P(\boldsymbol{S} \in \{\boldsymbol{s}_3, \boldsymbol{s}_6\}), P(\boldsymbol{S} \in \{\boldsymbol{s}_4, \boldsymbol{s}_7\}), P(\boldsymbol{S} \in \{\boldsymbol{s}_6, \boldsymbol{s}_7\})$.

as well as the following counterfactual outcomes.

Always-takers	$E(Y(t_1)\|\boldsymbol{S} = \boldsymbol{s}_1)$	-	$E(Y(t_3)\|\boldsymbol{S} = \boldsymbol{s}_8)$
Switchers	$E(Y(t_1)\|\boldsymbol{S} = \boldsymbol{s}_2)$	-	$E(Y(t_3)\|\boldsymbol{S} = \boldsymbol{s}_5)$
Partially Identified	$E(Y(t_1)\|\boldsymbol{S} \in \{\boldsymbol{s}_3, \boldsymbol{s}_4, \boldsymbol{s}_5\})$	$E(Y(t_2)\|\boldsymbol{S} \in \{\boldsymbol{s}_4, \boldsymbol{s}_7\})$	$E(Y(t_3)\|\boldsymbol{S} \in \{\boldsymbol{s}_2, \boldsymbol{s}_4, \boldsymbol{s}_7\})$
		$E(Y(t_2)\|\boldsymbol{S} \in \{\boldsymbol{s}_6, \boldsymbol{s}_7\})$	
		$E(Y(t_2)\|\boldsymbol{S} \in \{\boldsymbol{s}_3, \boldsymbol{s}_6\})$	
		$E(Y(t_2)\|\boldsymbol{S} \in \{\boldsymbol{s}_3, \boldsymbol{s}_4\})$	

The identification results above state that only four out of nine response-type probabilities are point-identified. Most of the counterfactual outcomes are partially identified. Only four counterfactual outcome means are point-identified, none of these for choice t_2. In contrast, the unordered response matrix (32) in Lemma **??** secures the point-identification of all response-type probabilities and most of the counterfactual outcome means.

D.3 MM under the Double Randomization Design

We consider the emergence of MM in a "Double Randomization" design in which two vouchers are randomly assigned to the same sample of prospective students. The first voucher offers a tuition discount that applies to a natural science major. The second one applies to social science majors. We can divide the students into four groups:

1. Group z_1 does not receive any voucher.
2. Group z_2 receives only the social sciences voucher (t_3).
3. Group z_3 receives only the natural sciences voucher (t_2).
4. Group z_4 receives both the social sciences and natural sciences voucher.

Assuming the social sciences and natural sciences vouchers are of the same amount and that students cannot double major (so that they can only apply one voucher at a time), the IV design described above can be summarized by the incentive matrix in (D.3).

$$\boldsymbol{L} = \begin{array}{cc} \begin{array}{ccc} t_1 & t_2 & t_3 \end{array} & \\ \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} & \begin{array}{c} z_1 \\ z_2 \\ z_3 \\ z_4 \end{array} \end{array} \tag{D.3}$$

Table 4: Choice Restrictions generated by Incentive Matrix (34)

1	$T_i(z_1) = t_1$	$\Rightarrow$	$T_i(z_2) \neq t_2$ and $T_i(z_3) \neq t_3$
2	$T_i(z_2) = t_1$	$\Rightarrow$	$T_i(z_1) \notin \{t_2, t_3\}$ and $T_i(z_3) \neq t_3$ and $T_i(z_4) \neq t_3$
3	$T_i(z_3) = t_1$	$\Rightarrow$	$T_i(z_1) \notin \{t_2, t_3\}$ and $T_i(z_2) \neq t_2$ and $T_i(z_4) \neq t_2$
4	$T_i(z_4) = t_1$	$\Rightarrow$	$T_i(z_1) \notin \{t_2, t_3\}$ and $T_i(z_2) \notin \{t_2, t_3\}$ and $T_i(z_3) \notin \{t_2, t_3\}$
5	$T_i(z_1) = t_2$	$\Rightarrow$	$T_i(z_2) \neq t_1$ and $T_i(z_3) \notin \{t_1, t_3\}$ and $T_i(z_4) \notin \{t_1, t_3\}$
6	$T_i(z_2) = t_2$	$\Rightarrow$	$T_i(z_1) \notin \{t_1, t_3\}$ and $T_i(z_3) \notin \{t_1, t_3\}$ and $T_i(z_4) \notin \{t_1, t_3\}$
7	$T_i(z_3) = t_2$	$\Rightarrow$	$T_i(z_4) \neq t_1$
8	$T_i(z_4) = t_2$	$\Rightarrow$	$T_i(z_1) \neq t_3$ and $T_i(z_3) \notin \{t_1, t_3\}$
9	$T_i(z_1) = t_3$	$\Rightarrow$	$T_i(z_2) \notin \{t_1, t_2\}$ and $T_i(z_3) \neq t_1$ and $T_i(z_4) \notin \{t_1, t_2\}$
10	$T_i(z_2) = t_3$	$\Rightarrow$	$T_i(z_4) \neq t_1$
11	$T_i(z_3) = t_3$	$\Rightarrow$	$T_i(z_1) \notin \{t_1, t_2\}$ and $T_i(z_2) \notin \{t_1, t_2\}$ and $T_i(z_4) \notin \{t_1, t_2\}$
12	$T_i(z_4) = t_3$	$\Rightarrow$	$T_i(z_1) \neq t_2$ and $T_i(z_2) \notin \{t_1, t_2\}$

This table presents all the choice restrictions generated by applying the choice rule (29) to each of the combination of choices $(t, t') \in \{t_1, t_2, t_3\}$ and instrumental values $(z, z') \in \{z_1, z_2, z_3, z_4\}$ of the incentive matrix (34).

Applying the Choice Rule (29) from above generates the choice restrictions of Table 4. These in turn generate the response matrix $\boldsymbol{R}$ in (D.4).

$$\boldsymbol{R} = \begin{array}{cc} \begin{array}{ccccccccc} \boldsymbol{s}_1 & \boldsymbol{s}_2 & \boldsymbol{s}_3 & \boldsymbol{s}_4 & \boldsymbol{s}_5 & \boldsymbol{s}_6 & \boldsymbol{s}_7 & \boldsymbol{s}_8 & \boldsymbol{s}_9 \end{array} & \\ \begin{bmatrix} t_1 & t_1 & t_1 & t_1 & t_1 & t_2 & t_2 & t_3 & t_3 \\ t_1 & t_1 & t_3 & t_3 & t_3 & t_2 & t_3 & t_3 & t_3 \\ t_1 & t_2 & t_1 & t_2 & t_2 & t_2 & t_2 & t_2 & t_3 \\ t_1 & t_2 & t_3 & t_2 & t_3 & t_2 & t_2 & t_3 & t_3 \end{bmatrix} & \begin{array}{c} z_1 \\ z_2 \\ z_3 \\ z_4 \end{array} \end{array} \tag{D.4}$$

Applying equations (**??**)–(**??**) to response-matrix (D.4) gives that all response-type probabilities are identified, $P(\boldsymbol{S} = \boldsymbol{s}_j)$; $j = 1, ..., 9$, as well as the following counterfactual outcomes:

Always-takers	$E(Y(t_0) \mid \boldsymbol{S} = \boldsymbol{s}_1)$	$E(Y(t_1) \mid \boldsymbol{S} = \boldsymbol{s}_6)$	$E(Y(t_2) \mid \boldsymbol{S} = \boldsymbol{s}_9)$
Switchers	$E(Y(t_0) \mid \boldsymbol{S} = \boldsymbol{s}_2)$ $E(Y(t_0) \mid \boldsymbol{S} = \boldsymbol{s}_3)$	$E(Y(t_1) \mid \boldsymbol{S} = \boldsymbol{s}_7)$	$E(Y(t_2) \mid \boldsymbol{S} = \boldsymbol{s}_8)$
Partially Identified	$E(Y(t_0) \mid \boldsymbol{S} \in \{\boldsymbol{s}_4, \boldsymbol{s}_5\})$	$E(Y(t_1) \mid \boldsymbol{S} \in \{\boldsymbol{s}_2, \boldsymbol{s}_4\})$ $E(Y(t_1) \mid \boldsymbol{S} \in \{\boldsymbol{s}_5, \boldsymbol{s}_8\})$	$E(Y(t_2) \mid \boldsymbol{S} \in \{\boldsymbol{s}_3, \boldsymbol{s}_5\})$ $E(Y(t_2) \mid \boldsymbol{S} \in \{\boldsymbol{s}_4, \boldsymbol{s}_7\})$

D.4 MM under the Extensive Margin Compliers Only (EMCO) Design

We revisit the incentive design described in (35), presented again in $\boldsymbol{L}$ (D.5) below

$$\boldsymbol{L} = \begin{array}{cc} & \begin{array}{ccc} t_0 & t_1 & t_2 \end{array} \\ & \left[\begin{array}{ccc} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{array}\right] \begin{array}{c} z_1 \\ z_2 \\ z_3 \\ z_4 \end{array} \end{array} \tag{D.5}$$

Applying the Choice Rule (29) to the incentive design summarized in (D.5) we generate the following choice restrictions. These choice restrictions will in turn be used to eliminate response types, i.e restrict supp($\boldsymbol{S}$).

Table 5: Choice Restrictions generated by Incentive Matrix (D.5)

1	$T_i(z_1) = t_1$	$\Rightarrow$	$T_i(z_2) \notin \{t_2, t_3\}$ and $T_i(z_4) \notin \{t_2, t_3\}$
2	$T_i(z_2) = t_1$	$\Rightarrow$	$\varnothing$
3	$T_i(z_3) = t_1$	$\Rightarrow$	$T_i(z_1) \notin \{t_2, t_3\}$ and $T_i(z_2) \notin \{t_2, t_3\}$ and $T_i(z_4) \notin \{t_2, t_3\}$
4	$T_i(z_4) = t_1$	$\Rightarrow$	$T_i(z_1) \notin \{t_2, t_3\}$ and $T_i(z_2) \notin \{t_2, t_3\}$
5	$T_i(z_1) = t_2$	$\Rightarrow$	$T_i(z_2) \neq t_3$ and $T_i(z_3) \notin \{t_1, t_3\}$ and $T_i(z_4) \notin \{t_1, t_3\}$
6	$T_i(z_2) = t_2$	$\Rightarrow$	$T_i(z_1) \notin \{t_1, t_3\}$ and $T_i(z_3) \notin \{t_1, t_3\}$ and $T_i(z_4) \notin \{t_1, t_3\}$
7	$T_i(z_3) = t_2$	$\Rightarrow$	$T_i(z_1) \neq t_3$ and $T_i(z_2) \neq t_3$ and $T_i(z_4) \neq t_3$
8	$T_i(z_4) = t_2$	$\Rightarrow$	$T_i(z_1) \notin \{t_1, t_3\}$ and $T_i(z_2) \neq t_3$ and $T_i(z_3) \notin \{t_1, t_3\}$
9	$T_i(z_1) = t_3$	$\Rightarrow$	$T_i(z_2) \neq t_2$ and $T_i(z_3) \notin \{t_1, t_2\}$ and $T_i(z_4) \notin \{t_1, t_2\}$
10	$T_i(z_2) = t_3$	$\Rightarrow$	$T_i(z_1) \notin \{t_1, t_2\}$ and $T_i(z_3) \notin \{t_1, t_2\}$ and $T_i(z_4) \notin \{t_1, t_2\}$
11	$T_i(z_3) = t_3$	$\Rightarrow$	$T_i(z_1) \neq t_2$ and $T_i(z_2) \neq t_2$ and $T_i(z_4) \neq t_2$
12	$T_i(z_4) = t_3$	$\Rightarrow$	$T_i(z_1) \notin \{t_1, t_2\}$ and $T_i(z_2) \neq t_2$ and $T_i(z_3) \notin \{t_1, t_2\}$

This table presents all the choice restrictions generated by applying the choice rule (29) to each of the combination of choices $(t, t') \in \{t_1, t_2, t_3\}$ and instrumental values $(z, z') \in \{z_1, z_2, z_3, z_4\}$ of the incentive matrix (D.5).

After exhausting the choice restrictions in Table 5 we are left with 7 out of a possible 81 response types. These response types are consolidated and displayed in the response matrix $\boldsymbol{R}$ (D.6) below.

$$\boldsymbol{R} = \begin{array}{c} \\ \\ \\ \\ \\ \end{array}\begin{array}{ccccccc|c} \boldsymbol{s}_1 & \boldsymbol{s}_2 & \boldsymbol{s}_3 & \boldsymbol{s}_4 & \boldsymbol{s}_5 & \boldsymbol{s}_6 & \boldsymbol{s}_7 & \\ t_1 & t_1 & t_1 & t_2 & t_2 & t_3 & t_3 & z_1 \\ t_1 & t_1 & t_1 & t_1 & t_2 & t_1 & t_3 & z_2 \\ t_1 & t_2 & t_3 & t_2 & t_2 & t_3 & t_3 & z_3 \\ t_1 & t_1 & t_1 & t_2 & t_2 & t_3 & t_3 & z_4 \end{array} \tag{D.6}$$

We can apply the identification results of Heckman and Pinto (2018) to response-matrix (D.6) in order to identify all the response-type probabilities $P(\boldsymbol{S} = \boldsymbol{s}_j)$; $j = 1, ..., 7$ as well as the following counterfactual outcomes:

Always-takers	$E(Y(t_1)\|S = s_1)$	$E(Y(t_2)\|S = s_5)$	$E(Y(t_3)\|S = s_7)$
Switchers		$E(Y(t_2)\|S = s_2)$	$E(Y(t_3)\|S = s_3)$
		$E(Y(t_2)\|S = s_4)$	$E(Y(t_3)\|S = s_6)$
Partially Identified	$E(Y(t_1)\|S \in \{s_2, s_3\})$		
	$E(Y(t_1)\|S \in \{s_4, s_6\})$		

D.5 MM under Orthogonal Array Design

We additionally examine an IV choice model based on the popular orthogonal array experimental design. Orthogonal arrays are a widely popular experimental design developed by CD Rao (Rao, 1946a,b, 1947, 1949). Orthogonal arrays are widely used in Agricultural and Industrial sciences to determine the optimum mix of treatments that maximize production yield. The method is based on the random assignment of a combinatorial arrangements of treatments for each randomization arm. We adapt this setup to an instrumental variable setting by exogenously providing incentives for one or more treatments instead of directly assigning agents to treatment arms. Below, we will see that this incentive structure allows for a broad range of identification results.

Formally, a binary orthogonal array is a matrix of zeros and ones such that any two-column submatrix displays all possible combinations of zeros and ones. In other words, the tuples

$$\{(0,0), (0,1), (1,0), (1,1)\}$$

are all rows in any two-column submatrix. An orthogonal array incentive design if its associated incentive matrix is a binary orthogonal array. The incentive matrix in (40) displays an example of an orthogonal array incentive design. In context of the college choice example, we can rationalize the orthogonal array incentive design (40) with the following research design:

1. Group z_1 receives a cash voucher if they choose to major in the natural sciences (t_2) or the social sciences (t_3).
2. Group z_2 receives no cash voucher.
3. Group z_3 receives a cash voucher if they do not go to college (t_1) or if they major in the natural sciences (t_2).
4. Group z_3 receives a cash voucher if they do not go to college (t_1) or if they major in the social sciences (t_3).

Table 6 displays the choice restrictions generated by applying the Choice Rule (29) to the orthogonal array incentive design (40). After using these choice restrictions to eliminate response types, we are left with nine total response types summarized in the response matrix (D.7).

Table 6: Choice Restrictions generated by Incentive Matrix (40)

1	$T_i(z_1) = t_1$	$\Rightarrow$	$T_i(z_2) \notin \{t_2, t_3\}$ and $T_i(z_3) \notin \{t_2, t_3\}$ and $T_i(z_4) \notin \{t_2, t_3\}$
2	$T_i(z_2) = t_1$	$\Rightarrow$	$T_i(z_3) \notin \{t_2, t_3\}$ and $T_i(z_4) \notin \{t_2, t_3\}$
3	$T_i(z_3) = t_1$	$\Rightarrow$	$T_i(z_2) \neq t_2$ and $T_i(z_4) \neq t_2$
4	$T_i(z_4) = t_1$	$\Rightarrow$	$T_i(z_2) \neq t_3$ and $T_i(z_3) \neq t_3$
5	$T_i(z_1) = t_2$	$\Rightarrow$	$T_i(z_2) \neq t_3$ and $T_i(z_3) \neq t_3$
6	$T_i(z_2) = t_2$	$\Rightarrow$	$T_i(z_1) \notin \{t_1, t_3\}$ and $T_i(z_3) \notin \{t_1, t_3\}$
7	$T_i(z_3) = t_2$	$\Rightarrow$	$T_i(z_1) \neq t_1$ and $T_i(z_2) \neq t_1$
8	$T_i(z_4) = t_2$	$\Rightarrow$	$T_i(z_1) \notin \{t_1, t_3\}$ and $T_i(z_2) \notin \{t_1, t_3\}$ and $T_i(z_3) \notin \{t_1, t_3\}$
9	$T_i(z_1) = t_3$	$\Rightarrow$	$T_i(z_2) \neq t_2$ and $T_i(z_4) \neq t_2$
10	$T_i(z_2) = t_3$	$\Rightarrow$	$T_i(z_1) \notin \{t_1, t_2\}$ and $T_i(z_4) \notin \{t_1, t_2\}$
11	$T_i(z_3) = t_3$	$\Rightarrow$	$T_i(z_1) \notin \{t_1, t_2\}$ and $T_i(z_2) \notin \{t_1, t_2\}$ and $T_i(z_4) \notin \{t_1, t_2\}$
12	$T_i(z_4) = t_3$	$\Rightarrow$	$T_i(z_1) \neq t_1$ and $T_i(z_2) \neq t_1$

This table presents all the choice restrictions generated by applying the choice rule (29) to each of the combination of choices $(t, t') \in \{t_1, t_2, t_3\}$ and instrumental values $(z, z') \in \{z_1, z_2, z_3, z_4\}$ of the incentive matrix (40).

$$\boldsymbol{R} = \begin{array}{c} \\ \\ \\ \\ \\ \end{array}\begin{array}{cccccccccc} \boldsymbol{s}_1 & \boldsymbol{s}_2 & \boldsymbol{s}_3 & \boldsymbol{s}_4 & \boldsymbol{s}_5 & \boldsymbol{s}_6 & \boldsymbol{s}_7 & \boldsymbol{s}_8 & \boldsymbol{s}_9 & \\ \left[\begin{array}{c} t_1 \\ t_1 \\ t_1 \\ t_1 \end{array}\right. & \begin{array}{c} t_2 \\ t_1 \\ t_1 \\ t_1 \end{array} & \begin{array}{c} t_2 \\ t_2 \\ t_2 \\ t_1 \end{array} & \begin{array}{c} t_2 \\ t_2 \\ t_2 \\ t_2 \end{array} & \begin{array}{c} t_2 \\ t_2 \\ t_2 \\ t_3 \end{array} & \begin{array}{c} t_3 \\ t_1 \\ t_1 \\ t_1 \end{array} & \begin{array}{c} t_3 \\ t_3 \\ t_1 \\ t_3 \end{array} & \begin{array}{c} t_3 \\ t_3 \\ t_2 \\ t_3 \end{array} & \left.\begin{array}{c} t_3 \\ t_3 \\ t_3 \\ t_3 \end{array}\right] & \begin{array}{c} z_1 \\ z_2 \\ z_3 \\ z_4 \end{array} \end{array} \tag{D.7}$$

This response matrix satisfies neither unordered nor ordered monotonicity. When the instrument switches from z_1 to z_4, agents in response type $\boldsymbol{s}_3$ move from treatment t_2 to treatment t_3 while agents in response type $\boldsymbol{s}_6$ move away from t_3 and towards t_1. This represents a violation of ordered monotonicity and also prevents t_3 from being ordered the highest or lowest in any ordering on $\mathcal{T}$ that would satisfy ordered monotonicity.[22] Similarly we can see a switch from z_3 to z_4 induces agents in response type $\boldsymbol{s}_3$ to move from treatment t_2 to treatment t_1 while inducing agents in response type $\boldsymbol{s}_7$ to move away from treatment t_1 and towards treatment t_3. This again represents a violation of unordered monotonicity and precents t_1 from being ordered either the highest or the lowest in any ordering $\mathcal{T}$ that would satisfy ordered monotonicity. Since all orderings on $\mathcal{T} = \{t_1, t_2, t_3\}$ must have either t_1 or t_3 as the largest or smallest element, this means there is no ordering on $\mathcal{T}$ that satisfies ordered monotonicity.

Despite this, we can once again use Theorem 3 to verify that this matrix does indeed satisfy MM. Thus we can still use 2SLS type estimands to recover interpretable causal parameters as defined in (20). Moreover, by applying (**??**)-(**??**) we can see that all response types probabilities $P(\boldsymbol{S} = \boldsymbol{s}_j)$, $j = 1, \ldots, 9$ are identified. Additionally, using (**??**)-(**??**) we obtain that the following counterfactual outcomes are identified

Always-takers	$E(Y(t_1)\vert\boldsymbol{S} = \boldsymbol{s}_1)$	$E(Y(t_2)\vert\boldsymbol{S} = \boldsymbol{s}_4)$	$E(Y(t_3)\vert\boldsymbol{S} = \boldsymbol{s}_9)$
Switchers	$E(Y(t_1)\vert\boldsymbol{S} = \boldsymbol{s}_3)$	$E(Y(t_2)\vert\boldsymbol{S} = \boldsymbol{s}_2)$	$E(Y(t_3)\vert\boldsymbol{S} = \boldsymbol{s}_5)$
	$E(Y(t_1)\vert\boldsymbol{S} = \boldsymbol{s}_7)$	$E(Y(t_2)\vert\boldsymbol{S} = \boldsymbol{s}_8)$	$E(Y(t_3)\vert\boldsymbol{S} = \boldsymbol{s}_6)$
Partially Identified	$E(Y(t_1)\vert\boldsymbol{S} \in \{\boldsymbol{s}_2, \boldsymbol{s}_6\})$	$E(Y(t_2)\vert\boldsymbol{S} \in \{\boldsymbol{s}_3, \boldsymbol{s}_5\})$	$E(Y(t_3)\vert\boldsymbol{S} \in \{\boldsymbol{s}_7, \boldsymbol{s}_8\})$

---

[22] If t_3 is ranked highest a movement away from t_3 represents moving towards a lower treatment while a towards t_3 represents moving towards a higher treatment. Vice versa, if t_3 is ranked lowest a movement towards t_3 represents moving towards a lower treatment while a movement away from t_3 represents moving towards a higher treatment.