NBER WORKING PAPER SERIES

UNORDERED MONOTONICITY

James J. Heckman
Rodrigo Pinto

Working Paper 23497
http://www.nber.org/papers/w23497

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2017

A version of this paper was first presented as the Presidential Address to the Econometric Society delivered in Taipei, June 22, 2014, under the title "Causal Models, Structural Equations, and Identification: Stratification and Instrumental Variables." An earlier version was presented by Heckman at the European Seminar on Bayesian Econometrics in Vienna, November 1, 2012, hosted by Sylvia Fruehwirth-Schnatter, under the title "Causal Analysis After Haavelmo: Definitions and a Unified Analysis of Identification of Causal Models." James J. Heckman is the Henry Schultz Distinguished Service Professor of Economics and Public Policy at the University of Chicago, Director of the Center for the Economics of Human Development, and a Research Fellow at the American Bar Foundation. Rodrigo Pinto is an assistant professor of economics at UCLA. We thank Joshua Shea and Anna Zi˛ for a close and perceptive read of the paper and helpful comments. We thank the participants in various seminars at the University of Chicago and members of the audience at various regional meetings of the Econometric Society for helpful comments. This research was supported in part by: the American Bar Foundation and NIH grant NICHD R37HD065072. The views expressed in this paper are solely those of the authors and do not necessarily represent those of the funders, nor the official views of the National Institutes of Health or National Bureau of Economic Research. The Web Appendix for this paper is posted at https://cehd.uchicago.edu/monotonicity_identifiability.

Unordered Monotonicity
James J. Heckman and Rodrigo Pinto
NBER Working Paper No. 23497
June 2017
JEL No. C16,C93,I21,J15

ABSTRACT

This paper presents a new monotonicity condition for unordered discrete choice models with multiple treatments. Unlike a less general version of mono-tonicity in binary and ordered choice models, monotonicity in unordered discrete choice models along with other standard assumptions does not necessarily identify causal effects defined by variation in instruments, although in some cases it does. Our condition implies and is implied by additive separability of the choice equations in terms of observables and unobservables. These results follow from properties of binary matrices developed in this paper. We investigate conditions under which Unordered Monotonicity arises as a consequence of choice behavior. We represent IV estimators of counterfactuals as solutions to discrete mixture problems.

James J. Heckman
Department of Economics
The University of Chicago
1126 E. 59th Street
Chicago, IL 60637
and IZA
and also NBER
jjh@uchicago.edu

Rodrigo Pinto
University of California, Los Angeles
Department of Economics
8283 Bunche Hall
Los Angeles, CA 90095
rodrig@econ.ucla.edu

A data appendix is available at http://www.nber.org/data-appendix/w23497

Contents

1 Introduction

The evaluation of economic policies is a central goal of econometrics.[1] Economists have long used instrumental variables (IV) to identify policy-relevant parameters.[2] Early econometricians used IV in linear models to identify parameters in systems of multiple equations. In those frameworks, economists could safely be agnostic about the models generating choices in estimating a variety of interesting policy counterfactuals if their instruments satisfied rank and exogeneity conditions.

This agnostic stance is not justified in models with heterogeneous responses in which decisions to take treatment are based on unobserved[3] components of those responses. Without additional assumptions, instrumental variables do not identify economically interpretable parameters. Choice mechanisms play a fundamental role in interpreting what instruments identify.

For binary and ordered versions of IV models, *monotonicity* facilitates interpretability. It requires that responses to changes in instruments move all people toward or against the same choices.[4] It is a condition about the uniformity of responses across all persons in response to changes in instruments.[5] In binary and ordered choice models, monotonicity coupled with standard IV assumptions allows economists to identify the causal effects on outcomes of changes in the choices induced by variation in the instruments.[6]

For a nonparametric binary choice Generalized Roy model, Vytlacil (2002) shows that monotonicity is equivalent to assuming that the treatment choice equation is characterized by an additively-separable latent-variable threshold-crossing model. Separability is defined in terms of observed and unobserved (by the economist) variables. Vytlacil (2006) extends his analysis to the case of *ordered* multiple choice models

[1] See the statement of purpose for the Econometric Society by Ragner Frisch (1933).

[2] Theil (1953, 1958) developed two stage least-squares—the leading instrumental variable estimator.

[3] By the economist.

[4] This concept is more accurately interpreted as "uniformity" and does not correspond to ordinary mathematical definitions of monotonicity. See Heckman and Vytlacil (2005, 2007a).

[5] See Heckman et al. (2006).

[6] For the binary choice model, this is the LATE parameter of Imbens and Angrist (1994). Their extension of LATE to situations with multiple choices assumes that indicators of choice are naturally ordered (e.g., years of schooling). It assumes a meaningful scalar aggregator can be constructed that is monotonic in the ordered indicators of choice (Angrist and Imbens, 1995). In general, LATE does not identify a variety of policy-relevant parameters. See Heckman and Vytlacil (2007b) or Heckman (2010).

where the order is placed on the possible outcome variables (e.g., years of schooling).

This paper contributes to the literature by analyzing a general model of unordered choices. We develop a new condition—*Unordered Monotonicity*—that applies to models of multiple choices without a natural order among the choice values. For example, consider the choice of a pet among the set { cat, dog, bird }. These choices are only ordered by the preferences of agents across choices and not necessarily by the characteristics of outcomes of choices. Unordered Monotonicity preserves the intuitive notion of weak uniformity of responses to changes in instruments across persons without imposing any cardinalization on choices or choice outcomes. It restricts the set that counterfactual choices can take as instruments vary. However, in a general model, unordered monotonicity along with standard instrumental variable assumptions do not necessarily identify causal parameters.

Unordered choice models are studied by Heckman et al. (2006), Heckman and Vytlacil (2007b), and Heckman et al. (2008) who identify a variety of economically relevant treatment effects. They assume that the equations generating choice of treatment are governed by additively separable threshold-crossing models. Their identification strategy relies critically on instruments that assume values on a continuum. They also invoke "identification at infinity," as does a large literature in structural economics.[7] In this paper, we show that these assumptions can be relaxed and under specified conditions identification of economically interpretable treatment effects can still be secured. We only rely on discrete-valued instruments—the case commonly encountered in empirical work.[8]

This paper establishes an equivalence result that connects Unordered Monotonicity with separability of choice equations. We do not impose separability on the underlying choice equations. However, Unordered Monotonicity implies and is implied by representations of choice equations that are additively separable in observed and unobserved variables. This equivalence arises from the properties of binary matrices that characterize choice sets. This paper introduces economists to the identifying and interpretive power of binary matrices. We show that the potential identifying prop-

[7] See Heckman and Vytlacil (2007b) and Blevins (2014).

[8] Continuity of instruments and full support produce identifiability in our model, but are not required. See Heckman and Pinto (2015a). In related work, Lee and Salanié (2016) use a general framework to investigate multivalued choice models defined by an arbitrary number of separable threshold-crossing rules. They show that the identification of causal effects is possible with enough variation in instrumental variables defined on a continuum.

erties of Unordered Monotonicity arise from the restrictions it poses on the kernels of discrete mixtures. For an empirical application of Unordered Monotonicity, see Pinto (2016a), who evaluates the Moving to Opportunity Experiment.

The paper proceeds in the following way. Section 2 defines a general model of multiple choices and categorical instrumental variables. Section 3 presents a general framework for studying identification in that model. Our framework is based on partitioning the population into strata corresponding to counterfactual treatment choices. Section 4 presents a new characterization of the IV identification problem using a finite mixture model with restrictions on admissible vectors of counterfactual choices. We state necessary and sufficient conditions for identifying causal parameters. We illustrate these conditions for a binary choice (LATE) model. We show the simplicity and power of our analytical framework by deriving Vytlacil's equivalence result (2002) in a transparent way. Section 5 defines *Unordered Monotonicity* and illustrates how Unordered Monotonicity arises from choice-theoretic models. Section 6 presents equivalence theorems that relate the properties of Unordered Monotonicity and the separability of choice equations. We interpret this equivalence in light of economic theory. Section 7 applies this analysis to identify causal parameters. Section 8 concludes.

2 A Choice-Theoretic Model of Instrumental Variables

Our model consists of five random variables defined on probability space $(\Omega, \mathcal{F}, P)$, two policy-invariant equations that determine causal relationships among the variables, and an independence condition:[9]

$$\textbf{Choice Equations}: \quad T = f_T(Z, \boldsymbol{V}) \tag{1}$$

$$\textbf{Outcome Equations}: \quad Y = f_Y(T, \boldsymbol{V}, \epsilon_Y) \tag{2}$$

$$\textbf{Independence Condition}: \quad \boldsymbol{V}, Z, \epsilon_Y \text{ are mutually independent.} \tag{3}$$

[9]By policy-invariant, we mean functions whose maps remain invariant under manipulation of the arguments. This is the notation of autonomy developed by Frisch (1938) and Haavelmo (1944). For a recent discussion of these conditions, see Heckman and Pinto (2015b).

Variables $(Z, T, Y, \epsilon_Y, \boldsymbol{V})$ have the following properties. **P1**: *Instrument* Z *is a categorical random variable with support* $\text{supp}(Z) = \{z_1, \ldots, z_{N_Z}\}$*;*[10] **P2**: *Treatment (or Choice) indicator* T *is a discrete-valued random variable with support* $\text{supp}(T) = \{t_1, \ldots, t_{N_T}\}$*;* **P3**: Y *is an observed random variable denoting outcomes arising from treatment;* **P4**: ϵ_Y *is an unobserved error term;*[11] **P5**: $\boldsymbol{V}$ *is a* confounder—*an unobserved random vector (possibly infinite dimensional) affecting both choices and outcomes.* We assume that the expectation of Y exists, i.e., $E(|Y|) < \infty$. We also assume that the distribution of T varies conditional on each value of Z, that is, $P(T = t|Z = z) > 0$ for all $t \in \text{supp}(T)$ and $z \in \text{supp}(Z)$. Vector $(Z_\omega; T_\omega; Y_\omega; \boldsymbol{V}_\omega)$ denotes the realization of these variables for an element $\omega \in \Omega$. To simplify notation, background variables unaffected by treatment are kept implicit. Our analysis is conditional on such variables.

Counterfactual outcome $Y(t)$ is defined by fixing the argument T of the outcome Equation (2) to $t \in \text{supp}(T)$, that is, $Y(t) = f_Y(t, \boldsymbol{V}, \epsilon_Y)$. The observed outcome Y (Equation (2)) is the output of a Quandt (1972) switching regression model:

$$Y = \sum_{t \in \text{supp}(T)} Y(t) \cdot \mathbf{1}[T = t] \;\equiv\; Y(T), \tag{4}$$

where $\mathbf{1}[\alpha]$ is an indicator function that takes value 1 if α is true and 0 otherwise. Counterfactual choice $T(z) = f_T(z, \boldsymbol{V})$ is defined by fixing the argument Z of the choice Equation (1) to $z \in \text{supp}(Z)$.[12] Observed choice is given by

$$T = \sum_{z \in \text{supp}(Z)} T(z) \cdot \mathbf{1}[Z = z] \;\equiv\; T(Z). \tag{5}$$

Remark 2.1. The binary Generalized Roy Model (Heckman and Vytlacil, 2007a) is a special case of this model in which $\boldsymbol{V}$ *is a scalar random variable* V*, the choice is binary* $T \in \{0, 1\}$*, and the choice equation is defined by an indicator function*

[10] The assumption that Z is a multivalued scalar is a convenience. We can vectorize a matrix of instruments into a scalar form. Thus, we accommodate multiple instruments defined in the usual way.

[11] Such errors terms are often called "shocks" in structural equation models. f_T is a random function that could be written as a deterministic function if we introduced shock ϵ_T of arbitrary dimension as an argument of the function, where ϵ_T is independent of $\boldsymbol{V}$ and ϵ_Y.

[12] *Fixing* is a causal operation that captures the notion of external (*ceteris paribus*) manipulation. It is central concept in the study of causality and dates back to (Haavelmo, 1943). See Heckman and Pinto (2015b) for a recent discussion of fixing and causality.

that is separable in Z and V, namely $T = f_T(Z,V) \equiv \mathbf{1}[\tau(Z) \geq V]$. In this paper, we analyze multiple choices and impose no restriction on the functional forms of the choice Equation (1) or outcome Equation (2). Instead, we make restrictions on counterfactual choices and examine how those restrictions affect the characterization of choice equations.

Independence condition (3) generates the following model properties:

$$\textbf{Exclusion Restriction:} \quad (\boldsymbol{V}, Y(t)) \perp\!\!\!\perp Z \tag{6}$$

$$\textbf{Conditional Independence (Matching) Property:} \quad Y(t) \perp\!\!\!\perp T|\boldsymbol{V}. \tag{7}$$

Equation (6) states that instrument Z is independent of counterfactual outcome $Y(t)$ and the confounding variable $\boldsymbol{V}$ that generates selection bias. It implies that instrument Z affects Y only through its effect on T. Equation (7) states that $Y(t)$ is independent of treatment choice T after conditioning on $\boldsymbol{V}$. Counterfactual outcomes can be evaluated by conditioning on $\boldsymbol{V}$:

$$E(Y(t)|V) = E(Y(t)|\boldsymbol{V}, T=t) = E\left(\left(\sum_{t' \in \text{supp}(T)} Y(t') \cdot \mathbf{1}[T=t']\right)|\boldsymbol{V}, T=t\right) = E(Y|\boldsymbol{V}, T=t). \tag{8}$$

Any solution to the problem of selection bias requires that the analyst controls for, or balances, unobserved $\boldsymbol{V}$ across treatment and control states.[13]

We control for $\boldsymbol{V}$ by partitioning the sample space Ω so that the treatment indicator T is independent of counterfactual outcomes within each partition set. Consider a partition of Ω: $\Omega = \cup_{n=1}^{N} \Omega_n$; $\Omega_n \cap \Omega_{n'} = \varnothing, \forall\, n, n' \in \{1, \ldots, N\}, n \neq n'$, with an associated indicator H_ω that takes the value $n \in \{1, \ldots, N\}$ if $\omega \in \Omega_n$, i.e., $H_\omega = \sum_{n=1}^{N} n \cdot \mathbf{1}[\omega \in \Omega_n]$. If the following relationship holds within each partition,

$$Y(t) \perp\!\!\!\perp T|(H=n); \;\; \forall\; n \in \{1, \ldots, N\}, \tag{9}$$

T is effectively randomly assigned conditional on $H = n$. If such partitions were known, one could apply the logic underlying Equation (8) to evaluate counterfactual outcome $E(Y(t)|H = n)$ using $E(Y|T = t, H = n)$. If T takes the value t with strictly positive probability in all partition sets, i.e., $\Pr(T = t|H = n) > 0$; $n \in \{1, \ldots, N\}$, $E(Y(t))$ can be constructed from $E(Y(t)) = \sum_{n=1}^{N} E(Y|T = t, H =$

[13]Counterfactual $E(Y(t))$ and conditional expectation $E(Y|T = t)$ differ if the conditional and unconditional distributions of $\boldsymbol{V}$ are different: $E(Y(t)) = \int E(Y(t)|\boldsymbol{V} = \boldsymbol{v})dF_{\boldsymbol{V}}(\boldsymbol{v}) \neq \int E(Y|\boldsymbol{V} = \boldsymbol{v}, T = t)dF_{\boldsymbol{V}|T=t}(\boldsymbol{v}) = E(Y|T = t)$ where $F_{\boldsymbol{V}}$ is the CDF of $\boldsymbol{V}$ and $F_{\boldsymbol{V}|T=t}$ is the CDF of $\boldsymbol{V}$ conditional on $T = t$. See Heckman and Pinto (2015b).

$n)P(H = n)$. Our identification strategy uses instrumental variable Z to generate partitions $\{\Omega_n\}_{n=1}^{N}$ that satisfy Equation (9). To do so we use *response vectors* which we define next.

3 Response Vectors and the Identification Problem

Central to our analysis is the concept of *Response Vector* $\boldsymbol{S}$, a N_Z-dimensional random vector of counterfactual treatment choices T for Z fixed at each value of its support:

$$\boldsymbol{S} = [T(z_1)\,,\ \ldots\ ,\ T(z_{N_Z})]' = [f_T(\boldsymbol{V}, z_1)\,,\ \ldots\ ,\ f_T(\boldsymbol{V}, z_{N_Z})]' \equiv f_S(\boldsymbol{V}), \tag{10}$$

where $T(z)$ denotes a counterfactual treatment choice when instrumental variable Z is fixed at $z \in \text{supp}(Z)$. Let $\text{supp}(\boldsymbol{S}) = \{\boldsymbol{s}_1, \cdots, \boldsymbol{s}_{N_S}\}$ denote the finite support of $\boldsymbol{S}$. The N_Z-dimensional vectors $\boldsymbol{s} \in \text{supp}(\boldsymbol{S})$ are termed *response-types* or *strata.*[14] $\boldsymbol{S}$ plays a fundamental role in our analysis. T is related to $\boldsymbol{S}$ in the following way:

$$T = [\mathbf{1}[Z = z_1]\,,\ \ldots\ ,\ \mathbf{1}[Z = z_{N_Z}]] \cdot S \equiv g_T(\boldsymbol{S}, Z).^{15} \tag{11}$$

Equation (10) uses the fact that after fixing $Z = z$, $\boldsymbol{S}$ is a function only of unobserved $\boldsymbol{V}$. Conditioning on $\boldsymbol{S}$ effectively conditions on the regions of $\boldsymbol{V}$ that map into $\boldsymbol{S}$ by Equation (10).[16] It is a coarse way of conditioning on $\boldsymbol{V}$.

3.1 Properties of Response Vectors

Lemma **L-1** establishes four useful properties of response vectors analogous to properties shared with $\boldsymbol{V}$.

[14]Different notions of response vectors are used in the literature. In our notation, response vectors correspond to the choices a person of type $\boldsymbol{V}$ would make when confronted by different values of Z. Robins and Greenland (1992) initiated the literature. Frangakis and Rubin (2002) use the term "principal strata." They do not explicitly model $\boldsymbol{V}$ or use the econometric framework (1)-(3) so the relationship between strata and $\boldsymbol{V}$ and the fact that conditioning on $\boldsymbol{S}$ is equivalent to conditioning on regions of $\boldsymbol{V}$ is only implicit in their analysis. $T(z)$ can potentially take as many as N_T values for each value $z \in \text{supp}(Z)$. Since Z has $|\,\text{supp}(Z)| = N_Z$ elements, $\text{supp}(\boldsymbol{S})$ can have at most N_T^{Nz} elements.

[15]Figure B.3 in Web Appendix B displays our IV model with response vector $\boldsymbol{S}$ as a Directed Acyclic Graph (DAG).

[16]The regions are distinct because $f_T(\cdot)$ is a function.

Lemma L-1. *The following relationships for* $\boldsymbol{S}$ *hold for IV model* (1)–(3)*:*

$$\text{(i) } Y(t) \perp\!\!\!\perp T|\boldsymbol{S}, \quad \text{(ii) } \boldsymbol{S} \perp\!\!\!\perp Z, \quad \text{(iii) } Y \perp\!\!\!\perp T|(\boldsymbol{S}, Z), \quad \text{(iv) } Y \perp\!\!\!\perp Z|(\boldsymbol{S}, T).$$

Proof. See Web Appendix A.1. □

Relationship (i) states that counterfactual outcomes $Y(t)$ for all $t \in \text{supp}(T)$ are independent of treatment choices conditional on $\boldsymbol{S}$. Thus $\boldsymbol{S}$ shares the same conditional independence (matching) properties as $\boldsymbol{V}$ in (7). Relationship (ii) states that the potential treatment choices in $\boldsymbol{S}$ are independent of the instrumental variables. Relationship (iii) states that outcomes are independent of treatment choices conditional on $\boldsymbol{S}$ and Z. Indeed, from Equation (11), T is deterministic conditional on $\boldsymbol{S}$ and Z. Relationship (iv) is closely related to (iii). It states that outcome Y is independent of instrumental variable Z when conditioned on $\boldsymbol{S}$ and T.

Remark 3.1. Response vector $\boldsymbol{S}$ *generates a partition of the sample space* Ω *that has independence property* (9)*. Function* $f_S : \text{supp}(\boldsymbol{V}) \to \text{supp}(\boldsymbol{S})$ *in* (10) *is constructed using function* f_T *defined by* (1)*. Thus, for each* $\omega \in \Omega$*, there is a single value* $\boldsymbol{v} \in \text{supp}(\boldsymbol{V})$ *such that* $\boldsymbol{V}_\omega = \boldsymbol{v}$ *and a single value* $\boldsymbol{s} \in \text{supp}(\boldsymbol{S})$ *such that* $f_S(\boldsymbol{v}) = \boldsymbol{s}$*. We define a partition of the sample space* Ω *by:*

$$\Omega_n = \{\omega \in \Omega; f_S(\boldsymbol{V}_\omega) = \boldsymbol{s}_n\} \text{ for each } \boldsymbol{s}_n \in \text{supp}(\boldsymbol{S}). \tag{12}$$

In partition (12)*,* $\boldsymbol{S}_\omega = \boldsymbol{s}_n$ *and* $\omega \in \Omega_n$ *are equivalent. This partition satisfies* (9) *because* $Y(t) \perp\!\!\!\perp T|(\omega \in \Omega_n)$ *holds due to item (i) of Lemma* ***L-1****. Hence treatment choice can be interpreted as being randomly assigned conditional on* $\boldsymbol{S}$*. Indeed, conditional on* $\boldsymbol{S}$*, treatment* T *only depends on* Z *which is statistically independent of* $\boldsymbol{V}$*.*

Response vector $\boldsymbol{S}$ is a *balancing score* for $\boldsymbol{V}$.[17] It exploits the properties of instruments Z to generate a coarse partition of unobserved variable $\boldsymbol{V}$ while maintaining the independence properties arising from conditioning on $\boldsymbol{V}$. The matching condition $Y(t) \perp\!\!\!\perp T|\boldsymbol{S}$ is analogous to $Y(t) \perp\!\!\!\perp T|\boldsymbol{V}$ in (7). If $\boldsymbol{S}$ (or $\boldsymbol{V}$) were known, counterfactual outcomes can be identified by conditioning on $\boldsymbol{S}$ or $\boldsymbol{V}$.[18] Thus, $\boldsymbol{S}$ plays the role of a control function (Heckman and Robb, 1985). From Equation (8), $Y(t) \perp\!\!\!\perp T|\boldsymbol{S}$ implies that $E(Y(t)|\boldsymbol{S} = \boldsymbol{s}) = E(Y|T = t, \boldsymbol{S} = \boldsymbol{s})$. If $P(T = t|\boldsymbol{S} = \boldsymbol{s}) > 0$ for all

[17] $\boldsymbol{S}$ being a balancing score means that properties of $\boldsymbol{V}$ are inherited by $\boldsymbol{S}$. Formally, $\boldsymbol{S} = f_S(\boldsymbol{V})$ is a surjective function of $\boldsymbol{V}$ that satisfies $Y(t) \perp\!\!\!\perp T|\boldsymbol{V} \Rightarrow Y(t) \perp\!\!\!\perp T|f_S(\boldsymbol{V})$, and $\sigma(\boldsymbol{S}) \subseteq \sigma(\boldsymbol{V})$ where σ denotes a σ-algebra in the probability space $(\Omega, \mathcal{F}, P)$.

[18] See Heckman (2008) for a survey of a wide array of methods that implement this principle.

$\boldsymbol{s} \in \text{supp}(\boldsymbol{S})$, counterfactual mean outcomes can be expressed as:

$$E(Y(t)) = \sum_{\boldsymbol{s}\in\text{supp}(\boldsymbol{S})} E(Y(t)|\boldsymbol{S}=\boldsymbol{s})P(\boldsymbol{S}=\boldsymbol{s}) = \sum_{\boldsymbol{s}\in\text{supp}(\boldsymbol{S})} E(Y|T=t,\boldsymbol{S}=\boldsymbol{s})P(\boldsymbol{S}=\boldsymbol{s}). \tag{13}$$

$\boldsymbol{S}$ acts as a coarse surrogate for $\boldsymbol{V}$ and identifies treatment effects *within strata* by balancing unobservables $\boldsymbol{V}$ across treatment states.

3.2 The Strata Identification Problem

The problem of identifying counterfactual mean outcomes defined over strata consists of identifying unobserved $E(Y(t)|\boldsymbol{S}=\boldsymbol{s})$ and $P(\boldsymbol{S}=\boldsymbol{s})$ for $\boldsymbol{s} \in \text{supp}(\boldsymbol{S})$ and $t \in \text{supp}(T)$, from observed $E(Y|T=t, Z=z)$ and $P(T=t|Z=z)$ for $z \in \text{supp}(Z)$ and $t \in \text{supp}(T)$. Theorem **T-1** uses the relationships of Lemma **L-1** to express unobserved objects in terms of observed ones.

Theorem T-1. *The following equality holds for the IV model* (1)–(3)*:*

$$E(\kappa(Y)\cdot \mathbf{1}[T=t]|Z) = \sum_{\boldsymbol{s}\in\text{supp}(\boldsymbol{S})} \mathbf{1}[T=t|\boldsymbol{S}=\boldsymbol{s},Z]E(\kappa(Y(t))|\boldsymbol{S}=\boldsymbol{s})P(\boldsymbol{S}=\boldsymbol{s}), \tag{14}$$

where $\kappa : \text{supp}(Y) \to \mathbb{R}$ is an arbitrary known function.

Proof. See Web Appendix A.2. □

Setting $\kappa(Y)$ to 1 generates the propensity score equality:[19]

$$P(T=t|Z=z) = \sum_{\boldsymbol{s}\in\text{supp}(\boldsymbol{S})} \mathbf{1}[T=t|\boldsymbol{S}=\boldsymbol{s},Z=z]P(\boldsymbol{S}=\boldsymbol{s}). \tag{15}$$

Replacing $\kappa(Y)$ by any variable X such that $X \perp\!\!\!\perp T|\boldsymbol{S}$, we obtain:[20]

$$E(X|T=t,Z)P(T=t|Z) = \sum_{\boldsymbol{s}\in\text{supp}(\boldsymbol{S})} \mathbf{1}[T=t|\boldsymbol{S}=\boldsymbol{s},Z]E(X|\boldsymbol{S}=\boldsymbol{s})P(\boldsymbol{S}=\boldsymbol{s}). \tag{16}$$

[19] If we set $\kappa(Y)=Y$, we equate expected values of observed outcomes with expected counterfactual outcomes. Setting $\kappa(Y)=\mathbf{1}[Y\le y]$, we equate the cumulative distribution function (CDF) of the observed outcome with the unobserved CDF of counterfactual outcomes.

[20] Candidates for X are baseline variables caused by $\boldsymbol{V}$. Knowledge of the X variables helps to identify the observed characteristics of persons within strata.

Remark 3.2. Equation (14) characterizes the problem of identifying counterfactual outcomes within strata. There are N_Z observed objects on the left-hand side for each $t \in \text{supp}(T)$ totalling $N_Z \cdot N_T$. Without further restrictions, the total number of latent response-types on the right-hand side is $N_T^{N_Z}$, i.e., the number of strata. Thus, the number of observed quantities ($N_T \cdot N_Z$) grows linearly in N_Z while the number of possible response-types ($N_T^{N_Z}$) grows geometrically in N_Z.[21] *Identification requires that constraints be placed on the number of admissible strata ($\boldsymbol{S}$). Choice theory can produce such restrictions, as can other assumptions, such as those about functional forms.*

Indicator $\mathbf{1}[T = t|\boldsymbol{S} = \boldsymbol{s}, Z = z]$ in Equation (14) is deterministic because T is deterministic given Z and $\boldsymbol{S}$ in Equation (11). Our identification strategy develops economically interpretable restrictions on these indicators that govern the choice of treatment as Z varies. Such restrictions reduce the number of admissible response-types and characterize the indicators $\mathbf{1}[T = t|\boldsymbol{S} = \boldsymbol{s}, Z = z]$, facilitating identification of causal parameters.

We note, for later use, that the probability of treatment choice conditional on response-types is

$$\begin{aligned} P(T = t|\boldsymbol{S} = \boldsymbol{s}) &= \sum_{z \in \text{supp}(Z)} \mathbf{1}[T = t|\boldsymbol{S} = \boldsymbol{s}, Z = z] P(Z = z|\boldsymbol{S} = \boldsymbol{s}), \\ &= \sum_{z \in \text{supp}(Z)} \mathbf{1}[T = t|\boldsymbol{S} = \boldsymbol{s}, Z = z] P(Z = z), \end{aligned} \tag{17}$$

where the last equality is a consequence of $\boldsymbol{S} \perp\!\!\!\perp Z$ (item (ii) of Lemma **L-1**).

Note that Equation (14) is a discrete mixture latent class model, a feature we exploit below.[22] Our paper differs from previous work on nonparametric instrumental variables. Instead of forming the usual nonparametric IV moment equations (see, e.g., Carrasco et al., 2007), we use instruments to construct strata that generate the kernels of finite mixture equations and choice theory to place restrictions on the kernels. We then use finite mixture methods to examine the identification of counterfactual outcomes.

[21] Across the two equation systems for T and scalar Y there are $(2 \cdot N_T - 1) \cdot N_Z$ observed quantities and $(N_T)^{2 \cdot N_Z + 1}$ unknown parameters.

[22] See, e.g., Clogg (1995) and Henry et al. (2014).

4 Identifying Response Probabilities and Counterfactual Outcomes

We now present general conditions for identifying response probabilities, counterfactual mean outcomes, and pre-program variables conditioned on strata. To do so, it is useful to express Equations (14)–(15) as a system of linear equations. Define $\boldsymbol{P}_Z(t) = [P(T = t|Z = z_1), \ldots, P(T = t|Z = z_{N_Z})]'$, the vector of observed choice probabilities ("propensity scores"). Define $\boldsymbol{P}_Z$ as the vector that stacks $\boldsymbol{P}_Z(t)$ across $t \in \text{supp}(T)$: $\boldsymbol{P}_Z = [\boldsymbol{P}_Z(t_1), \ldots, \boldsymbol{P}_Z(t_{N_T})]'$. $\boldsymbol{Q}_Z(t)$ is defined in an analogous fashion for outcomes defined for different values of T (i.e., multiplied by the treatment indicators). In a similar fashion, $\boldsymbol{L}_Z(t)$ stands for vector $\boldsymbol{X}$ such that $\boldsymbol{X} \perp\!\!\!\perp T|\boldsymbol{S}, Z$. The left-hand sides of Equations (14) and (16) are given respectively by: $\boldsymbol{Q}_Z(t) = [E(\kappa(Y) \cdot \mathbf{1}[T = t]|Z = z_1), \ldots, E(\kappa(Y) \cdot \mathbf{1}[T = t]|Z = z_{N_Z})]'$, and $\boldsymbol{L}_Z(t) = [E(\boldsymbol{X} \cdot \mathbf{1}[T = t]|Z = z_1), \ldots, E(\boldsymbol{X} \cdot \mathbf{1}[T = t]|Z = z_{N_Z})]'$, where $\boldsymbol{L}_Z = [\boldsymbol{L}_Z(t_1), \ldots, \boldsymbol{L}_Z(t_{N_T})]'$.

Let $\boldsymbol{P}_S$ be the vector of unobserved response probabilities $\boldsymbol{P}_S = [P(\boldsymbol{S} = \boldsymbol{s}_1), \ldots, P(\boldsymbol{S} = \boldsymbol{s}_{N_S})]'$ and $\boldsymbol{L}_S = [E(\boldsymbol{X} \cdot \mathbf{1}[\boldsymbol{S} = \boldsymbol{s}_1]), \ldots, E(\boldsymbol{X} \cdot \mathbf{1}[\boldsymbol{S} = \boldsymbol{s}_{N_S}])]'$ be the unobserved vector of $\boldsymbol{X}$-expectations times response indicators. We denote the vector of the expected outcomes multiplied by response indicators by: $\boldsymbol{Q}_S(t) = [E(\kappa(Y(t)) \cdot \mathbf{1}[\boldsymbol{S} = \boldsymbol{s}_1]), \ldots, E(\kappa(Y(t)) \cdot \mathbf{1}[\boldsymbol{S} = \boldsymbol{s}_{N_S}])]'$.

The following notation and concepts are used throughout the rest of this paper. Define response matrix $\boldsymbol{R}$ as an array of response-types defined over $\text{supp}(\boldsymbol{S})$, i.e., $\boldsymbol{R} = [\boldsymbol{s}_1, \ldots, \boldsymbol{s}_{N_S}]$. To avoid trivial degeneracies we delete redundant rows (where different values of Z produce the same pattern for T) and redundant columns (where the same choices are made for the same value of Z. Matrix $\boldsymbol{R}$ has dimension $N_Z \times N_S$. An element in the i-th row and n-th column of $\boldsymbol{R}$ is denoted by $\boldsymbol{R}[i, n] = (T|Z = z_i, \boldsymbol{S} = \boldsymbol{s}_n), i \in \{1, \cdots, N_Z\}, n \in \{1, \ldots, N_S\}$. We use $\boldsymbol{R}[i, \cdot]$ to denote the i-th row of $\boldsymbol{R}$, $\boldsymbol{R}[\cdot, n]$ for the n-th column $\boldsymbol{R}$. Let $\boldsymbol{B}_t$ denote a binary matrix of the same dimension as $\boldsymbol{R}$ and whose elements take value 1 if the respective element in $\boldsymbol{R}$ is equal to t and 0 otherwise. Notationally, we define an element in the i-th row and n-th column of matrix $\boldsymbol{B}_t$ by $\boldsymbol{B}_t[i, n] = \mathbf{1}[T = t|Z = z_i, \boldsymbol{S} = \boldsymbol{s}_n], i \in \{1, \cdots, N_Z\}, n \in \{1, \ldots, N_S\}$. We also use the short-hand notation $\boldsymbol{B}_t = \mathbf{1}[\boldsymbol{R} = t]$ to denote $\boldsymbol{B}_t$. Let $\boldsymbol{B}_T$ be a binary matrix of dimension $(N_Z \cdot N_T) \times N_S$ generated by stacking $\boldsymbol{B}_t$ as t ranges over $\text{supp}(T)$: $\boldsymbol{B}_T = [\boldsymbol{B}'_{t_1}, \ldots, \boldsymbol{B}'_{t_{N_T}}]'$.

In this notation, Equations (14), (15), and (16) can be written respectively as

$$\boldsymbol{Q}_Z(t) = \boldsymbol{B}_t \boldsymbol{Q}_S(t), \tag{18}$$

$$\boldsymbol{P}_Z = \boldsymbol{B}_T \boldsymbol{P}_S \tag{19}$$

$$\boldsymbol{L}_Z = \boldsymbol{B}_T \boldsymbol{L}_S. \tag{20}$$

If $\boldsymbol{B}_t$ and $\boldsymbol{B}_T$ were invertible, $\boldsymbol{Q}_S(t)$, $\boldsymbol{P}_S$, and $\boldsymbol{L}_S$ would be identified. However, such inverses do not always exist. In their place, we can use generalized inverses.

Let $\boldsymbol{B}_T^+$ and $\boldsymbol{B}_t^+$ be the Moore-Penrose pseudo-inverses[23] of matrices $\boldsymbol{B}_T$ and $\boldsymbol{B}_t$ respectively for $t \in \text{supp}(T)$. The following expressions are useful for characterizing the identification of response probabilities and counterfactual means:

$$\boldsymbol{K}_T = \boldsymbol{I}_{N_S} - \boldsymbol{B}_T^+ \boldsymbol{B}_T \text{ and } \boldsymbol{K}_t = \boldsymbol{I}_{N_S} - \boldsymbol{B}_t^+ \boldsymbol{B}_t; \;\; t \in \text{supp}(T), \tag{21}$$

where $\boldsymbol{I}_{N_S}$ denotes an identity matrix of dimension N_S. $\boldsymbol{K}_T$ and $\boldsymbol{K}_t$ are orthogonal projection matrices.[24]

Applying the Moore-Penrose inverse to Equations (18) and (19), we obtain:

$$\boldsymbol{P}_S = \boldsymbol{B}_T^+ \boldsymbol{P}_Z + \boldsymbol{K}_T \boldsymbol{\lambda} \tag{22}$$

$$\boldsymbol{Q}_S(t) = \boldsymbol{B}_t^+ \boldsymbol{Q}_Z(t) + \boldsymbol{K}_t \tilde{\boldsymbol{\lambda}} \tag{23}$$

where $\boldsymbol{\lambda}$ and $\tilde{\boldsymbol{\lambda}}$ are arbitrary N_S-dimensional vectors (same dimension as $\boldsymbol{P}_S$). In this notation, Theorem **T-2** states general conditions for identification of response probabilities and counterfactual means.

Theorem T-2. *For IV model (1)–(3), if there exists a real-valued N_S-dimensional vector $\boldsymbol{\xi}$ such that $\boldsymbol{\xi}'\boldsymbol{K}_T = \boldsymbol{0}$, then $\boldsymbol{\xi}'\boldsymbol{P}_S$ and $\boldsymbol{\xi}'\boldsymbol{L}_S$ are identified. In addition, if there exists a real-valued N_S-dimensional vector $\boldsymbol{\zeta}$ such that $\boldsymbol{\zeta}'\boldsymbol{K}_t = \boldsymbol{0}$, then $\boldsymbol{\zeta}'\boldsymbol{Q}_S(t)$ is identified.*

Proof. See Web Appendix A.3. □

Theorem **T-2** shows the identifying properties of the response matrix. For ex-

[23]The Moore-Penrose inverse of a matrix $\boldsymbol{A}$ is denoted by $\boldsymbol{A}^+$ and is defined by the four following properties: (1) $\boldsymbol{A}\boldsymbol{A}^+\boldsymbol{A} = \boldsymbol{A}$; (2) $\boldsymbol{A}^+\boldsymbol{A}\boldsymbol{A}^+ = \boldsymbol{A}^+$; (3) $\boldsymbol{A}^+\boldsymbol{A}$ is symmetric; (4) $\boldsymbol{A}\boldsymbol{A}^+$ is symmetric. The Moore-Penrose matrix $\boldsymbol{A}^+$ of a real matrix $\boldsymbol{A}$ is unique and always exists (Magnus and Neudecker, 1999).

[24]See Web Appendix A.3.

ample, suppose that $\boldsymbol{B}_T$ has full column-rank. Then $\boldsymbol{B}_T^+\boldsymbol{B}_T = \boldsymbol{I}_{N_S}$ and $\boldsymbol{K}_T = \boldsymbol{0}$. Therefore $\boldsymbol{\xi}'\boldsymbol{P}_S$ is identified for any real vector $\boldsymbol{\xi}$ of dimension N_S. In particular, $\boldsymbol{\xi}'\boldsymbol{P}_S$ is identified when $\boldsymbol{\xi}$ is set to be each column vector of the identity matrix $\boldsymbol{I}_{N_S}$. In that case, each n-th column of $\boldsymbol{I}_{N_S}$ identifies $P(\boldsymbol{S} = \boldsymbol{s}_n)$ and all the response-type probabilities are identified.[25]

Note that full-rank for $\boldsymbol{B}_T$ does not imply full-rank for each $\boldsymbol{B}_t, t \in \text{supp}(T)$. Therefore, the identification of the response-type probabilities does not automatically produce identification of corresponding mean counterfactual outcomes. Corollary **C-1** formalizes this discussion.

Corollary C-1. *The following relationships hold for the IV model* (1)–(3)*:*

$$\text{Vectors } \boldsymbol{P}_S \text{ and } \boldsymbol{L}_S \text{ are point-identified } \Leftrightarrow \text{rank}(\boldsymbol{B}_T) = N_S, \tag{24}$$

$$\text{Vector } \boldsymbol{Q}_S(t) \text{ is point-identified } \Leftrightarrow \text{rank}(\boldsymbol{B}_t) = N_S. \tag{25}$$

Also, if (25) *holds, then* $E(\kappa(Y(t)))$ *is identified by* $\iota'\boldsymbol{B}_t^+\boldsymbol{Q}_Z(t)$*, where* ι *is a* N_S*-dimensional vector of 1s.*

Proof. See Web Appendix A.5. □

Versions of Corollary **C-1** are found in the literature on the identifiability of finite mixtures.[26] Given binary matrices $\boldsymbol{B}_T$, $\boldsymbol{B}_t$, and $t \in \{1, N_T\}$ the problem of identifying $\boldsymbol{P}_S, \boldsymbol{L}_S$, and $\boldsymbol{Q}_S(t)$ is equivalent to the problem of identifying finite mixtures of distributions where the $\boldsymbol{B}_T$ and $\boldsymbol{B}_t$ play the roles of kernels of mixtures. Mixture components are the corresponding counterfactual outcomes conditional on the response-types and mixture probabilities are the response-type probabilities.

One approach to identifiability is to simply assume that conditions (24) and (25) apply to $\boldsymbol{R}$. A more satisfactory approach, and the one taken here and in Pinto (2016a), investigates how alternative specifications of choice relationships generate response matrices $\boldsymbol{R}$ that satisfy the identifiability requirements of Theorem **T-2** and Corollary **C-1**.

[25] See Section A.4 of the Web Appendix for bounds on the response-type probabilities and counterfactual outcomes.

[26] See, e.g., Yakowitz and Spragins (1968) and Prakasa Rao (1992).

4.1 Example: Binary Choice (LATE)

To familiarize the reader with our notation and concepts, and anticipate our generalization of it, consider the binary choice model implicit in the Local Average Treatment Effect (LATE). Treatment variable T takes two values: $T_\omega = t_1$ if agent ω chooses to be treated and $T_\omega = t_0$ if not. Instrument Z is binary valued ($\text{supp}(Z) = \{z_0, z_1\}$) with the property $0 < P(T = t_1|Z = z_0) < P(T = t_1|Z = z_1) < 1$. A standard example is the problem of identifying the causal effect of college education on income Y. Agent ω decides between going to college ($T_\omega = t_1$) or not ($T_\omega = t_0$). Instrumental variable Z represents randomly assigned college scholarships. For example, $Z_\omega = z_1$ if a scholarship is assigned to agent ω and $Z_\omega = z_0$ if agent ω does not receive a scholarship.

The response vector is $\boldsymbol{S} = [T(z_0)\ ,\ T(z_1)]'$. Without further restrictions, $\boldsymbol{S}$ can take four possible values described by the following response matrix:

$$\begin{array}{cccccc} & \boldsymbol{s}_1 & \boldsymbol{s}_2 & \boldsymbol{s}_3 & \boldsymbol{s}_4 & \\ \boldsymbol{R} = & \left[\begin{array}{c} t_1 \\ t_1 \end{array}\right. & \begin{array}{c} t_0 \\ t_1 \end{array} & \begin{array}{c} t_1 \\ t_0 \end{array} & \left.\begin{array}{c} t_0 \\ t_0 \end{array}\right] & \begin{array}{l} \text{values for } T(z_0) \\ \text{values for } T(z_1) \end{array} \end{array} . \tag{26}$$

In the language of LATE, the response-types $\boldsymbol{s}_1, \boldsymbol{s}_2, \boldsymbol{s}_3, \boldsymbol{s}_4$ are always-takers, compliers, defiers, and never-takers, respectively. $\boldsymbol{B}_{t_1}$ is the binary matrix that has the same dimension as $\boldsymbol{R}$, whose elements take value 1 if the corresponding element in $\boldsymbol{R}$ is t_1 and value 0 if the element in $\boldsymbol{R}$ is t_0. Thus, $\boldsymbol{B}_{t_1} = \mathbf{1}[\boldsymbol{R} = t_1]$ and $\boldsymbol{B}_{t_0} = \mathbf{1}[\boldsymbol{R} = t_0]$ indicate whether elements in $\boldsymbol{R}$ are equal to t_1 or t_0, respectively.[27]

The 4×4 binary matrix $\boldsymbol{B}_T = [\boldsymbol{B}'_{t_0}\ ,\ \boldsymbol{B}'_{t_1}]'$ has rank equal to 3, which is less than the number of response-types $N_S = 4$. Therefore, by Corollary **C-1**, neither response-type probabilities nor the counterfactual outcomes are point identified. To identify them, it is necessary to reduce the number of response-types.

LATE solves this non-identification problem by assuming that each agent ω can only change his decision in one direction as the instrument varies. The *monotonicity* condition of Imbens and Angrist (1994) is:

Assumption A-1. ***Monotonicity for the Binary Choice Model:*** *The following*

[27] We also have that $\boldsymbol{B}_{t_1} = \iota_{N_Z}\iota'_{N_S} - \boldsymbol{B}_{t_0}$, where ι_N denotes a N-dimensional vector of elements equal to 1.

inequalities hold for any $z, z' \in \text{supp}(Z)$ *:*

$$\mathbf{1}[T_\omega(z) = t_1] \geq \mathbf{1}[T_\omega(z') = t_1] \, \forall \, \omega \in \Omega \quad \text{or} \quad \mathbf{1}[T_\omega(z) = t_1] \leq \mathbf{1}[T_\omega(z') = t_1] \, \forall \, \omega \in \Omega. \tag{27}$$
[28]

In our example, condition **A-1** assumes that each agent is inclined to decide towards college if a scholarship is granted, i.e., $\mathbf{1}[T_\omega(z_1) = t_1] \geq \mathbf{1}[T_\omega(z_0) = t_1]$ for all $\omega \in \Omega$. This eliminates the response-type $\boldsymbol{s}_3$ (the defiers) in matrix (26), generating the following matrices:

$$\boldsymbol{R} = \begin{array}{c} \begin{array}{ccc} \boldsymbol{s}_1 & \boldsymbol{s}_2 & \boldsymbol{s}_4 \end{array} \\ \left[\begin{array}{ccc} t_1 & t_0 & t_0 \\ t_1 & t_1 & t_0 \end{array} \right] \end{array}, \boldsymbol{B}_{t_1} = \begin{array}{c} \begin{array}{ccc} \boldsymbol{s}_1 & \boldsymbol{s}_2 & \boldsymbol{s}_4 \end{array} \\ \left[\begin{array}{ccc} 1 & 0 & 0 \\ 1 & 1 & 0 \end{array} \right] \end{array}, \boldsymbol{B}_{t_0} = \begin{array}{c} \begin{array}{ccc} \boldsymbol{s}_1 & \boldsymbol{s}_2 & \boldsymbol{s}_4 \end{array} \\ \left[\begin{array}{ccc} 0 & 1 & 1 \\ 0 & 0 & 1 \end{array} \right] \end{array}, \; \boldsymbol{B}_T = \left[\begin{array}{c} \boldsymbol{B}_{t_0} \\ \boldsymbol{B}_{t_1} \end{array} \right]. \tag{28}$$

Under monotonicity condition **A-1** the three response-type probabilities $(P(\boldsymbol{s}_1), P(\boldsymbol{s}_2), P(\boldsymbol{s}_4))$ and the four counterfactual outcomes $(E(Y(t_0)|\boldsymbol{S} = \boldsymbol{s}_2), E(Y(t_1)|\boldsymbol{S} = \boldsymbol{s}_2), E(Y(t_0)|\boldsymbol{S} = \boldsymbol{s}_1), E(Y(t_0)|\boldsymbol{S} = \boldsymbol{s}_4))$ are identified. These results can be demonstrated by applying Theorem **T-2** and Corollary **C-1**. For instance the rank of the binary matrix $\boldsymbol{B}_T$ in (28) is 3, which is also the number of response-types. Thus, by Corollary **C-1**, all the response probabilities $\boldsymbol{P}_S$ are identified. The identification of counterfactual outcomes depends on the properties of matrices $\boldsymbol{K}_{t_0}$, $\boldsymbol{K}_{t_1}$ that are calculated using the pseudo-inverse matrices $\boldsymbol{B}^+_{t_0}, \boldsymbol{B}^+_{t_1}$ as described in (21):

$$\boldsymbol{B}^+_{t_0} = \left[\begin{array}{cc} 0 & 0 \\ 1 & -1 \\ 0 & 1 \end{array} \right] \Rightarrow \boldsymbol{K}_{t_0} = \left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array} \right] \text{and } \boldsymbol{B}^+_{t_1} = \left[\begin{array}{cc} 1 & 0 \\ -1 & 1 \\ 0 & 0 \end{array} \right] \Rightarrow \boldsymbol{K}_{t_1} = \left[\begin{array}{ccc} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{array} \right].$$

The observed vectors of propensity scores and conditional outcome expectations are $\boldsymbol{P}_Z = [P(T = t|Z = z_0), P(T = t|Z = z_1)]'$ and $\boldsymbol{Q}_Z(t) = [E(Y \cdot \mathbf{1}[T = t]|Z = z_0), E(Y \cdot \mathbf{1}[T = t]|Z = z_1)]'$, for $t \in \{t_1, t_0\}$. The unobserved 3×1 vectors of response-type probabilities and counterfactual outcomes are given by $\boldsymbol{P}_S = [P(\boldsymbol{S} =$

[28] Imbens and Angrist (1994) do not use indicator functions. This is an innovation of this paper. They compare the values of the counterfactual choices directly, e.g., $T_\omega(z) \geq T_\omega(z')$, assuming the T are ordered. In their analysis, the values that choice T takes must be ordered. Our approach does not require T to be ordered. The two monotonicity criteria are equivalent for the binary choice model.

$\boldsymbol{s}_1), P(\boldsymbol{S} = \boldsymbol{s}_2), P(\boldsymbol{S} = \boldsymbol{s}_4)]'$ and $\boldsymbol{Q}_S(t) = [E(Y(t)|\boldsymbol{S} = \boldsymbol{s}_1)P(\boldsymbol{S} = \boldsymbol{s}_1), E(Y(t)|\boldsymbol{S} = \boldsymbol{s}_2)P(\boldsymbol{S} = \boldsymbol{s}_2), E(Y(t)|\boldsymbol{S} = \boldsymbol{s}_4)P(\boldsymbol{S} = \boldsymbol{s}_4)]'$. Applying Equations (22) and (23), $E(Y(t_0)|\boldsymbol{S} = \boldsymbol{s}_2) = \frac{\boldsymbol{\zeta}'\boldsymbol{Q}_S(t_0)}{\boldsymbol{\zeta}'\boldsymbol{P}_S}$ where $\boldsymbol{\zeta} = [0,1,0]'$, so that $\boldsymbol{\zeta}'\boldsymbol{P}_S = P(\boldsymbol{S} = \boldsymbol{s}_2)$ is the population probability of the compliers. Note that $\boldsymbol{\zeta}'\boldsymbol{K}_{t_0} = \boldsymbol{0}$, thus, by Theorem **T-2**, $E(Y(t_0)|\boldsymbol{S} = \boldsymbol{s}_2)$ is identified. From Equations (22) and (23), we have:

$$E(Y(t_0)|\boldsymbol{S} = \boldsymbol{s}_2) = \frac{\boldsymbol{\zeta}'\boldsymbol{B}_{t_0}^{+}\boldsymbol{Q}_Z(t_0)}{\boldsymbol{\zeta}'\boldsymbol{B}_{t_0}^{+}\boldsymbol{P}_Z(t_0)} = \frac{E(Y\cdot\mathbf{1}[T=t_0]|Z=z_0) - E(Y\cdot\mathbf{1}[T=t_0]|Z=z_1)}{P(T=t_0|Z=z_0) - P(T=t_0|Z=z_1)}.$$

By a parallel argument, the counterfactual outcome $E(Y(t_1)|\boldsymbol{S} = \boldsymbol{s}_2) = \frac{\boldsymbol{\zeta}'\boldsymbol{Q}_S(t_1)}{\boldsymbol{\zeta}'\boldsymbol{P}_S}$. Since $\boldsymbol{\zeta}'\boldsymbol{K}_{t_1} = \boldsymbol{0}$, by Theorem **T-2**, $E(Y(t_1)|\boldsymbol{S} = \boldsymbol{s}_2)$ is identified from the expression

$$E(Y(t_1)|\boldsymbol{S} = \boldsymbol{s}_2) = \frac{\boldsymbol{\zeta}'\boldsymbol{B}_{t_1}^{+}\boldsymbol{Q}_Z(t_1)}{\boldsymbol{\zeta}'\boldsymbol{B}_{t_1}^{+}\boldsymbol{P}_Z(t_1)} = \frac{E(Y\cdot\mathbf{1}[T=t_1]|Z=z_1) - E(Y\cdot\mathbf{1}[T=t_1]|Z=z_0)}{P(T=t_1|Z=z_1) - P(T=t_1|Z=z_0)}.$$

LATE is the causal effect for compliers $E(Y(t_1) - Y(t_0)|\boldsymbol{S} = \boldsymbol{s}_2)$. Since $P(T = t_0|Z = z) = 1 - P(T = t_1|Z = z)$, $\boldsymbol{\zeta}'\boldsymbol{B}_{t_1}^{+}\boldsymbol{P}_Z(t_1) = \boldsymbol{\zeta}'\boldsymbol{B}_{t_0}^{+}\boldsymbol{P}_Z(t_0) = P(\boldsymbol{S} = \boldsymbol{s}_2)$. Putting these ingredients together,

$$\begin{aligned} &E(Y(t_0) - Y(t_0)|\boldsymbol{S} = \boldsymbol{s}_2) \\ =&\frac{\boldsymbol{\zeta}'\left(\boldsymbol{B}_{t_1}^{+}\boldsymbol{Q}_Z(t_1) - \boldsymbol{B}_{t_0}^{+}\boldsymbol{Q}_Z(t_0)\right)}{\boldsymbol{\zeta}'\boldsymbol{B}_{t_1}^{+}\boldsymbol{P}_Z(t_1)} = \frac{E(Y|Z=z_1) - E(Y|Z=z_0)}{P(T=t_1|Z=z_1) - P(T=t_1|Z=z_0)}. \end{aligned}$$

LATE is the causal effect conditioned on the values of $\boldsymbol{V}$ associated with strata $\boldsymbol{s}_2$. It does not identify the average treatment effect $E(Y(t_1) - Y(t_0))$ because we cannot identify $Y(t_1)$ for $\boldsymbol{s}_4$ (t_0-always-taker) nor $Y(t_0)$ for $\boldsymbol{s}_1$ (t_1-always-taker). The counterfactual outcomes for the always-takers can be expressed in terms of $\boldsymbol{Q}_S(t)$ and $\boldsymbol{P}_S$ by:

$$E(Y(t_0)|\boldsymbol{S} = \boldsymbol{s}_4) = \frac{\boldsymbol{\zeta}_0'\boldsymbol{Q}_S(t_0)}{\boldsymbol{\zeta}_0'\boldsymbol{P}_S};\ \boldsymbol{\zeta}_0 = [0,0,1]' \text{ and } E(Y(t_1)|\boldsymbol{S} = \boldsymbol{s}_1) = \frac{\boldsymbol{\zeta}_1'\boldsymbol{Q}_S(t_1)}{\boldsymbol{\zeta}_1'\boldsymbol{P}_S};\ \boldsymbol{\zeta}_1 = [1,0,0]'.$$

Since $\boldsymbol{\zeta}_0'\boldsymbol{K}_{t_0} = \boldsymbol{0}$ and $\boldsymbol{\zeta}_1'\boldsymbol{K}_{t_1} = \boldsymbol{0}$, by Theorem **T-2**, $E(Y(t_0)|\boldsymbol{S} = \boldsymbol{s}_4)$ and $E(Y(t_1)|\boldsymbol{S} = \boldsymbol{s}_1)$ are identified. In Section 7, we use the properties of the generalized inverse to extend our analysis to a general model of multiple choices.

4.2 Revisiting Vytlacil's Equivalence Theorem

A by-product of our analysis is a simple derivation of Vytlacil's (2002) fundamental equivalence result. He shows that monotonicity condition **A-1** holds if and only if the treatment choice can be expressed as a function that is separable in Z and $\boldsymbol{V}$, i.e., there exist deterministic functions, $\varphi : \text{supp}(\boldsymbol{V}) \to \mathbb{R}$ and $\tau : \text{supp}(Z) \to \mathbb{R}$ such that:

$$\Big(\mathbf{1}[T = t_1] | \boldsymbol{V} = \boldsymbol{v}, Z = z\Big) = \mathbf{1}[\tau(z) \geq \varphi(\boldsymbol{v})]. \tag{29}$$

Monotonicity **A-1** generates a key property of the binary matrix $\boldsymbol{B}_{t_1} = \mathbf{1}[\boldsymbol{R} = t_1]$. We can always reorder its rows and columns so that $\boldsymbol{B}_{t_1}$ becomes a lower-triangular matrix.[29] Consider the binary choice model where T takes values in $\{t_0, t_1\}$ and Z takes values in $\{z_1, \ldots, z_{N_Z}\}$ that are indexed by increasing values of the propensity score, i.e., $P(T = t_1|z_1) \leq \cdots \leq P(T = t_1|z_{N_Z})$. Arrange the columns of binary matrix $\boldsymbol{B}_{t_1}$ in decreasing order of the column-sums. Under monotonicity (**A-1**), $\boldsymbol{B}_{t_1}$ has dimension $N_Z \times (N_Z + 1)$ and is lower triangular. An explicit expression for $\boldsymbol{B}_{t_1}$ is given by Equation (28) for $N_Z = 2$.[30] Under triangularity, for all $i \in \{1, \cdots, N_Z\}$, $n \in \{1, \cdots, N_Z + 1\}$,

$$\boldsymbol{B}_{t_1}[i, n] = 1 \text{ for } i \geq n \quad \text{and} \quad \boldsymbol{B}_{t_1}[i, n] = 0 \text{ for } i < n. \tag{30}$$

Propensity score equality (15) generates the following expressions:

$$\begin{aligned} P(T = t_1 | Z = z_i) &= \sum_{n'=1}^{N_S} \mathbf{1}[T = t_1 | Z = z_i, \boldsymbol{S} = \boldsymbol{s}_{n'}] \cdot P(\boldsymbol{S} = \boldsymbol{s}_{n'}) \\ &= \sum_{n'=1}^{N_Z+1} \boldsymbol{B}_{t_1}[i, n'] \cdot P(\boldsymbol{S} = \boldsymbol{s}_{n'}) \\ &= \sum_{n'=1}^{i} P(\boldsymbol{S} = \boldsymbol{s}_{n'}). \end{aligned} \tag{31}$$

The second equality uses the definition of an element in the i-th row and n-th column of $\boldsymbol{B}_{t_1}[i, n']$, that is $\boldsymbol{B}_{t_1}[i, n'] = \mathbf{1}[T = t_1 | Z = z_i, \boldsymbol{S} = \boldsymbol{s}_{n'}]$ and because $N_S = N_Z + 1$

[29] Recall $\boldsymbol{R}$ does not have redundant rows or columns.

[30] In Section 6, we present a generalization of the triangular property for matrices called "lonesum matrices."

due to monotonicity (A-1). The third equality uses triangularity property (30). Thus the following inequalities hold:

$$\text{Since } P(T=t_1|Z=z_i) = \sum_{n'=1}^{i} P(\boldsymbol{S}=\boldsymbol{s}_{n'}),$$

$$\text{then } P(T=t_1|Z=z_i) \geq \sum_{n'=1}^{n} P(\boldsymbol{S}=\boldsymbol{s}_{n'}) \qquad \text{for } i \geq n \tag{32}$$

$$\text{and } \quad P(T=t_1|Z=z_i) < \sum_{n'=1}^{n} P(\boldsymbol{S}=\boldsymbol{s}_{n'}) \qquad \text{for } i < n. \tag{33}$$

We can combine Equations (30) and (32)–(33) to express the elements $\boldsymbol{B}_{t_1}[i,n]$ as:

$$\boldsymbol{B}_{t_1}[i,n] = \mathbf{1}\left[P(T=t_1|Z=z_i) \geq \phi(\boldsymbol{s}_n)\right], \tag{34}$$

$$\text{where } \phi(\boldsymbol{s}_n) = P(S \in \{\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n\}) = \sum_{n'=1}^{n} P(\boldsymbol{S}=\boldsymbol{s}_{n'}). \tag{35}$$

Vytlacil's theorem emerges since $\boldsymbol{B}_{t_1}[i,n] = \mathbf{1}[T=t_1|Z=z_i,\boldsymbol{S}=\boldsymbol{s}_n]$ and $\boldsymbol{S}$ is a balancing score for $\boldsymbol{V}$, i.e., $\boldsymbol{S}=f_S(\boldsymbol{V})$. Thus, for any $\boldsymbol{v} \in \text{supp}(\boldsymbol{V})$ there is an $\boldsymbol{s} \in \text{supp}(\boldsymbol{S})$ such that $\boldsymbol{s}=f_S(\boldsymbol{v})$, and

$$\mathbf{1}\Big[T=t_1|Z=z,\boldsymbol{V}=\boldsymbol{v}\Big]=\mathbf{1}\Big[T=t_1|Z=z,\boldsymbol{S}=f_S(\boldsymbol{v})\Big] = \mathbf{1}\Big[\underbrace{P(T=t_1|Z=z)}_{\tau(z)} \geq \underbrace{\phi(f_S(\boldsymbol{v}))}_{\varphi(\boldsymbol{v})}\Big]. \tag{36}$$

This expression captures the key idea that the response variable $\boldsymbol{S}$ summarizes $\boldsymbol{V}$. Section 6 generalizes Vytlacil's analysis to the case of a general unordered model. The triangularity property generating separability carries over to that general setting.

5 Multiple Unordered Choices

In the published literature, when LATE is extended to analyze multiple choices, T is assumed to be a scalar index defined over an ordered finite set of natural numbers $\{1,\ldots,N_T\}$ where the index is monotonically increasing (or decreasing) in the indicators of t (Angrist and Imbens, 1995). Treatment effects are defined in terms of variations in this index:

Assumption A-2. ***Ordered Monotonicity.*** *The following inequalities hold for any* $z, z' \in \text{supp}(Z)$*, and each treatment* $t \in \text{supp}(T)$*:*

$$T_\omega(z) \geq T_\omega(z') \, \forall \, \omega \in \Omega \quad \text{or} \quad T_\omega(z) \leq T_\omega(z') \, \forall \, \omega \in \Omega. \tag{37}$$

Under standard assumptions about IV, Assumption **A-2** is equivalent to the assumption that choices are generated by an ordered choice model (Vytlacil, 2004). To extend monotonicity to the unordered case, we retain the core feature of a monotonic relationship: shifts in Z move all agents toward or against making treatment choice t in $\text{supp}(T)$. We do not require any order among the values of T, nor do we rely on a scalar representation of T. Instead, we replace comparisons of T with inequalities that compare *indicator functions* of the values taken by T for each pair of values z, z' in $\text{supp}(Z)$. If the support of T has no natural order, Assumption **A-2** is meaningless.

This section extends the literature to define a concept of monotonicity for an *unordered* choice model. We discuss restrictions on the response matrix $\boldsymbol{R}$ that follow from this definition. We present some examples that build intuition.

5.1 Monotonicity for Unordered Models

Assumption A-3. ***Unordered Monotonicity.*** *The following inequalities hold for any* $z, z' \in \text{supp}(Z)$*, and each treatment* $t \in \text{supp}(T)$*:*

$$\mathbf{1}[T_\omega(z) = t] \geq \mathbf{1}[T_\omega(z') = t] \, \forall \, \omega \in \Omega \quad \text{or} \quad \mathbf{1}[T_\omega(z) = t] \leq \mathbf{1}[T_\omega(z') = t] \, \forall \, \omega \in \Omega, \tag{38}$$

where $\mathbf{1}[T_\omega(z) = t]$ *indicates whether or not agent* ω *chooses treatment* $t \in \text{supp}(T)$ *when* Z *is set to* z*.*

Using indicator functions, we can make pairwise comparisons for all values of Z for each choice $t \in \text{supp}(T)$ without imposing an arbitrary ordering on the values of the treatment choices T or creating a scalar index of T. Condition (38) preserves the key intuitive notion of monotonicity: a shift in an instrument moves all agents uniformly toward or against each possible choice. Assumption **A-3** prohibits non-uniform movements induced by the instruments and is ruled out in Theorem **T-3** below.

In the case of binary treatment, Ordered Monotonicity **A-2** and Unordered Monotonicity **A-3** generate the same monotonicity restriction **A-1**.[31] In Appendix C, we present a simple example that demonstrates the benefits of using choice indicators rather than cardinal measures of outcomes to define monotonicity.

5.2 Linking Unordered Monotonicity to Choice Theory

Under Unordered Monotonicity, treatment choice can be characterized as the solution to a problem in which agents maximize utility $\Psi(t, z, \boldsymbol{v})$, the utility arising from choosing $t \in \text{supp}(T)$ for agent ω whose unobserved variable $\boldsymbol{V}$ takes value $\boldsymbol{v}$ when the instrument Z is set at z. We present a formal analysis of the properties of $\Psi(t, z, \boldsymbol{v})$ generated by Unordered Monotonicity in Section 6. In this section we build economic intuition of how Unordered Monotonicity arises. We use revealed preference arguments to restrict $\boldsymbol{R}$ and generate monotonicity conditions. We give examples where plausible restrictions on choice theory, coupled with standard instrumental variable conditions, produce identification of various strata counterfactuals and response-type probabilities. We also examine cases in which the point identification of response-type probabilities fails.

Consider an example in which each agent buys a single car from three possible options: $\{a, b, c\}$. Let $T_\omega = t_j$ if agent ω buys car j, $\text{supp}(T) = \{t_a, t_b, t_c\}$. Instruments are randomly assigned car-specific vouchers that offer price discounts to the car (or cars) specified by an offered voucher. We use z_a, z_b, z_c for vouchers that offer a discount to cars a, b, and c respectively. We use z_{bc} for the voucher whose discount can be used to buy car b or c. z_{no} denotes no discount. If the voucher assigned to agent ω is z_a, he faces a price-discount if he decides to buy car a. Instead, if agent ω decides to buy car b or c, the agent will pay full price. If the agent were assigned voucher z_{bc} then the cars b and c become cheaper and car a is full price. We compare experimental designs that randomly assign different combinations of 3 out of the 5 voucher-types described above. Each agent ω is assumed to buy some car. In this section and in Web Appendix D, we give some examples of how choice restrictions facilitate identification and where they fail.

Our main example carried throughout the rest of this paper considers vouchers in $\text{supp}(Z) = \{z_{no}, z_a, z_{bc}\}$. The response vector $\boldsymbol{S}$ is given by the 3-dimensional vector of counterfactual choices: $\boldsymbol{S} = [T(z_{no}), T(z_a), T(z_{bc})]'$. Each of the three counterfac-

[31] Heckman and Pinto (2015c) show that in the general case of multivalued treatments, Ordered Monotonicity **A-2** and Unordered Monotonicity **A-3** do not imply each other.

tual choices $T(z), z \in \{z_{no}, z_a, z_{bc}\}$ takes values in $\{t_a, t_b, t_c\}$, which gives a total of 27 ($= 3^3$) possible response-types.[32] Without restrictions on admissible strata, the model of strata-contingent counterfactuals is not identified.[33] There are four intuitive monotonicity relationships arising from changes in z:

$$\mathbf{1}[T_\omega(z_{no}) = t_a] \leq \mathbf{1}[T_\omega(z_a) = t_a], \tag{39}$$

$$\mathbf{1}[T_\omega(z_{bc}) = t_a] \leq \mathbf{1}[T_\omega(z_a) = t_a], \tag{40}$$

$$\mathbf{1}[T_\omega(z_{no}) \in \{t_b, t_c\}] \leq \mathbf{1}[T_\omega(z_{bc}) \in \{t_b, t_c\}], \tag{41}$$

$$\mathbf{1}[T_\omega(z_a) \in \{t_b, t_c\}] \leq \mathbf{1}[T_\omega(z_{bc}) \in \{t_b, t_c\}]. \tag{42}$$

Relationship (39) states that the agent is induced toward buying car a when the instrument changes from no voucher (z_{no}) to a voucher for car a (z_a). Relationship (40) states that the agent is induced toward buying car a when the instrument changes from a voucher to buy b or c (z_{bc}) to a voucher for car a (z_a). Relationship (41) states that the agent is induced toward buying either car b or c when the instrument changes from no voucher (z_{no}) to a voucher for either car b or c (z_{bc}). Relationship (42) states that the agent is induced toward buying either car b or c when the instrument changes from a voucher for car a (z_a) to a voucher that applies to either car b or c (z_{bc}). Monotonicity relationships (39)–(42) eliminate 12 response-types out of the 27 possible ones, leaving the 15 admissible response-types presented in Table 1.[34]

Table 1: Response Matrix Generated by Monotonicity Relationships (39)–(42)

Instrumental Variables	Choices	s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9	s_{10}	s_{11}	s_{12}	s_{13}	s_{14}	s_{15}
		Response-types of S														
No Voucher	$T(z_{no})$	t_a	t_a	t_a	t_b	t_b	t_b	t_b	t_b	t_b	t_c	t_c	t_c	t_c	t_c	t_c
Voucher for a	$T(z_a)$	t_a	t_a	t_a	t_a	t_a	t_b	t_b	t_c	t_c	t_a	t_a	t_b	t_b	t_c	t_c
Voucher for b or c	$T(z_{bc})$	t_a	t_b	t_c	t_b	t_c	t_b	t_c	t_b	t_c	t_b	t_c	t_b	t_c	t_b	t_c

Thus, by Corollary **C-1**, our model for counterfactuals is not identified. In addition, some of the remaining strata are not consistent with Unordered Monotonicity **A-**

[32] See Web Appendix Table D.1.

[33] If we assume an ordered choice model, we can readily secure identification. If we only assume a partially ordered model we lose identification. Heckman and Pinto (2015c) discuss these cases.

[34] See the elimination analysis in Table D.1 of Web Appendix D.

3. More stringent application of revealed preference analysis can generate additional choice restrictions. Let $\Lambda_\omega(z,t)$ be the consumption set of agent ω when assigned instrument $z \in \text{supp}(Z)$ when treatment is set to $t \in \text{supp}(T)$. Let $\gamma \in \Lambda_\omega(z,t)$ represent a consumption good. Agent ω is assumed to maximize a utility function u_ω defined over consumption goods γ and choice t. Thus, the choice function $Ch_\omega : \text{supp}(Z) \to \text{supp}(T)$ of agent ω when the instrument is set to value $z \in \text{supp}(Z)$ is:

$$Ch_\omega(z) = \underset{t \in \text{supp}(T)}{\text{argmax}} \Big(\max_{g \in \Lambda_\omega(z,t)} u_\omega(g,t) \Big). \tag{43}$$

For budget set $\Lambda_\omega(z,t)$ for agent ω, we assume the following relationships:

$$\Lambda_\omega(z_{no},t_a) = \Lambda_\omega(z_{bc},t_a) \subset \Lambda_\omega(z_a,t_a), \tag{44}$$

$$\Lambda_\omega(z_{no},t_b) = \Lambda_\omega(z_a,t_b) \subset \Lambda_\omega(z_{bc},t_b), \tag{45}$$

$$\Lambda_\omega(z_{no},t_c) = \Lambda_\omega(z_a,t_c) \subset \Lambda_\omega(z_{bc},t_c). \tag{46}$$

Relationship (44) compares the budget sets of agent ω for each possible voucher assignment given the car choice is fixed at a. The budget set of agent ω is enlarged when she has a voucher for car a (z_a) compared to when she does not (z_a is the only voucher that applies to car a). Thus, assigning consumer ω who buys car a, voucher z_a provides additional income. Vouchers z_{no} and z_{bc} offer no discount for car a and produce the same budget set for this choice. Relationship (45) examines the agent's budget set if ω purchases car b. The budget set of agent ω is enlarged if she has a voucher that subsidizes car b when compared to vouchers that do not affect the choice set (z_a, z_{no}). Relationship (46) examines the agent's budget set when car c is assigned and is consistent with the budget analysis of relationship (45).[35] For this example, the Weak Axiom of Revealed Preference (WARP) generates the following choice rule:

$$\text{if } Ch_\omega(z) = t \text{ and } \Lambda_\omega(z,t) \subseteq \Lambda_\omega(z',t) \text{ and } \Lambda_\omega(z',t') \subseteq \Lambda_\omega(z,t') \Rightarrow Ch_\omega(z') \neq t'.\text{[36]} \tag{47}$$

In particular, choice rule (47) applied to budget set relationships (44)–(46) generates the choice restrictions 1–6 in Table 2.

Under additional assumptions about choice, we generate additional restrictions

[35] Under this rationale, it follows that: $\Lambda_\omega(z_b,t_a) = \Lambda_\omega(z_{no},t_a)$, $\Lambda_\omega(z_b,t_b) = \Lambda_\omega(z_{bc},t_b)$, and $\Lambda_\omega(z_b,t_c) = \Lambda_\omega(z_{no},t_c)$.

[36] See Pinto (2016a).

Table 2: Choice Restrictions Generated by Revealed Preference Analysis for $\text{supp}(Z) = \{z_{no}, z_a, z_{bc}\}$

Choice Restriction **1** :	$Ch_\omega(z_{no}) = t_a$	$\Rightarrow$	$Ch_\omega(z_a) = t_a$		
Choice Restriction **2** :	$Ch_\omega(z_{no}) = t_b$	$\Rightarrow$	$Ch_\omega(z_a) \neq t_c$	and	$Ch_\omega(z_{bc}) \neq t_a$
Choice Restriction **3** :	$Ch_\omega(z_{no}) = t_c$	$\Rightarrow$	$Ch_\omega(z_a) \neq t_b$	and	$Ch_\omega(z_{bc}) \neq t_a$
Choice Restriction **4** :	$Ch_\omega(z_a) = t_b$	$\Rightarrow$	$Ch_\omega(z_{no}) = t_b$	and	$Ch_\omega(z_{bc}) \neq t_a$
Choice Restriction **5** :	$Ch_\omega(z_a) = t_c$	$\Rightarrow$	$Ch_\omega(z_{no}) = t_c$	and	$Ch_\omega(z_{bc}) \neq t_a$
Choice Restriction **6** :	$Ch_\omega(z_{bc}) = t_a$	$\Rightarrow$	$Ch_\omega(z_{no}) = t_a$	and	$Ch_\omega(z_a) = t_a$
Choice Restriction **7** :	$Ch_\omega(z_{no}) \neq t_a$	$\Rightarrow$	$Ch_\omega(z_{bc}) = Ch_\omega(z_{no})$		

on the admissible strata. It is reasonable to assume that if an agent decides to buy a car without a discount, then the agent will not alter his choice if assigned a voucher that makes his choice of car cheaper. Specifically consider the agent who decides between cars b and c when voucher assignment shifts from z_{no} to z_{bc}. There is no discount under z_{no} whereas z_{bc} offers a discount for either car. If most of the income increase is spent on goods, then the agent's car choice remains the same.[37] Under this condition, an income increase should not decrease its consumption of a good. If the agent is already consuming one unit of car b and his income is increased, then the agent will not decrease his car consumption, hence the agent still buys car b if the voucher changes from z_{no} to z_{bc}.[38] This restriction on choice generates the 7 admissible response types in Table 2. The choice restrictions of Table 2 eliminate 20 out of the 27 possible response-types generating the admissible response matrix in Table 3.[39]

Table 3: Response-Types Generated by Revealed Preference Analysis for $\text{supp}(Z) = \{z_{no}, z_a, z_{bc}\}$

Instrumental Variables	Choices	$\boldsymbol{s}_1$	$\boldsymbol{s}_2$	$\boldsymbol{s}_3$	$\boldsymbol{s}_4$	$\boldsymbol{s}_5$	$\boldsymbol{s}_6$	$\boldsymbol{s}_7$
		Response-types of S						
No Voucher	$T(z_{no})$	t_a	t_a	t_a	t_b	t_b	t_c	t_c
Voucher for car a	$T(z_a)$	t_a	t_a	t_a	t_a	t_b	t_a	t_c
Voucher for car b or c	$T(z_{bc})$	t_a	t_b	t_c	t_b	t_b	t_c	t_c

[37] This would occur if utility is quasilinear in γ.
[38] A stronger assumption is homothetic preferences on consumption goods.
[39] See the elimination analysis in Table D.5 of Web Appendix D.

For the response matrix of Table 3, the rank of the indicator matrix $\boldsymbol{B}_T$ associated with this response matrix is equal to 7 which is also equal to the number of response-types. From Corollary **C-1**, response-type probabilities are identified. We can also identify mean counterfactual outcomes defined in terms of the strata in the table. The response matrix of Table 3 is generated by the nine Unordered Monotonicity relationships of Table 4.[40] The choice restrictions generated by the revealed preference analysis in Table 2 produce Unordered Monotonicity **A-3**.

Remark 5.1. The response matrix in Table 3 is ***uniquely*** *generated by the Unordered Monotonicity relationships of Table 4. By uniquely we mean that a change in the direction of any of these inequalities produces a response matrix that differs from the one in Table 3. This property is useful for testing the model assumptions as each monotonicity relationship implies a propensity score inequality that can be tested on observed data.*

Table 4: An Identified Pattern of Response Matrices

	Monotonicity Relationships			Implied Propensity Score Inequalities		
Relation 1	$\mathbf{1}[T_\omega(z_{no}) = t_a]$	$\leq$	$\mathbf{1}[T_\omega(z_a) = t_a]$	$P(T = t_a \vert Z = z_{no})$	$\leq$	$P(T = t_a \vert Z = z_a)$
Relation 2	$\mathbf{1}[T_\omega(z_{no}) = t_a]$	$\geq$	$\mathbf{1}[T_\omega(z_{bc}) = t_a]$	$P(T = t_a \vert Z = z_{no})$	$\geq$	$P(T = t_a \vert Z = z_{bc})$
Relation 3	$\mathbf{1}[T_\omega(z_a) = t_a]$	$\geq$	$\mathbf{1}[T_\omega(z_{bc}) = t_a]$	$P(T = t_a \vert Z = z_a)$	$\geq$	$P(T = t_a \vert Z = z_{bc})$
Relation 4	$\mathbf{1}[T_\omega(z_{no}) = t_b]$	$\geq$	$\mathbf{1}[T_\omega(z_a) = t_b]$	$P(T = t_b \vert Z = z_{no})$	$\geq$	$P(T = t_a \vert Z = z_a)$
Relation 5	$\mathbf{1}[T_\omega(z_{no}) = t_b]$	$\leq$	$\mathbf{1}[T_\omega(z_{bc}) = t_b]$	$P(T = t_b \vert Z = z_{no})$	$\leq$	$P(T = t_a \vert Z = z_{bc})$
Relation 6	$\mathbf{1}[T_\omega(z_a) = t_b]$	$\leq$	$\mathbf{1}[T_\omega(z_{bc}) = t_b]$	$P(T = t_b \vert Z = z_a)$	$\leq$	$P(T = t_a \vert Z = z_{bc})$
Relation 7	$\mathbf{1}[T_\omega(z_{no}) = t_c]$	$\geq$	$\mathbf{1}[T_\omega(z_a) = t_c]$	$P(T = t_c \vert Z = z_{no})$	$\geq$	$P(T = t_c \vert Z = z_a)$
Relation 8	$\mathbf{1}[T_\omega(z_{no}) = t_c]$	$\leq$	$\mathbf{1}[T_\omega(z_{bc}) = t_c]$	$P(T = t_c \vert Z = z_{no})$	$\leq$	$P(T = t_c \vert Z = z_{bc})$
Relation 9	$\mathbf{1}[T_\omega(z_a) = t_c]$	$\leq$	$\mathbf{1}[T_\omega(z_{bc}) = t_c]$	$P(T = t_c \vert Z = z_a)$	$\leq$	$P(T = t_c \vert Z = z_{bc})$

Unordered Monotonicity can arise under different configurations of the instrumental variable. Thus, in the previous example, consider changing the support of the instrumental variable Z from $\{z_{no}, z_a, z_{bc}\}$ to $\{z_{no}, z_b, z_{bc}\}$. We can apply the same revealed preference analysis of the first example to $\{z_{no}, z_b, z_{bc}\}$. This analysis generates the response matrix shown in Table 5 which is also uniquely generated by nine inequalities consistent with Unordered Monotonicity **A-3**. The response matrix

[40] See the elimination analysis in Table D.6 of Web Appendix D.

also identifies response-type probabilities and an associated set of counterfactual outcomes. However, three out of seven response-types in Table 5 differ from the ones in Table 3.

Table 5: Response-Types Generated by Revealed Preference Analysis for $\text{supp}(Z) = \{z_{no}, z_b, z_{bc}\}$.

Instrumental Variables	Choices	$\boldsymbol{s}_1$	$\boldsymbol{s}_2$	$\boldsymbol{s}_3$	$\boldsymbol{s}_4$	$\boldsymbol{s}_5$	$\boldsymbol{s}_6$	$\boldsymbol{s}_7$
		Response-types of S						
No Voucher	$T(z_{no})$	t_a	t_a	t_a	t_a	t_b	t_c	t_c
Voucher for car b	$T(z_b)$	t_a	t_a	t_b	t_b	t_b	t_c	t_b
Voucher for car b or c	$T(z_{bc})$	t_a	t_c	t_b	t_c	t_b	t_c	t_c

Choice restrictions alone do not necessarily produce identifiability. For an example, see Web Appendix D.2. We further note that Unordered Monotonicity **A-3** is not a necessary condition for identification of model parameters. In Web Appendix D.3, we modify the example of Table 5 by assuming that Z takes values in $\text{supp}(Z) = \{z_c, z_b, z_{bc}\}$. WARP alone generates the response matrix described in Table 6.[41] The rank of its associated binary matrix $\boldsymbol{B}_T$ is equal to 7. Thus, response-type probabilities are identified. However, the response matrix in Table 6 is not consistent with Unordered Monotonicity **A-3**. There is no sequence of monotonic relationships consistent with **A-3** that generates this response matrix. For example, consider the change in voucher assignment from voucher for c (z_c) to voucher for b (z_b) in Table 6. This change induces those in $\boldsymbol{s}_4$ to move towards t_a (from t_c to t_a), while those in $\boldsymbol{s}_2$ to move away from t_a (from t_a to t_b). This pattern of counterfactual choices is inconsistent with monotonicity.[42] Unordered Monotonicity in the general choice model coupled with standard IV conditions, does not necessarily guarantee identifiability, unlike its counterpart in the binary choice model and in the ordered choice model.[43] Moreover, revealed preference analysis may or may not identify the choice model, depending on the patterns of restrictions imposed on the variation in

[41]See Table D.13 in Web Appendix D.3 for the elimination analysis.
[42]This claim is formally proved in the next section using condition (iii) of Theorem **T-3**.
[43]See the discussion in Web Appendix D.2.

the instruments.[44]

Table 6: Response-Types Generated by Revealed Preference Analysis for $\text{supp}(Z) = \{z_c, z_b, z_{bc}\}$.

Instrumental Variables	Count. Choices	Response-types of $\boldsymbol{S}$ $\boldsymbol{s}_1$	$\boldsymbol{s}_2$	$\boldsymbol{s}_3$	$\boldsymbol{s}_4$	$\boldsymbol{s}_5$	$\boldsymbol{s}_6$	$\boldsymbol{s}_7$
Voucher for c	$T(z_c)$	t_a	t_a	t_b	t_c	t_c	t_c	t_c
Voucher for b	$T(z_b)$	t_a	t_b	t_b	t_a	t_b	t_b	t_c
Voucher for b or c	$T(z_{bc})$	t_a	t_b	t_b	t_c	t_b	t_c	t_c

6 Equivalent Conditions for Characterizing Unordered Monotonicity

This section presents and interprets general properties shared by all response matrices that satisfy Unordered Monotonicity **A-3**. We explore a variety of ways to express **A-3** including separability of choice equations.

6.1 Properties of Binary Matrices

To establish a relationship between identifiability and the properties of response matrix $\boldsymbol{R}$, it is helpful to use concepts from the literature on binary matrices. A binary matrix is *lonesum* if it is uniquely determined by its row and column sums.[45] We establish that response matrix $\boldsymbol{R}$ is an *unordered monotone response matrix* (henceforth "monotone") if each binary matrix derived from it, $\boldsymbol{B}_t = \mathbf{1}[\boldsymbol{R} = t]; t \in \text{supp}(T)$, is *lonesum*. Lonesum matrices can be used to characterize monotonicity conditions in choice models. We show that identification and equivalence results arise from the properties of lonesum matrices.

[44]This paper does not consider issues of estimation and inference. If certain parameters are overidentified from different instrument configurations, the obvious approach is to combine estimators using efficient GMM (Hansen, 1982).

[45]See Ryser (1957), Brualdi (1980), Brualdi and Ryser (1991), and Sachnov and Tarakanov (2002) for surveys of the properties of binary matrices.

Let $r_{i,t}$ be the i-th row sum of the binary matrix $\boldsymbol{B}_t$: $r_{i,t} = \sum_{n=1}^{N_S} \boldsymbol{B}_t[i,n]$. Let $c_{n,t}$ denote the sum of the n-th column of $\boldsymbol{B}_t$, that is, $c_{n,t} = \sum_{i=1}^{N_Z} \boldsymbol{B}_t[i,n]$. The *maximal* of matrix $\boldsymbol{B}_t$ is a matrix whose i-th row is given by $r_{i,t}$ elements 1 followed by 0s. Two matrices are *equivalent* if one can be transformed into the other by a series of row and/or column permutations.

Table 7 displays matrix $\boldsymbol{B}_{t_a} = \mathbf{1}[\boldsymbol{R} = t_a]$, where $\boldsymbol{R}$ is the response matrix of Table 3. The first column of Table 7 gives the row sums of $\boldsymbol{B}_{t_a}$. The last row of Table 7 presents its column sums. To show that matrix $\boldsymbol{B}_{t_a}$ is lonesum, reorder its columns and rows based on decreasing values of column sums and increasing values of row sums. The maximal of $\boldsymbol{B}_{t_a}$ is obtained by a reordering of B_{t_a} based only on row and column sums. Note that there are different orderings for different t. The reordered matrix of Table 3 is given in Table 8. It is a maximal matrix because the matrix rows are described by elements 1 followed by 0s. For example, if a maximal matrix has 7 columns and its first row sum is 1, the first row is $[1,0,0,0,0,0,0]$. Thus a maximal matrix is uniquely determined by its row sums. Therefore we conclude that $\boldsymbol{B}_{t_a}$ is a lonesum matrix. One can check that matrices $\boldsymbol{B}_{t_b}$ and $\boldsymbol{B}_{t_c}$ of Table 3 are also lonesum. Thus, following our definition, response matrix $\boldsymbol{R}$ of Table 3 is unordered monotone. In our analysis of LATE in Section 4.1, $\boldsymbol{B}_{t_1}$ and $\boldsymbol{B}_{t_0}$ are both lonesum.

Table 7: Row and Column Sums of Matrix $\boldsymbol{B}_{t_a}$ of Response Matrix in Table 3

Row Sum	Row Index	**Matrix** $\boldsymbol{B}_{t_a} = \mathbf{1}[\boldsymbol{R} = t_a]$ of Table 3						
		$\boldsymbol{s}_1$	$\boldsymbol{s}_2$	$\boldsymbol{s}_3$	$\boldsymbol{s}_4$	$\boldsymbol{s}_5$	$\boldsymbol{s}_6$	$\boldsymbol{s}_7$
3	r_{1,t_a}	1	1	1	0	0	0	0
5	r_{2,t_a}	1	1	1	1	0	1	0
1	r_{3,t_a}	1	0	0	0	0	0	0
	Column Index	c_{1,t_a}	c_{2,t_a}	c_{3,t_a}	c_{4,t_a}	c_{5,t_a}	c_{6,t_a}	c_{7,t_a}
	Column Sum	3	2	2	1	0	1	0

Table 8: Reordered Matrix $\boldsymbol{B}_{t_a}$ According to Increasing Values of Row Sums and Decreasing Values of Column Sums

Row Sum	Row Index	Reordered Rows and Columns by Sums						
		s_1	s_2	s_3	s_4	s_6	s_5	s_7
1	r_{3,t_a}	1	0	0	0	0	0	0
3	r_{1,t_a}	1	1	1	0	0	0	0
5	r_{2,t_a}	1	1	1	1	1	0	0
	Column Index	c_{1,t_a}	c_{2,t_a}	c_{3,t_a}	c_{4,t_a}	c_{6,t_a}	c_{5,t_a}	c_{7,t_a}
	Column Sum	3	2	2	1	1	0	0

6.2 The Main Theorem

The following conditions are necessary and sufficient for characterizing Unordered Monotonicity **A-3**:

Theorem T-3. *The following statements are equivalent characterizations of* ***A-3*** *for the IV model* (1)–(3)*:*

(i) $\boldsymbol{R}$ *is an unordered monotone response matrix, i.e., each binary matrix* $\boldsymbol{B}_t = \mathbf{1}[\boldsymbol{R} = t]$, $t \in \text{supp}(T)$ *is lonesum;*

(ii) *For any* $t, t', t'' \in \text{supp}(T)$, *there are no* 2×2 *sub-matrices of* $\boldsymbol{R}$ *of the type:*

$$\begin{pmatrix} t & t' \\ t'' & t \end{pmatrix} \text{ or } \begin{pmatrix} t' & t \\ t & t'' \end{pmatrix}, \text{where } t' \neq t \text{ and } t'' \neq t.\text{[46]} \tag{48}$$

(iii) ***Unordered Monotonicity:*** *For any* $z, z' \in \text{supp}(Z)$, *and for each treatment* $t \in \text{supp}(T)$, *we have that:*[47]

$$\mathbf{1}[T_\omega(z) = t] \geq \mathbf{1}[T_\omega(z') = t] \; \forall \, \omega \in \Omega \quad \text{or} \quad \mathbf{1}[T_\omega(z) = t] \leq \mathbf{1}[T_\omega(z') = t] \; \forall \, \omega \in \Omega.$$

(iv) ***Unordered Separability:*** *Treatment choice can be represented by separable choice functions in* $\boldsymbol{V}$ *and* Z, *i.e., there exist functions* $\varphi : \text{supp}(\boldsymbol{V}) \times$

[46] These are called "prohibited patterns" or "forbidden patterns" in the literature.

[47] Alternatively, this can be written as: for any $z, z' \in \text{supp}(Z)$ and $t \in \text{supp}(T)$, we have that

$$\big(\mathbf{1}[T = t] | Z = z, \boldsymbol{V} = \boldsymbol{v}\big) \geq \big(\mathbf{1}[T = t] | Z = z', \boldsymbol{V} = \boldsymbol{v}\big) \text{ for all } \boldsymbol{v} \in \text{supp}(\boldsymbol{V})$$
$$\text{or } \big(\mathbf{1}[T = t] | Z = z, \boldsymbol{V} = \boldsymbol{v}\big) \leq \big(\mathbf{1}[T = t] | Z = z', \boldsymbol{V} = \boldsymbol{v}\big) \text{ for all } \boldsymbol{v} \in \text{supp}(\boldsymbol{V}).$$

$\text{supp}(T) \to \mathbb{R}$ *and* $\tau : \text{supp}(Z) \times \text{supp}(T) \to \mathbb{R}$ *such that:*

$$\mathbf{1}[T = t | \boldsymbol{V} = \boldsymbol{v}, Z = z] = \mathbf{1}[\Psi(t, z, \boldsymbol{v}) \geq 0] = \mathbf{1}[\varphi(\boldsymbol{v}, t) + \tau(z, t) \geq 0]. \quad (49)$$

Proof. See Web Appendix A.6. □

Condition (i) states our main condition for equivalence: if and only if response matrix $\boldsymbol{R}$ is unordered monotone, each indicator matrix formed from it ($\boldsymbol{B}_t = \mathbf{1}[\boldsymbol{R} = t]$) is lonesum. Condition (ii) states that if $\boldsymbol{R}$ is an unordered monotone response matrix, each 2×2 sub-matrix in $\boldsymbol{R}$ is not of the form in (48). Condition (iii) states that the conditions preceding it hold if and only if Unordered Monotonicity **A-3** holds. As previously noted, condition (iii) implies monotonicity **A-1** for the binary choice model. Condition (iv) is a separability property that characterizes the choice functions. Vytlacil's equivalence theorem (2002) is generated by the equivalence of conditions (iii) and (iv) when we specialize the model to the case of a binary treatment.[48]

6.3 Interpreting T-3

Condition (i) describes a key property of response matrices: the lonesum property of treatment choice indicators. Lonesum matrices are not only useful for characterizing Unordered Monotonicity, but they are key concepts for investigating properties of choice models.[49]

Condition (i) of **T-3** implies that $\boldsymbol{B}_t$ is fully characterized by its column and row sums. This condition implies that the response matrix $\boldsymbol{R}$ is also characterized by its row and column sums. However, the reverse is not true. We illustrate this in Remark *6.1*:

Remark 6.1. If $\boldsymbol{R}$ *is an unordered monotone response, each matrix* $\boldsymbol{B}_t$ *is lonesum and therefore fully characterized by its column and row sums* $r_{i,t}, c_{n,t}, t \in \text{supp}(T), i \in \{1, \ldots, N_Z\}, n \in \{1, \ldots, N_S\}$. *Since Response matrix* $\boldsymbol{R}$ *can be written as* $\boldsymbol{R} = \sum_{t \in \text{supp}(T)} t \boldsymbol{B}_t$, $\boldsymbol{R}$ *is characterized by its column and row sums* $r_{i,t}, c_{n,t}$ *as well. However, the reverse is not true.* $\boldsymbol{R}$ *being characterized by its column and row sums does not*

[48]See Web Appendix E for a derivation.

[49]Heckman and Pinto (2015c) show that lonesum matrices also play a key role in equivalence results for ordered monotonicity. Pinto (2016b) develops a framework for the design of social interventions using lonesum matrices that rely on revealed preference relationships to identify causal parameters. He shows that incentive designs based on lonesum matrices generate a range of monotonicity conditions.

imply that $\boldsymbol{R}$ *is an unordered monotone response. To illustrate this claim, let response matrix* $\boldsymbol{R}$ *be defined by:*

$$\boldsymbol{R} = \begin{pmatrix} t_1 & t_2 \\ t_2 & t_3 \end{pmatrix}, \; \textit{thus} \; \underbrace{\begin{matrix} r_{1,t_1} = 1, & r_{1,t_2} = 1, & r_{1,t_3} = 0, \\ r_{2,t_1} = 0, & r_{2,t_2} = 1, & r_{2,t_3} = 1, \end{matrix}}_{\textit{row sums}} ; \underbrace{\begin{matrix} c_{1,t_1} = 1, & c_{1,t_2} = 1, & c_{1,t_3} = 0, \\ c_{2,t_1} = 0, & c_{2,t_2} = 1, & c_{2,t_3} = 1. \end{matrix}}_{\textit{column sums}}$$

$\boldsymbol{R}$ *is not lonesum because* $\boldsymbol{B}_2 = \mathbf{1}[\boldsymbol{R} = t_2]$ *exhibits one of the prohibited patterns* (50)*. Therefore, it is not unordered monotone. Nevertheless,* $\boldsymbol{R}$ *is fully characterized by its column sums and row sums:* $r_{1,t_1} = 1$ *and* $c_{1,t_1} = 1 \Rightarrow \boldsymbol{R}[1,1] = t_1$*;* $r_{2,t_3} = 1$ *and* $c_{2,t_3} = 1 \Rightarrow \boldsymbol{R}[2,2] = t_3$*;* $r_{1,t_2} = 1$ *and* $\boldsymbol{R}[1,1] = t_1 \Rightarrow \boldsymbol{R}[1,2] = t_2$*;* $r_{2,t_2} = 1$ *and* $\boldsymbol{R}[2,2] = t_3 \Rightarrow \boldsymbol{R}[2,1] = t_2$*.*

All response matrices for the case of binary treatment are equivalent under monotonicity **A-1**. This property does not hold for the general unordered case:

Remark 6.2. Consider the binary choice model in which the instrument takes N_Z *values and* T *takes values in* $\{0,1\}$*. Unordered Monotonicity generates a monotonicity inequality for each pair of* Z*-values. Different sets of inequalities generate different response matrices. However, each of these response matrices is equivalent to the same lower triangular binary matrix with* N_Z *rows and* $N_Z + 1$ *columns (see the example in Section 4.1) and produces an identified model. However, in the case of multiple choices, Unordered Monotonicity does not generate response matrices that are equivalent to the same matrix. For example, the response matrices of Tables 3 and 5 are monotone responses but they are not equivalent, because one matrix cannot be transformed into another by row and/or column permutations. The response matrices in Tables 3 and 5 consist of seven response-types for* $N_T = 3$ *and* $N_Z = 3$*. There are 27 possible response-types for* $N_T = 3$ *and* $N_Z = 3$*. The combination of 7 response-types out of these 27 generates 888,030 possible response matrices. Among them, 66 response matrices satisfy Unordered Monotonicity condition (iii).*[50] *Response matrices of Tables 3 and 5 are two examples of these matrices.*

Condition (ii) of **T-3** imposes a restriction on counterfactual choices that does not depend on the number of treatment choices in supp(T) or the number of values that Z takes. The condition rules out two-way flows generated by changes in instruments. Thus the response matrix of Table 6 is not unordered monotone. The forbidden type of condition (ii) is obtained using the first and second rows of response-types $\boldsymbol{s}_2$ and $\boldsymbol{s}_4$ in Table 6.[51] The change from z_c to z_b shifts people away from a in $\boldsymbol{s}_2$ but toward

[50] Web Appendix F presents all of the 66 response matrices that consist of distinct sets of 7 response-types generated by Unordered Monotonicity **A-3**.

[51] In this case, we obtain the following forbidden sub-matrix: $\begin{pmatrix} t_a & t_c \\ t_b & t_a \end{pmatrix}$.

a in $\boldsymbol{s}_4$.

*Remark 6.3. We note that a consequence of condition (ii) in **T-3** is that under **A-3**, no 2×2 sub-matrix of any $\boldsymbol{B}_t$, $t \in \text{supp}(T)$ is of the type:*

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ nor } \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{50}$$

Unordered Monotonicity **A-3** holds if and only if no prohibited patterns (50) occur for any $\boldsymbol{B}_t$, $t \in \text{supp}(T)$. An example clarifies the equivalence between the requirements for Unordered Monotonicity **A-3** and the absence of prohibited patterns (50). Suppose that $\big(\mathbf{1}[T = t]|Z = z, \boldsymbol{V} = \boldsymbol{v}\big) \geq \big(\mathbf{1}[T = t]|Z = z', \boldsymbol{V} = \boldsymbol{v}\big)$ holds for all $\boldsymbol{v} \in \text{supp}(\boldsymbol{V})$. Then it must be the case that:

$$\big(\mathbf{1}[T = t]|Z = z, \boldsymbol{S} = \boldsymbol{s}\big) \geq \big(\mathbf{1}[T = t]|Z = z', \boldsymbol{S} = \boldsymbol{s}\big) \tag{51}$$

holds for all $\boldsymbol{s} \in \text{supp}(\boldsymbol{S})$ because for each $\boldsymbol{v} \in \text{supp}(\boldsymbol{V})$ there is $\boldsymbol{s} \in \text{supp}(\boldsymbol{S})$ such that $\boldsymbol{s} = f_S(\boldsymbol{v})$ (see (10)) and $(T|\boldsymbol{S} = \boldsymbol{s}, Z = z) = (T|\boldsymbol{V} = \boldsymbol{v}, Z = z)$. Inequality (51) generates three possible sub-vectors of dimension 2×1 that indicate whether T is equal to t when Z takes value z and z' or any response-type $\boldsymbol{s} \in \text{supp}(\boldsymbol{S})$:

$$\begin{pmatrix} (\mathbf{1}[T = t]|Z = z, S = \boldsymbol{s}) \\ (\mathbf{1}[T = t]|Z = z', S = \boldsymbol{s}) \end{pmatrix} \in \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\} \text{ for all } \boldsymbol{s} \in \text{supp}(\boldsymbol{S}). \tag{52}$$

The matrix generated by a combination of sub-vectors in (52) for any two response-types $\boldsymbol{s}, \boldsymbol{s}' \in \text{supp}(\boldsymbol{S})$ is:

$$\begin{pmatrix} (\mathbf{1}[T = t]|Z = z, \boldsymbol{S} = \boldsymbol{s}) & (\mathbf{1}[T = t]|Z = z, \boldsymbol{S} = \boldsymbol{s}') \\ (\mathbf{1}[T = t]|Z = z', \boldsymbol{S} = \boldsymbol{s}) & (\mathbf{1}[T = t]|Z = z', \boldsymbol{S} = \boldsymbol{s}') \end{pmatrix}.$$

It cannot be of the form:

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

which are prohibited patterns (50). Hence the weak inequality $\big(\mathbf{1}[T = t]|Z = z, \boldsymbol{V} = \boldsymbol{v})\big) \geq \big(\mathbf{1}[T = t]|Z = z', \boldsymbol{V} = \boldsymbol{v})\big) \,\forall\, \boldsymbol{v} \in \text{supp}(\boldsymbol{V})$ implies that $\boldsymbol{B}_t$ is lonesum. On the other hand, suppose that $\boldsymbol{v}, \boldsymbol{v}' \in \text{supp}(\boldsymbol{V})$ are such that $(\mathbf{1}[T = t]|Z = z, \boldsymbol{V} = \boldsymbol{v}) >$

[51] These are termed "prohibited" or "forbidden" patterns (see Ryser, 1957).

$(\mathbf{1}[T=t]|Z=z',\boldsymbol{V}=\boldsymbol{v})$ and $(\mathbf{1}[T=t]|Z=z,\boldsymbol{V}=\boldsymbol{v}') < (\mathbf{1}[T=t]|Z=z',\boldsymbol{V}=\boldsymbol{v}')$. Then there must exist $\boldsymbol{s},\boldsymbol{s}' \in \text{supp}(\boldsymbol{S})$ where $\boldsymbol{s}=f_S(\boldsymbol{v}),\boldsymbol{s}'=f_S(\boldsymbol{v}')$ that generates the prohibited pattern:

$$\begin{pmatrix} (\mathbf{1}[T=t]|Z=z,\boldsymbol{S}=\boldsymbol{s}) & (\mathbf{1}[T=t]|Z=z,\boldsymbol{S}=\boldsymbol{s}') \\ (\mathbf{1}[T=t]|Z=z',\boldsymbol{S}=\boldsymbol{s}) & (\mathbf{1}[T=t]|Z=z',\boldsymbol{S}=\boldsymbol{s}') \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{53}$$

The equality of the first columns in the right and left side of Equation (53) means that, for some type $\boldsymbol{s}$, treatment t is chosen when the instrument shifts from z' to z. The equality of the second columns of Equation (53) states the opposite. For some type $\boldsymbol{s}'$, the instrument shift from z' to z causes treatment t *not* to be chosen. This behavior violates the intuitive notion and formal definition of monotonicity because the instrument shifts some agents to change their choice towards t while others change their choice away from t.[52] Condition (iii) is implied by super (or sub) modularity of $\Psi(t,z,\boldsymbol{v})$ in terms of $\boldsymbol{v}$ and z for all t, but that condition is stronger than what is required to produce **A-3**. Strictly speaking, the requirement is that component-wise $\mathbf{sgn}(\frac{\Delta\Psi(T,z,\boldsymbol{v})}{\Delta z})$ is the same for all $\boldsymbol{V}=\boldsymbol{v}$ for each $T=t$ and $Z=z$.

6.4 Understanding Condition (iv) of T-3

We draw on and generalize the binary-treatment model of Sections 4.1–4.2 to build the intuition underlying condition (iv). In the binary case, monotonicity implies that $\boldsymbol{B}_{t_1}$ is lower triangular (28).[53] Triangularity generates Equation (36) which expresses treatment choice T as an indicator function that is separable in the observed propensity score $P(T=t_1|Z)$, which depends on Z, and a sum of response-type probabilities, which depends on V.

Theorem **T-3** applies to choice models with multiple treatments, which include the binary case. If Unordered Monotonicity (condition (iii) of **T-3**) holds, then each binary matrix $\boldsymbol{B}_t, t \in \text{supp}(T)$ is characterized solely by its row and column sums so that $\boldsymbol{B}_t, t \in \text{supp}(T)$ are lonesum (Condition (i) of **T-3**). This property can be understood as a generalization of the lower triangular property in the binary case, but applied to each $\boldsymbol{B}_t$.[54] Generalized triangularity generates condition (iv) which

[52]Violation of Condition (ii) is not necessarily a violation of rationality. Table 6 is based on an application of WARP, but violates Condition (ii) and Unordered Monotonicity.

[53]Recall that we eliminate all redundancies in the rows or columns of $\boldsymbol{R}$.

[54]With the caveat that we eliminate any redundancies in the rows and columns of $\boldsymbol{R}$.

characterizes treatment choice as an indicator function that is separable in Z and $\boldsymbol{V}$. We present a detailed discussion of this condition in Appendix G.

To interpret separability condition (iv), suppose that agent ω with $\boldsymbol{V}_\omega = \boldsymbol{v} \in \text{supp}(\boldsymbol{V})$ chooses $t \in \text{supp}(T)$ when an instrumental variable is set to $z \in supp(Z)$, so that $\mathbf{1}[T = t|\boldsymbol{V} = \boldsymbol{v}, Z = z] = 1$. According to condition (iv), there exist functions φ and τ such that $\varphi(\boldsymbol{v}, t) + \tau(z, t) \geq 0$.[55] It is clear that expressions of this type rule out the prohibited patterns (50) and therefore generate Unordered Monotonicity. What is less obvious is that (iii) implies representation (iv), which is not necessarily unique.[56]

Note that $\mathbf{1}[T = t|\boldsymbol{V} = \boldsymbol{v}, Z = z] = 1$ implies that $\mathbf{1}[T = t'|\boldsymbol{V} = \boldsymbol{v}, Z = z] = 0$ for all $t' \in \text{supp}(T) \setminus \{t\}$. Therefore it must be the case that $\varphi(\boldsymbol{v}, t') + \tau(z, t') < 0$ for all t' that differs from t. In particular, condition (iv) implies that:

$$\mathbf{1}[T = t|\boldsymbol{V} = \boldsymbol{v}, Z = z] = 1 \Leftrightarrow t = \underset{t' \in \text{supp}(T)}{\text{argmax}} \big(\Psi(t', z, \boldsymbol{v})\big) = \underset{t' \in \text{supp}(T)}{\text{argmax}} \big(\varphi(\boldsymbol{v}, t') + \tau(z, t')\big). \tag{54}$$

Condition (iv) does *not* claim that the functions φ and τ are unique. Indeed if t maximizes $\varphi(\boldsymbol{v}, t') + \tau(z, t')$, it also maximizes $m(\varphi(\boldsymbol{v}, t') + \tau(z, t'))$ where m is any strictly increasing function.

Condition (iv) does not impose rationality or perfect foresight. Suppose that agent ω decides among t_1, t_2, t_3 and that his treatment choice is generated by maximization of a utility function $\Psi(t, z, \boldsymbol{v})$ where $\boldsymbol{V}_\omega = \boldsymbol{v}$ and $Z_\omega = z$. Condition (iv) states that if Unordered Monotonicity **A-3** holds, the *maximized* choice value $\Psi(t, z, \boldsymbol{v})$ can be characterized as arising from the *maximization* of a separable function $\varphi(\boldsymbol{v}, t) + \tau(z, t)$. Specifically, if ω chooses t_1, then t_1 is the maximum among $\Psi(t, z, \boldsymbol{v})$ for $t \in \{t_1, t_2, t_3\}$. In this case, t_1 also maximizes $\varphi(\boldsymbol{v}, t) + \tau(z, t)$ for $t \in \{t_1, t_2, t_3\}$:

$$t_1 = \underset{t \in \{t_1, t_2, t_3\}}{\text{argmax}} \, \Psi(\boldsymbol{v}, t, z) \quad \Leftrightarrow \quad t_1 = \underset{t \in \{t_1, t_2, t_3\}}{\text{argmax}} \big(\varphi(\boldsymbol{v}, t) + \tau(z, t)\big).$$

[55] Web Appendix H discusses the threshold property of condition (iv) in greater detail.

[56] Consider a binary choice model: $T = \mathbf{1}[\alpha + V \cdot Z \geq 0]$ $T \in \{0, 1\}$ where (α, V) is a random vector and V, Z, and α are scalar. Suppose that we impose the requirement that $V > 0$ while Z is unrestricted. This is a monotone response model that is nonseparable. However, it can be represented as a separable model: $T = \mathbf{1}[\frac{\alpha}{V} + Z \geq 0]$. This highlights the non-uniqueness of the representation of $\Psi(t, Z, V)$ and that separability is only one characterization of preferences consistent with **A-3**.

Condition (iv) does *not* imply that the ranking of treatment utilities generated by $\Psi(t, z, \boldsymbol{v})$ is necessarily the same as the ranking generated by $\varphi(\boldsymbol{v}, t) + \tau(z, t)$. For instance, if $\Psi(t_1, z, \boldsymbol{v}) > \Psi(t_2, z, \boldsymbol{v}) > \Psi(t_3, z, \boldsymbol{v})$ then ω prefers t_1 to t_2, and t_2 to t_3. This does not necessarily imply that $\varphi(\boldsymbol{v}, t_1) + \tau(z, t_1) > \varphi(\boldsymbol{v}, t_2) + \tau(z, t_2) > \varphi(\boldsymbol{v}, t_3) + \tau(z, t_3)$. Indeed, $\varphi(\boldsymbol{v}, t_1) + \tau(z, t_1) > \varphi(\boldsymbol{v}, t_3) + \tau(z, t_3) > \varphi(\boldsymbol{v}, t_2) + \tau(z, t_2)$ may also occur. It is the ranking of t_1 relative to the next best that generates agent choices of t_1. Variation in instruments only identifies preferences relative to the next best choice and not an order among the remaining elements in the choice set.

To formalize this discussion, we establish that Unordered Monotonicity arises if we assume that utilities of a choice compared to the next best choice can be represented as additively separable functions:[57]

$$u(v, t) + h(z, t) = \Psi(t, z, \boldsymbol{v}) - \max_{t' \in \text{supp}(T) \setminus \{t\}} \Psi(t', z, \boldsymbol{v}).$$

The following theorem formalizes this point.

Theorem T-4. *If there exist functions* $u : \text{supp}(\boldsymbol{V}) \times \text{supp}(T) \to \mathbb{R}$ *and* $h : \text{supp}(Z) \times \text{supp}(T) \to \mathbb{R}$ *such that*

$$u(\boldsymbol{v}, t) + h(z, t) = \left(\Psi(t, z, \boldsymbol{v}) - \max_{t' \in \text{supp}(T) \setminus \{t\}} \Psi(t', z, \boldsymbol{v}) \right) \quad \forall\, \boldsymbol{v} \in \text{supp}(\boldsymbol{V}),\ z \in \text{supp}(Z),$$

then the response matrix $\boldsymbol{R}$ *associated with this choice model is unordered monotone.*

Proof. See Web Appendix A.7. □

As before, the separable representation is not necessarily unique.

Remark 6.4. ***T-4*** *imposes stronger functional form assumptions than* ***T-3****. Summa-*

[57] This transformation does not change an agent's preferences towards choices in $\text{supp}(T)$:

$$\Psi(t, z, \boldsymbol{v}) \geq \Psi(t', z, \boldsymbol{v}) \Leftrightarrow \left(\Psi(t, z, \boldsymbol{v}) - \max_{t' \in \text{supp}(T) \setminus \{t\}} \Psi(t', z, \boldsymbol{v}) \right) \geq \left(\Psi(t', z, \boldsymbol{v}) - \max_{t' \in \text{supp}(T) \setminus \{\tilde{t}\}} \Psi(t', z, \boldsymbol{v}) \right).$$

rizing:

$$\textbf{Unordered Monotonicity} \Rightarrow$$
$$\left(\underset{t\in\text{supp}(T)}{\text{argmax}}\, \varphi(\boldsymbol{v},t)+\tau(z,t)\right) = \left(\underset{t\in\text{supp}(T)}{\text{argmax}}\left(\Psi(t,z,\boldsymbol{v}) - \max_{t'\in\text{supp}(T)\setminus\{t\}} \Psi(t',z,\boldsymbol{v})\right)\right)$$
$$\text{while} \quad u(\boldsymbol{v},t)+h(z,t) = \left(\Psi(t,z,\boldsymbol{v}) - \max_{t'\in\text{supp}(T)\setminus\{t\}} \Psi(t',z,\boldsymbol{v})\right) \Rightarrow \textbf{Unordered Monotonicity}$$

Heckman et al. (2006, 2008) assume separability in the underlying preference functions and show that IV estimates a LATE that compares the outcome of one choice to the outcome for the next best option. Our condition is weaker. Theorem **T-4** states that Unordered Monotonicity only requires that the utility of a choice relative to the next best choice be separable. To clarify, the impact of instrument Z on the treatment choice is summarized by the term $h(z,t)$. Suppose Z changes from z' to z. If $h(z,t)-h(z',t)>0$, each agent is induced towards t. If $h(z,t)-h(z',t)<0$ agents are induced against t. This analysis applies for all pairwise values of $(z,z')\in \text{supp}(Z)\times\text{supp}(Z)$ and for all $t\in\text{supp}(T)$. The collection of all of these inequalities characterizes Unordered Monotonicity **A-3**.

6.5 Verifying Unordered Monotonicity condition A-3

Verifying condition (ii) of Theorem **T-3** is a daunting combinatorial task. It would require checking each 2×2 sub-matrix in $\boldsymbol{R}$, which is impractical for large $\boldsymbol{R}$. We show that a single calculation based on a simple multiplication of binary matrices suffices to check condition **A-3**. Our criterion is based on a binary matrix $\boldsymbol{M}$:

$$\begin{aligned}
&\text{For each } t_j\in\text{supp}(T)=\{t_1,\ldots,t_{N_T}\},\\
&\qquad \text{let } \boldsymbol{M}_{t_j} = [\underbrace{\mathbf{1}_{N_Z,N_S},\ \cdots\ ,\ \mathbf{1}_{N_Z,N_S}}_{j-1 \text{ times}}, \boldsymbol{B}_{t_j}, \underbrace{\mathbf{0}_{N_Z,N_S},\ \cdots\ ,\ \mathbf{0}_{N_Z,N_S}}_{N_T-j \text{ times}}],\\
&\qquad \text{then } \boldsymbol{M} = [\boldsymbol{M}'_{t_1},\ \ldots\ ,\ \boldsymbol{M}'_{t_{N_T}}]',
\end{aligned} \tag{55}$$

where $\mathbf{1}_{N_Z,N_S}$ is a matrix of elements 1 and $\mathbf{0}_{N_Z,N_S}$ is a matrix of elements 0 of same dimension. Matrix $\boldsymbol{M}$ is block diagonal with matrices $\boldsymbol{B}_t$ on the diagonal, where, again, we eliminate any redundancies. $\boldsymbol{M}$ has elements 1 below this diagonal and elements 0 above it.

Theorem T-5. *For the IV model (1)–(3), the response matrix* $\boldsymbol{R}$ *is an unordered* *monotone response, that is, each binary matrix* $\boldsymbol{B}_t = \mathbf{1}[\boldsymbol{R} = t]$; $t \in \text{supp}(T)$ *is lonesum, if and only if,*

$$\iota_c'\Big(\big(\boldsymbol{M}'(\iota_r\iota_c' - \boldsymbol{M})\big) \odot \big(\boldsymbol{M}'(\iota_r\iota_c' - \boldsymbol{M})\big)' \Big)\iota_c = 0, \tag{56}$$

where ι_r *is an* $N_T \cdot N_Z$ *vector* $1s$ *and* ι_c *is an* $N_T \cdot N_S$ *vector* $1s$*. Moreover, if Equation* (56) *holds, then matrix* $\boldsymbol{M}$ *is lonesum.*

Proof. See Web Appendix A.8. □

Unordered monotonicity condition **A-3** holds if and only if this value is equal to zero. Moreover, if equation (56) holds, then all the conditions stated in Theorem **T-3** also hold.

7 Identification of Counterfactuals

This section establishes which components of a model are identified under Unordered Monotonicity **A-3**. We build on our analysis of binary LATE in Section 4.1. We generalize the notions of "compliers" and "always-takers" to a general unordered choice model.

To this end, it is helpful to introduce some additional notation. Let $\Sigma_t(i)$ be the *set* of response-types in which t appears exactly i times:

$$\Sigma_t(i) = \left\{ \boldsymbol{s}_n, \text{ such that } \boldsymbol{s}_n \in \text{supp}(\boldsymbol{S}) \text{ and } \sum_{j'=1}^{N_Z} B_t[j', n] = i \right\} \text{ where } i \in \{0, \ldots, N_Z\}. \tag{57}$$

For example, $\Sigma_{t_a}(2)$ for the response matrix of Table 3 consists of the response-types for which the value t_a appears exactly twice. They are $\Sigma_{t_a}(2) = \{\boldsymbol{s}_2, \boldsymbol{s}_3\}$ (see Table 9). Those are also the response-types whose column-sum of $\boldsymbol{B}_{t_a}$ in Table 7 is 2.

Table 9: Partition of Response-Types in Table 3 where $\text{supp}(Z) = \{z_{no}, z_a, z_{bc}\}$

Instrumental Variables	Count. Choices	$\boldsymbol{s}_1$	$\boldsymbol{s}_2$	$\boldsymbol{s}_3$	$\boldsymbol{s}_4$	$\boldsymbol{s}_5$	$\boldsymbol{s}_6$	$\boldsymbol{s}_7$
		Response-types of S						
No Voucher	$T(z_{no})$	t_a	t_a	t_a	t_b	t_b	t_c	t_c
Voucher for car a	$T(z_a)$	t_a	t_a	t_a	t_a	t_b	t_a	t_c
Voucher for car b or c	$T(z_{bc})$	t_a	t_b	t_c	t_b	t_b	t_c	t_c
Response-types in $\Sigma_{t_a}(0)$						$\boldsymbol{s}_5$		$\boldsymbol{s}_7$
Response-types in $\Sigma_{t_a}(1)$					$\boldsymbol{s}_4$		$\boldsymbol{s}_6$	
Response-types in $\Sigma_{t_a}(2)$			$\boldsymbol{s}_2$	$\boldsymbol{s}_3$				
Response-types in $\Sigma_{t_a}(3)$		$\boldsymbol{s}_1$						
t_a-Compliers			$\boldsymbol{s}_2$	$\boldsymbol{s}_3$	$\boldsymbol{s}_4$		$\boldsymbol{s}_6$	
t_a-Always-takers		$\boldsymbol{s}_1$						
t_a-Never-takers						$\boldsymbol{s}_5$		$\boldsymbol{s}_7$

For each $t \in \text{supp}(T)$, we can partition the set of response-types by the number of times a treatment value t appears: $\text{supp}(\boldsymbol{S}) = \cup_{i=0}^{N_Z} \Sigma_t(i)$. Table 9 displays these partitions for $\Sigma_{t_a}(i); i = 0, \ldots, 3$. Let $\boldsymbol{b}_t(i)$ be the N_S-dimensional binary row-vector that indicates if response-type $\boldsymbol{s}$ belongs to $\Sigma_t(i)$, that is, $\boldsymbol{b}_t(i)[n] = 1$ if $\boldsymbol{s}_n \in \Sigma_t(i)$ and zero otherwise. In the example of Table 3, $\boldsymbol{b}_{t_a}(2) = [0, 1, 1, 0, 0, 0, 0]$. Using this notation, we prove the following identification theorem:

Theorem T-6. *If Unordered Monotonicity **A-3** holds for the IV model (1)–(3), then, for all $t \in \text{supp}(T)$ and for all $i \in \{1, \ldots, N_Z\}$, the probabilities $P(\boldsymbol{S} \in \Sigma_t(i))$ and the outcome counterfactual expectations $E(\kappa(Y(t))|\boldsymbol{S} \in \Sigma_t(i))$ are identified from:*

$$P(\boldsymbol{S} \in \Sigma_t(i)) = \boldsymbol{b}_t(i)\boldsymbol{B}_t^+\boldsymbol{P}_Z(t) \quad \text{and} \quad E(\kappa(Y(t))|\boldsymbol{S} \in \Sigma_t(i)) = \frac{\boldsymbol{b}_t(i)\boldsymbol{B}_t^+\boldsymbol{Q}_Z(t)}{\boldsymbol{b}_t(i)\boldsymbol{B}_t^+\boldsymbol{P}_Z(t)}, \tag{58}$$

where $\kappa : \text{supp}(Y) \to \mathbb{R}$ denotes an arbitrary function in the support of Y.

Proof. See Web Appendix A.9. □

This expression uses the tools for identification based on the generalized inverse developed in Section 4.

Remark 7.1. *A direct implication of Theorem* ***T-6*** *is that if there exists* $t, t' \in \text{supp}(T)$ *and* $i, i' \in \{1, \ldots, N_Z\}$ *such that* $\Sigma_t(i) = \Sigma_{t'}(i')$ *then* $E(Y(t)-Y(t')|\Sigma_t(i))$ *is identified.*

We generalize the terminology of Angrist et al. (1996) to the case of multiple treatments. The appropriate generalization is t-specific. In the binary case, there is no need to specify a particular t since the specification of one value automatically implies the other possible value.

$\Sigma_t(0)$ consists of response-types in which t does not occur. We call these t-*Never-takers*. $\Sigma_t(N_Z)$ consists of a single response-type whose elements are all t. It stands for the t-*Always-takers*. The set of remaining response-types are named t-*Compliers* $\equiv \bigcup_{i=1}^{N_Z-1} \Sigma_t(i)$ and consists of all strata for which the choice of treatment t varies as Z ranges in its support. Those sets can alternatively be defined as:

$$\begin{aligned}
t\text{-}\textit{Never-takers} &= \{\boldsymbol{s} \in \text{supp}(\boldsymbol{S}); P(T=t|\boldsymbol{S}=\boldsymbol{s})=0\} \equiv \Sigma_t(0);\\
t\text{-}\textit{Always-takers} &= \{\boldsymbol{s} \in \text{supp}(\boldsymbol{S}); P(T=t|\boldsymbol{S}=\boldsymbol{s})=1\} \equiv \Sigma_t(N_Z);\\
t\text{-}\textit{Compliers} &= \{\boldsymbol{s} \in \text{supp}(\boldsymbol{S}); 0 < P(T=t|\boldsymbol{S}=\boldsymbol{s})<1\} \equiv \bigcup_{i=1}^{N_Z-1} \Sigma_t(i).
\end{aligned}$$

These sets for the response matrix of Table 3 are: t_a-*Always-takers* $= \{\boldsymbol{s}_1\}$, t_a-*Never-takers* $= \{\boldsymbol{s}_5, \boldsymbol{s}_7\}$, and t_a-*Compliers* $= \{\boldsymbol{s}_2, \boldsymbol{s}_3, \boldsymbol{s}_4, \boldsymbol{s}_6\}$ (see Table 9). Corollaries **C-2**–**C-3** present identification results based on this terminology:

Corollary C-2. *For the IV model* (1)–(3) *in which Unordered Monotonicity* ***A-3*** *holds, the following probabilities are identified for each* $t \in \text{supp}(T)$*:*

$$\begin{aligned}
P(\boldsymbol{S} \in t\text{-}\textit{Always-takers}) &= P(\boldsymbol{S} \in \Sigma_t(N_Z)) = \boldsymbol{b}_t(N_Z)\boldsymbol{B}_t^{+}\boldsymbol{P}_Z(t);\\
P(\boldsymbol{S} \in t\text{-}\textit{Compliers}) &= P(\boldsymbol{S} \in \cup_{i=1}^{N_Z-1}\Sigma_t(i)) = \left(\sum_{i=1}^{N_Z-1} \boldsymbol{b}_t(i)\right)\boldsymbol{B}_t^{+}\boldsymbol{P}_Z(t);\\
P(\boldsymbol{S} \in t\text{-}\textit{Never-takers}) &= P(\boldsymbol{S} \in \Sigma_t(0))\\
&= 1 - P(\boldsymbol{S} \in t\text{-}\textit{Always-takers}) - P(\boldsymbol{S} \in t\text{-}\textit{Compliers}).
\end{aligned}$$

Proof. See Web Appendix A.10. □

Corollary C-3. *Assume the IV model* (1)–(3) *for which Unordered Monotonicity* ***A-3*** *holds. The mean counterfactual outcomes for the* t*-Always-takers and* *t-Compliers for each $t \in \text{supp}(T)$ are identified by:*

$$E(Y(t)|t\text{-}Always\text{-}takers) = E(Y(t)|\boldsymbol{S} \in \Sigma_t(N_Z)) = \frac{\boldsymbol{b}_t(N_Z)\boldsymbol{B}_t^+\boldsymbol{Q}_Z(t)}{\boldsymbol{b}_t(N_Z)\boldsymbol{B}_{t_a}^+\boldsymbol{P}_Z(t)};$$

$$E(Y(t)|t\text{-}Compliers) = \sum_{i=1}^{N_Z-1} E(Y(t)|\boldsymbol{S} \in \Sigma_t(i)) \cdot \frac{P(\boldsymbol{S} \in \Sigma_t(i))}{P(\boldsymbol{S} \in t\text{-}Compliers)},$$

$$\text{where} \sum_{i=1}^{N_Z-1} \frac{P(\boldsymbol{S} \in \Sigma_t(i))}{P(\boldsymbol{S} \in t\text{-}Compliers)} = 1;$$

$$\text{Also } E(Y(t)|t\text{-}Compliers) = \frac{\left(\sum_{i=1}^{N_Z-1}\boldsymbol{b}_t(i)\right)\boldsymbol{B}_t^+\boldsymbol{Q}_Z(t)}{\left(\sum_{i=1}^{N_Z-1}\boldsymbol{b}_t(i)\right)\boldsymbol{B}_{t_a}^+\boldsymbol{P}_Z(t)}. \tag{59}$$

Proof. See Web Appendix A.11. □

Corollary **C-2** relies on the result in **T-6** that $P(\boldsymbol{S} \in \Sigma_t(i))$ is identified for all $i \in \{1, \ldots, N_Z\}$. Corollary **C-3** is obtained by setting $\kappa(Y) = Y$ and using the fact that $E(Y(t)|\boldsymbol{S} \in \Sigma_t(i))$ is identified for all $i \in \{1, \ldots, N_Z\}$. To illustrate these corollaries we present an example.

Remark 7.2. Corollary ***C-2*** *states that the expected value of counterfactual outcomes for response-types $s \in \text{supp}(S)$ such that $P(T = t|\boldsymbol{S} = \boldsymbol{s}) = 1$ (the* t-Always-takers*) or $0 < P(T = t|\boldsymbol{S} = \boldsymbol{s}) < 1$ (*t-Compliers*) are identified. According to Remark 3.1, these response-types refer to the values $v \in \text{supp}(\boldsymbol{V})$ such that $0 < P(T = t|\boldsymbol{V} = \boldsymbol{v}) \leq 1$. Therefore,* ***C-3*** *implies that $E(Y(t)|\boldsymbol{V} \in \{\boldsymbol{v}; 0 < P(T = t|\boldsymbol{V} = \boldsymbol{v}) \leq 1\})$ is identified. The remaining set of response-types are the* t-Never-takers, *which consists of the response-types $\boldsymbol{s} \in \text{supp}(S)$ such that $P(T = t|\boldsymbol{S} = \boldsymbol{s}) = 0$. This set refers to the set of values $\boldsymbol{v} \in \text{supp}(\boldsymbol{V})$ such that $P(T = t|\boldsymbol{V} = \boldsymbol{v}) = 0$. If the set of* t-Never-takers *is empty, then all response-types belong to either* t-Always-takers *or* t-Compliers *and $E(Y(t))$ is identified.*

Example 7.1. According to **C-3**, the counterfactual outcome mean for t_a-*Compliers* in the response matrix $\boldsymbol{R}$ of Table 3 is given by:

$$E(Y(t_a)|t_a - Compliers) = E(Y(t_a)|S \in \{\boldsymbol{s}_2, \boldsymbol{s}_3, \boldsymbol{s}_4, \boldsymbol{s}_6\}) = \frac{\left(\sum_{i=1}^{2}\boldsymbol{b}_{t_a}(i)\right)\boldsymbol{B}_{t_a}^+\boldsymbol{Q}_Z(t_a)}{\left(\sum_{i=1}^{2}\boldsymbol{b}_{t_a}(i)\right)\boldsymbol{B}_{t_a}^+\boldsymbol{P}_Z(t_a)} \tag{60}$$

The components of Equation (60) that can be estimated from observed data are:

$$\boldsymbol{P}_Z(t_a) = [P(T = t_a|Z = z_{no}), P(T = t_a|Z = z_a), P(T = t_a|Z = z_{bc})]';$$
$$\boldsymbol{Q}_Z(t_a) = [E(Y \cdot \mathbf{1}[T = t_a]|Z = z_{no}), E(Y \cdot \mathbf{1}[T = t_a]|Z = z_a), E(Y \cdot \mathbf{1}[T = t_a]|Z = z_{bc})]'.$$

The components of (60) that depend on the response matrix are:

$$\sum_{i=1}^{2} \boldsymbol{b}_{t_a}(i) = [0,1,1,1,0,1,0];$$

$$\boldsymbol{B}_{t_a} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \Rightarrow \boldsymbol{B}_{t_a}^{+} = \begin{bmatrix} 0 & 0 & 1 \\ 1/2 & 0 & -1/2 \\ 1/2 & 0 & -1/2 \\ -1/2 & 1/2 & 0 \\ 0 & 0 & 0 \\ -1/2 & 1/2 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Equation (60) produces the following expression:

$$E(Y(t_a)|t_a\text{-}Compliers) = \frac{E(Y \cdot \mathbf{1}[T=t_a]|Z=z_a) - E(Y \cdot \mathbf{1}[T=t_a]|Z=z_{bc})}{P(T=t_a|Z=z_a) - P(T=t_a|Z=z_{bc})}.$$

Web Appendix A.16 presents additional results on identification.

7.1 Maximum Number of Admissible Response-types to Secure Identification

The identification of strata probabilities $P(\boldsymbol{s}) \in \boldsymbol{S}$ can be achieved under weaker conditions than are required for identifying counterfactual outcomes. The identification of response-type probabilities depends on the column-rank of $\boldsymbol{B}_T$ while the identification of counterfactual outcomes for a choice $t \in \text{supp}(T)$ depends on the column-rank of $\boldsymbol{B}_t$. The rank of $\boldsymbol{B}_T$ is always greater than the rank of each $\boldsymbol{B}_t; t \in \text{supp}(T)$ because $\boldsymbol{B}_T$ is generated by stacking $\boldsymbol{B}_t$ across $t \in supp(T)$ (Equation (24)).

We characterize the maximum number of response-types N_S in $\boldsymbol{R}$ that facilitate the identification of all response-type probabilities, that is, the maximum N_S such that $N_S \leq \text{rank}(\boldsymbol{B}_T)$.

Theorem T-7. *Consider the IV model (1)–(3). Let $\boldsymbol{R}$ be the response matrix consisting of N_S response-types. If response-type probabilities are point identified, then it must be the case that:*

$$N_S \leq 1 + (N_T - 1)N_Z - \sum_{i=1}^{N_Z} \sum_{j=1}^{N_T} \mathbf{1}[P(T=t_j|Z=z_i)=0],$$

where N_Z is the number of possible values that the instrument takes and N_T is the *number of possible values that the treatment choice T takes. In particular, if $P(T = t|Z = z) > 0$ for all $z \in \text{supp}(Z)$ and $t \in$ supp(T) then the maximum number of response-types N_S in $\boldsymbol{R}$ for the model to be identified is:*

$$N_S = 1 + (N_T - 1)N_Z.$$

Proof. See Web Appendix A.12. □

To identify choice-response probabilities, choice restrictions should eliminate at least $N_T^{N_Z} - [1 + (N_T - 1)N_Z]$ response-types to generate identification of response-type probabilities. **T-6** shows that even if we are not able to identify each probability $P(\boldsymbol{S} = \boldsymbol{s}); \boldsymbol{s} \in \text{supp}(\boldsymbol{S})$, it may still be possible to identify counterfactual means $E(Y(t)|S \in \Sigma_t(i))$ and associated probabilities $P(\boldsymbol{S} \in \Sigma_t(i))$ for strata $\Sigma_t(i)$. See Web Appendix A.13 for additional results on identification of strata probabilities.

8 Summary and Conclusions

This paper extends the literature on instrumental variables in general unordered choice models with heterogenous responses which affect choice and present a new monotonicity condition. Using discrete instruments, we identify the counterfactuals associated with general unordered multiple discrete choice models. We represent IV equations using discrete mixtures. Identification is achieved by imposing restrictions on the kernels of the mixtures. We generalize the notion of monotonicity to unordered choice models. We do not invoke separability of preferences or identification at infinity to achieve these results, although a version of separability of choice equations emerges as one representation of the underlying choice process.

Unordered monotonicity can sometimes be justified by economic choice models. It can be represented in multiple ways. These representations are linked to properties of binary matrices that characterize the admissible response-types generated by the available instrumental variables. We develop a variety of criteria to determine if Unordered Monotonicity is satisfied. We interpret each of these criteria and explain how they can be used in practice. We show that "principal strata" in the statistics literature are coarse versions of control functions.

This paper demonstrates the power of binary matrices in generating identification and in unifying seemingly diverse approaches to identification. The broader lesson of this paper is that in general unordered discrete choice models restrictions on choice behavior, encoded in our generalized version of monotonicity, play a fundamental role in identifying counterfactuals using instrumental variables.

References

Angrist, J. D. and G. W. Imbens (1995): "Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity," *Journal of the American Statistical Association*, 90, 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, 444–455.

Blevins, J. R. (2014): "Nonparametric Identification of Dynamic Decision Processes with Discrete and Continuous Choices," *Quantitative Economics*, 5, 531–554.

Brualdi, R. A. (1980): "Matrices of zeros and ones with fixed row and column sum vectors," *Linear Algebra and Its Applications*, 33, 159–231.

Brualdi, R. A. and H. J. Ryser (1991): *Combinatorial Matrix Theory*, Encyclopedia of Mathematics and its Applications, New York, NY: Cambridge University Press.

Carrasco, M., J.-P. Florens, and E. Renault (2007): "Linear Inverse Problems in Structural Econometrics Estimation Based on Spectral Decomposition and Regularization," in *Handbook of Econometrics*, ed. by J. J. Heckman and E. Leamer, Amsterdam: Elsevier, vol. 6B, 5633–5751.

Clogg, C. C. (1995): "Latent Class Models," in *Handbook of Statistical Modeling for the Social and Behavioral Sciences*, ed. by G. Arminger, C. C. Clogg, and M. E. Sobel, New York: Plenum Press, chap. 6, 311–359.

Frangakis, C. E. and D. Rubin (2002): "Principal stratification in causal inference," *Biometrics*, 58, 21–29.

Frisch, R. (1933): "Editor's Note," *Econometrica*, 1, 1–4.

——— (1938): "Autonomy of Economic Relations: Statistical versus Theoretical Relations in Economic Macrodynamics," Paper given at League of Nations. Reprinted in D.F. Hendry and M.S. Morgan (1995), *The Foundations of Econometric Analysis*, Cambridge University Press.

Haavelmo, T. (1943): "The Statistical Implications of a System of Simultaneous Equations," *Econometrica*, 11, 1–12.

——— (1944): "The Probability Approach in Econometrics," *Econometrica*, 12, iii–vi and 1–115.

Hansen, L. P. (1982): "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50, 1029–1054.

Heckman, J. J. (2008): "The Principles Underlying Evaluation Estimators with an Application to Matching," *Annales d'Economie et de Statistiques*, 91–92, 9–73.

——— (2010): "Building Bridges Between Structural and Program Evaluation Approaches to Evaluating Policy," *Journal of Economic Literature*, 48, 356–398.

Heckman, J. J. and R. Pinto (2015a): "Alternative Ways to Identify P_S," Unpublished manuscript, University of Chicago, August 2015.

——— (2015b): "Causal Analysis after Haavelmo," *Econometric Theory*, 31, 115–151.

——— (2015c): "Comparing Ordered and Unordered Choice Models," Unpublished manuscript, University of Chicago, July 2015.

Heckman, J. J. and R. Robb (1985): "Alternative Methods for Evaluating the Impact of Interventions: An Overview," *Journal of Econometrics*, 30, 239–267.

Heckman, J. J., S. Urzúa, and E. J. Vytlacil (2006): "Understanding Instrumental Variables in Models with Essential Heterogeneity," *Review of Economics and Statistics*, 88, 389–432.

——— (2008): "Instrumental Variables in Models with Multiple Outcomes: The General Unordered Case," *Les Annales d'Economie et de Statistique*, 91–92, 151–174.

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

——— (2007a): "Econometric Evaluation of Social Programs, Part I: Causal Models, Structural Models and Econometric Policy Evaluation," in *Handbook of Econometrics*, ed. by J. Heckman and E. Leamer, Amsterdam: Elsevier, vol. 6B, 4779–4874.

——— (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 *Handbook of Econometrics*, ed. by J. Heckman and E. Leamer, Amsterdam: Elsevier, vol. 6B, chap. 71, 4875–5143.

Henry, M., Y. Kitamura, and B. Salanié (2014): "Partial Identification of Finite Mixtures in Econometric Models," *Quantitative Economics*, 5, 123–144.

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

Lee, S. and B. Salanié (2016): "Identifying Effects of Multivalued Treatments," Unpublished manuscript.

Magnus, J. and H. Neudecker (1999): *Matrix Differential Calculus with Applications in Statistics and Econometrics*, Wiley, 2 ed.

Pinto, R. (2016a): "Learning from Noncompliance in Social Experiments: The Case of Moving to Opportunity." Unpublished paper based on Ph.D. Thesis, University of Chicago, Department of Economics.

——— (2016b): "Randomized Biased-Controlled Trials: Adding Incentives to the Design of Experiments." Mpublished Working Paper.

Prakasa Rao, B. L. S. (1992): "Identifiability for Mixtures of Distributions," in *Identifiability in Stochastic Models: Characterization of Probability Distributions*, Boston, MA: Academic Press, chap. 8, 183–228.

Quandt, R. E. (1972): "A New Approach to Estimating Switching Regressions," *Journal of the American Statistical Association*, 67, 306–310.

Robins, J. M. and S. Greenland (1992): "Identifiability and Exchangeability for Direct and Indirect Effects," *Epidemiology*, 3, 143–155.

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

Sachnov, V. N. and V. E. Tarakanov (2002): *Combinatorics of Nonnegative Matrices*, no. 213 in Translations of Mathematical Monographs, Providence, RI: American Mathematical Society, translated by Valentin F. Kolchin.

Theil, H. (1953): *Esimation and Simultaneous Correlation in Complete Equation Systems*, The Hague: Central Planning Bureau, mimeographed memorandum.

——— (1958): *Economic Forecasts and Policy (Contributions to Economic Analysis no. 15)*, Amsterdam: North-Holland Publishing Company.

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

——— (2004): "Ordered Discrete Choice Selection Models: Equivalence, Nonequivalence, and Representation Results," Unpublished manuscript, Stanford University, Department of Economics.

——— (2006): "Ordered Discrete-Choice Selection Models and Local Average Treatment Effect Assumptions: Equivalence, Nonequivalence, and Representation Results," *Review of Economics and Statistics*, 88, 578–581.

Yakowitz, S. J. and J. D. Spragins (1968): "On the Identifiability of Finite Mixtures," *Annals of Mathematical Statistics*, 39, 209–214.