Web Appendix for

Unordered Monotonicity

James J. Heckman and Rodrigo Pinto

Revised October 2, 2017

Contents

A Mathematical Appendix: Proofs of Lemmas and Theorems

A.1 Proof of Lemma L-1:

Proof. Using the setup of (B.1) and (B.2), it is easy to see that the independence relationship $(\epsilon_{\boldsymbol{V}}, \epsilon_Y) \perp\!\!\!\perp \epsilon_Z$ implies the exclusion restriction $(\boldsymbol{V}, Y(t)) \perp\!\!\!\perp Z$ of Equation (6). Moreover, $\epsilon_Y \perp\!\!\!\perp \epsilon_Z | \epsilon_{\boldsymbol{V}}$ implies the matching (conditional independence) property $Y(t) \perp\!\!\!\perp T | \boldsymbol{V}$ of Equation (7). The independence relationship, $\boldsymbol{V} \perp\!\!\!\perp Z$, implies that $f_{\boldsymbol{S}}(\boldsymbol{V}) \perp\!\!\!\perp Z$, hence $\boldsymbol{S} \perp\!\!\!\perp Z$, as stated in item (ii) of the lemma. Equation (11), i.e., $T = g_T(\boldsymbol{S}, Z)$, implies that T is deterministic conditional on $\boldsymbol{S}$ and Z. Thus, $Y \perp\!\!\!\perp T | (\boldsymbol{S}, Z)$ holds as stated in item (iii) of the lemma.

From $(\boldsymbol{V}, Y(t)) \perp\!\!\!\perp Z$ and the fact that $\boldsymbol{S} = f_{\boldsymbol{S}}(\boldsymbol{V})$, we obtain $(\boldsymbol{S}, Y(t)) \perp\!\!\!\perp Z$. We can apply the Weak Union Property of conditional independence relationships of the Graphoid axioms to obtain $Y(t) \perp\!\!\!\perp Z | \boldsymbol{S}$.[1] But T only depends on Z when conditioned on $\boldsymbol{S}$ (Equation (11)), thus we have that $Y(t) \perp\!\!\!\perp f_T(Z, \boldsymbol{S}) | \boldsymbol{S}$ which is equivalent to $Y(t) \perp\!\!\!\perp T | \boldsymbol{S}$ as stated in item (i) of the lemma.

Independence relationship $Y(t) \perp\!\!\!\perp Z | \boldsymbol{S}$ implies that $Y(t) \perp\!\!\!\perp (f_T(Z, \boldsymbol{S}), Z) | \boldsymbol{S}$ also holds. This relation is equivalent to $Y(t) \perp\!\!\!\perp (T, Z) | \boldsymbol{S}$. By Weak Union and Decomposition we have that $Y(t) \perp\!\!\!\perp Z | (\boldsymbol{S}, T)$ holds. In particular, $Y(t) \perp\!\!\!\perp Z | (\boldsymbol{S}, T = t)$ holds for all $t \in \text{supp}(T)$. From representation (4):

[1] The Graphoid axioms are a set of conditional independence relations first presented by Dawid (1979):

$$
\begin{aligned}
\text{Symmetry: } & X \perp\!\!\!\perp Y | Z \Rightarrow Y \perp\!\!\!\perp X | Z. \\
\text{Decomposition: } & X \perp\!\!\!\perp (W, Y) | Z \Rightarrow X \perp\!\!\!\perp Y | Z. \\
\text{Weak Union: } & X \perp\!\!\!\perp (W, Y) | Z \Rightarrow X \perp\!\!\!\perp W | (Y, Z). \\
\text{Contraction: } & X \perp\!\!\!\perp Y | Z \text{ and } X \perp\!\!\!\perp W | (Y, Z) \Rightarrow X \perp\!\!\!\perp (W, Y) | Z. \\
\text{Intersection: } & X \perp\!\!\!\perp W | (Y, Z) \text{ and } X \perp\!\!\!\perp Y | (W, Z) \Rightarrow X \perp\!\!\!\perp (W, Y) | Z. \\
\text{Redundancy: } & X \perp\!\!\!\perp Y | X.
\end{aligned}
$$

The intersection relation is only valid for strictly positive probability distributions.

$$
\begin{aligned}
&\left(Y(t) \perp\!\!\!\perp Z | (\boldsymbol{S}, T = t)\right) \\
\Rightarrow &\left(\sum_{t' \in \text{supp}(T)} Y(t') \cdot \mathbf{1}[T = t'] \perp\!\!\!\perp Z | (\boldsymbol{S}, T = t)\right) \\
\Rightarrow &\left(Y \perp\!\!\!\perp Z | (\boldsymbol{S}, T = t)\right) \,\forall\, t \in \text{supp}(T),
\end{aligned}
$$

which proves item (iv) of the lemma. □

A.2 Proof of Theorem T-1:

Proof. We make no assumption about the functional form of f_Y in the outcome equation $Y = f_Y(T, \boldsymbol{V}, \epsilon_Y)$ in (2) except $E(|Y|) < \infty$. Without loss of generality, the outcome equation can be replaced by $Y = \kappa(f_Y(T, \boldsymbol{V}, \epsilon_Y))$. To prove the theorem, it suffices to show that the following relationship holds for the expectation of Y.

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

The first equality comes from the law of iterated expectations. The second equality comes from Bayes' theorem. The first term of the third equality comes from the fact the T is deterministic conditioned on Z and $\boldsymbol{S}$. The second term in the expression comes from $Y \perp\!\!\!\perp Z | (\boldsymbol{S}, T)$, a consequence of Lemma **L-1**. The third term comes from $\boldsymbol{S} \perp\!\!\!\perp Z$ as established in Lemma **L-1**. The fourth equality comes from conditional independence $Y(t) \perp\!\!\!\perp T | \boldsymbol{S}$ of Lemma **L-1**, which implies that $E(Y | T = t, \boldsymbol{S} = \boldsymbol{s}) = E(Y(t) | \boldsymbol{S} = \boldsymbol{s})$. □

A.3 Proof of Theorem T-2:

Restrictions on the response matrix $\boldsymbol{R}$ generate identification of mean counterfactuals *defined on strata*. We rely on Lemma **L-2**—stated below—to prove Theorem **T-2**. Lemma **L-2** states the general solution for a system of linear equations. We refer to Magnus and Neudecker (1999) for a general discussion of linear systems.

Lemma L-2. A general solution for $\boldsymbol{x}$ in the system of linear equations represented by $\boldsymbol{b} = \boldsymbol{B}\boldsymbol{x} \Rightarrow \boldsymbol{x}$ is given by:

$$\boldsymbol{b} = \boldsymbol{B}\boldsymbol{x} \Rightarrow \boldsymbol{x} = \boldsymbol{B}^{+}\boldsymbol{b} + (\boldsymbol{I} - \boldsymbol{B}^{+}\boldsymbol{B})\boldsymbol{\lambda} \tag{A.1}$$

where $\boldsymbol{\lambda}$ is an arbitrary real-valued $|\boldsymbol{b}|$-dimension vector, $\boldsymbol{I}$ is an identity matrix of the same dimension and $\boldsymbol{B}^{+}$ is the Moore-Penrose Pseudoinverse of matrix $\boldsymbol{B}$.

Proof. In this proof we use the definition of the Moore-Penrose Pseudoinverse $\boldsymbol{B}^{+}$ and the fact that the matrix $\boldsymbol{B}^{+}$ is unique for a real-valued matrix $\boldsymbol{B}$. Matrix $\boldsymbol{B}^{+}$ has the following properties: (1) $\boldsymbol{B}\boldsymbol{B}^{+}\boldsymbol{B} = \boldsymbol{B}$; (2) $\boldsymbol{B}^{+}\boldsymbol{B}\boldsymbol{B}^{+} = \boldsymbol{B}^{+}$; (3) $\boldsymbol{B}^{+}\boldsymbol{B} = (\boldsymbol{B}^{+}\boldsymbol{B})'$; and (4) $\boldsymbol{B}\boldsymbol{B}^{+} = (\boldsymbol{B}\boldsymbol{B}^{+})'$. Properties (2)–(3) imply that $\boldsymbol{Q} = \boldsymbol{B}^{+}\boldsymbol{B}$ is an orthogonal projection operator, so $\boldsymbol{Q}^2 = \boldsymbol{Q}$ and $\boldsymbol{Q}' = \boldsymbol{Q}$:

$$\boldsymbol{Q}^2 = \boldsymbol{B}^{+}\boldsymbol{B}\boldsymbol{B}^{+}\boldsymbol{B} = \boldsymbol{B}^{+}\boldsymbol{B} = \boldsymbol{Q} \qquad \text{due to property (2)}$$

$$\boldsymbol{Q}' = (\boldsymbol{B}^{+}\boldsymbol{B})' = \boldsymbol{B}^{+}\boldsymbol{B} = \boldsymbol{Q} \qquad \text{due to property (3).}$$

Any vector $\boldsymbol{x}$ can be decomposed by a orthogonal $\boldsymbol{Q}$ projection as: $\boldsymbol{x} = \boldsymbol{Q}\boldsymbol{x} + (\boldsymbol{I} - \boldsymbol{Q})\boldsymbol{x}$. In our case, we have that $\boldsymbol{x} = \boldsymbol{B}^{+}\boldsymbol{B}\boldsymbol{x} + (\boldsymbol{I} - \boldsymbol{B}^{+}\boldsymbol{B})\boldsymbol{x}$. If vector $\boldsymbol{x}$ is a solution to the system

$\boldsymbol{b} = \boldsymbol{Bx}$, then it must be that:

$$\boldsymbol{Bx} = \boldsymbol{b} \Rightarrow \boldsymbol{x} = \boldsymbol{B}^{+}\boldsymbol{b} + (\boldsymbol{I} - \boldsymbol{B}^{+}\boldsymbol{B})\boldsymbol{x}$$

$$\text{Moreover } \boldsymbol{b} = \boldsymbol{Bx} \Rightarrow \boldsymbol{b} = \boldsymbol{B}\big(\boldsymbol{B}^{+}\boldsymbol{b} + \boldsymbol{B}(\boldsymbol{I} - \boldsymbol{B}^{+}\boldsymbol{B})\boldsymbol{x}\big)$$

$$\text{But: } \boldsymbol{B}(\boldsymbol{I} - \boldsymbol{B}^{+}\boldsymbol{B}) = \boldsymbol{0} \text{ due to property (4) of } \boldsymbol{B}^{+}$$

$$\text{Thus : } \boldsymbol{B}(\boldsymbol{I} - \boldsymbol{B}^{+}\boldsymbol{B})\boldsymbol{\lambda} = \boldsymbol{0} \text{ for any real valued } \boldsymbol{\lambda}$$

$$\Rightarrow \boldsymbol{b} = \boldsymbol{B}\big(\boldsymbol{B}^{+}\boldsymbol{b} + (\boldsymbol{I} - \boldsymbol{B}^{+}\boldsymbol{B})\boldsymbol{\lambda}\big)$$

$$\therefore \widetilde{\boldsymbol{x}} = \boldsymbol{B}^{+}\boldsymbol{b} + (\boldsymbol{I} - \boldsymbol{B}^{+}\boldsymbol{B})\boldsymbol{\lambda} \text{ is also a solution as } \boldsymbol{b} = \boldsymbol{B}\tilde{\boldsymbol{x}} \text{ holds.}$$

$$\text{Thus } \widetilde{\boldsymbol{x}} = \boldsymbol{B}^{+}\boldsymbol{b} + \boldsymbol{K}\boldsymbol{\lambda} \text{ such that } \boldsymbol{K} = (\boldsymbol{I} - \boldsymbol{B}^{+}\boldsymbol{B}) \text{ is a general solution.}$$

□

We now use Lemma **L-2** to prove Theorem **T-2**.

Proof of T-2:

Proof. We apply the general solution for the matrix form of a system of linear equations to Equation (19) in the text. This generates $\boldsymbol{P_S} = \boldsymbol{B}_T^{+}\boldsymbol{P_Z} + \boldsymbol{K}_T\boldsymbol{\lambda}$. By hypothesis $\xi'\boldsymbol{K}_T = 0$, and thus $\xi'\boldsymbol{P}_S = \boldsymbol{\xi}'\boldsymbol{B}_T^{+}\boldsymbol{P_Z}$, which makes $\boldsymbol{\xi}'\boldsymbol{P_S}$ identified. We can apply the same rationale to Equation (20) which identifies $\xi'\boldsymbol{L_S}$. By the same token, applying this analysis to (18), $\boldsymbol{Q_S}(t) = \boldsymbol{B}_T^{+}\boldsymbol{Q}_Z(t) + \boldsymbol{K}_t\boldsymbol{\lambda}$. Thus $\zeta'\boldsymbol{K}_t = \boldsymbol{0}$ implies that $\boldsymbol{\zeta}'\boldsymbol{Q_S}(t) = \boldsymbol{\zeta}'\boldsymbol{B}_T^{+}\boldsymbol{Q}_Z(t)$ is identified.

□

A.4 Bounds for Response-Type Probabilities and Counterfactual Outcomes

Lemma **L-3** below uses linear Equations (18)–(19) and Lemma **L-2** to generate simple bounds for response-type probabilities and counterfactual outcomes:

Lemma L-3. *For the IV model (1)–(3), bounds for response-type probabilities $\boldsymbol{P_S}$ given a response matrix $\boldsymbol{R}$ are given by:*

$$\boldsymbol{P_S} \in \left[\max\left(\boldsymbol{0}_{N_{\boldsymbol{S}}}, \boldsymbol{B}_T^{+}\boldsymbol{P}_Z + \min_{\boldsymbol{\lambda}\in\mathbb{R}^{N_{\boldsymbol{S}}}} (\boldsymbol{K}_T\boldsymbol{\lambda})\right), \min\left(\boldsymbol{\iota}_{N_{\boldsymbol{S}}}, \boldsymbol{B}_T^{+}\boldsymbol{P}_Z + \max_{\boldsymbol{\lambda}\in\mathbb{R}^{N_{\boldsymbol{S}}}} (\boldsymbol{K}_T\boldsymbol{\lambda})\right)\right], \tag{A.2a}$$

where $\boldsymbol{\lambda}$ is an arbitrary real-valued vector of dimension $N_{\boldsymbol{S}}$. Bounds on $\boldsymbol{\lambda}$ come from the fact that $P_{\boldsymbol{S}}$ is a vector with probabilities defined on the unit simplex. Bounds for the expectation of outcomes by strata are given by:

$$\left(\boldsymbol{B}_t^{+}\boldsymbol{Q}_Z(t) + \min_{\boldsymbol{\xi}\in\mathbb{R}^{N_{\boldsymbol{S}}}} (\boldsymbol{K}_t\boldsymbol{\xi})\right) \leq \boldsymbol{Q_S}(t) \leq \left(\boldsymbol{B}_t^{+}\boldsymbol{Q}_Z(t) + \max_{\boldsymbol{\xi}\in\mathbb{R}^{N_{\boldsymbol{S}}}} (\boldsymbol{K}_t\boldsymbol{\xi})\right).[2] \tag{A.2b}$$

where $\boldsymbol{\xi}$ is an arbitrary real-valued vector of dimension $N_{\boldsymbol{S}}$.

Proof. Equations (A.2a) and (A.2b) follow directly from the application of the general linear solution (A.1) of Lemma **L-2** to the system of linear equations of Equations (19) and (18) respectively. The admissible ranges of λ in Equation (A.2a) comes from using the fact that $\boldsymbol{P_S}$ are probabilities. □

A.5 Proof of Corollary C-1:

Proof. According to **T-2**, Vectors $\boldsymbol{P}_S$ and $\boldsymbol{L}_S$ are point-identified if and only if $\boldsymbol{\xi}'\boldsymbol{K}_T = \boldsymbol{0}$ for any $\boldsymbol{\xi}'$. Thus it must be the case that $\boldsymbol{K}_T = \boldsymbol{0}$. Since $\boldsymbol{K}_T = (\boldsymbol{I}_{N_{\boldsymbol{S}}} - \boldsymbol{B}_T^{+}\boldsymbol{B}_T)$, $\boldsymbol{K}_T = \boldsymbol{0}$ if and only if $\boldsymbol{I}_{N_{\boldsymbol{S}}} = \boldsymbol{B}_T^{+}\boldsymbol{B}_T$ which holds if and only if rank$(\boldsymbol{B}_T) = N_{\boldsymbol{S}}$, that is, $\boldsymbol{B}_T$ has full column-rank. From Theorem **T-2**, $\boldsymbol{P_S}$ is identified from $\boldsymbol{B}_T^{+}\boldsymbol{P_Z}$ if and only if rank$(\boldsymbol{B}_T) = N_{\boldsymbol{S}}$. The second equation follows from the same rationale. $\boldsymbol{K}_t = \boldsymbol{0}$ if and only if rank$(\boldsymbol{B}_t) = N_{\boldsymbol{S}}$. According to Theorem **T-2**, if $\boldsymbol{K}_t = \boldsymbol{0}$, then $\boldsymbol{Q_S}(t) = \boldsymbol{B}_t^{+}\boldsymbol{Q}_Z(t)$, and thereby $E(\kappa(Y(t)))$

[2] These bounds are not sharp because we do not use the full distribution of the data generating process in constructing them.

can be expressed as:

$$\begin{aligned} E(\kappa(Y(t))) &= \sum_{n=1}^{N_{\boldsymbol{S}}} E(\kappa(Y(t))|\boldsymbol{S}=\boldsymbol{s}_n)P(\boldsymbol{S}=\boldsymbol{s}_n) \\ &= \iota'_{N_{\boldsymbol{S}}}\boldsymbol{Q}_{\boldsymbol{S}}(t), \\ &= \iota'_{N_{\boldsymbol{S}}}\boldsymbol{B}_t^{+}\boldsymbol{Q}_Z(t), \end{aligned}$$

where $\iota_{N_{\boldsymbol{S}}}$ is a $N_{\boldsymbol{S}}$-dimensional vector of $1s$. □

A.6 Proof of Theorem T-3

We first establish a series of lemmas and then turn to the main proof.[3]

A.6.1 Lemma L-4

Lemma L-4. Every sub-matrix of a lonesum matrix is lonesum.

Proof. Suppose that there is a binary matrix $\boldsymbol{B}$ whose sub-matrix $\tilde{\boldsymbol{B}}$ is not lonesum. Thus $\tilde{\boldsymbol{B}}$ cannot be uniquely recovered by its row and column sums. This fact is not altered if all elements in $\boldsymbol{B}$ other than $\tilde{\boldsymbol{B}}$ were known. In particular, $\boldsymbol{B}$ cannot be lonesum. □

A.6.2 Lemma L-5

Lemma L-5. If a binary matrix is lonesum, then no 2×2 sub-matrix takes the form of the prohibited patterns (52), that is,

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \text{nor} \quad \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Proof. We first prove that the prohibited patterns are **not** lonesum. Consider a matrix 2×2 binary matrix $\boldsymbol{B}$ whose column-sums and row-sums are equal to 1. Matrix $\boldsymbol{B}$ can be equal to either $\boldsymbol{B}_1$ or $\boldsymbol{B}_2$ (defined below). Indeed the column-sums and row-sums of both $\boldsymbol{B}_1$ and

[3] Lemmas L-4–L-8 provide simple proofs of the properties of binary matrices used in this paper. For an extensive discussion of binary matrices, see Brualdi (1980); Brualdi and Ryser (1991); Ryser (1957); Sachnov and Tarakanov (2002).

$\boldsymbol{B}_2$ are equal to 1.

$$\boldsymbol{B}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \text{and} \quad \boldsymbol{B}_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

As a consequence, any 2×2 binary matrix $\boldsymbol{B}$ whose column-sums and row-sums are equal to 1 cannot be uniquely identified based only on its column and row sums, i.e., $\boldsymbol{B}$ is not lonesum. In particular, $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$ are **not** lonesum. But $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$ exhibit the prohibited pattern (52). Thus the prohibited patterns $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$ are **not** lonesum.

Now according to Lemma **L-4**, all sub-matrices of a lonesum matrix are also lonesum. But $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$ are **not** lonesum. Thus **no** 2×2 sub-matrix of a lonesum matrix can be equal to either of the prohibited patterns $\boldsymbol{B}_1$ or $\boldsymbol{B}_2$.

□

A.6.3 Lemma L-6

Lemma L-6. If no 2×2 sub-matrix of a binary matrix $\boldsymbol{B}$ takes the form of the prohibited patterns (52) then $\boldsymbol{B}$ is equivalent to its maximal.

Proof. The proof of the Lemma is done by proving its contrapositive form, that is, if $\boldsymbol{B}$ is **not** equivalent to its maximal, then prohibited patterns (52) **must** arise. Without loss of generality, let the columns of $\boldsymbol{B}$ be ordered in decreasing column sum. Suppose $\boldsymbol{B}$ is not maximal. Then there must exist a row i whose element of the j^{th} column is 0 followed by the element 1 in column $j+1$. But the j^{th} column sum is greater or equal than the column sum of $j+1$. Thus, there must exist at least one row i' whose j^{th} column is 1 followed by the element 0 in column $j+1$. This generates the prohibited pattern of Lemma **L-5**. □

A.6.4 Lemma L-7

Lemma L-7. If a binary matrix $\boldsymbol{B}$ is equivalent to its maximal, then its maximal can be generated by reordering its columns in decreasing column sum.

Proof. The maximal of $\boldsymbol{B}$ is a matrix whose rows present a sequence of elements 1 followed by elements 0. Thereby the maximal of $\boldsymbol{B}$ has decreasing column sums. Thus, it suffices to prove that matrix $\widetilde{\boldsymbol{B}}$, generated by permuting the $\boldsymbol{B}$-columns in decreasing column sum, is unique. Suppose it is not, then there must exist two distinct columns, say $\boldsymbol{B}[\cdot, j]$ and $\boldsymbol{B}[\cdot, j']$ of same column sum. Then it must be the case that there exist two rows i, i' such that $\boldsymbol{B}[i, j] = 1, \boldsymbol{B}[i, j'] = 0$ and $\boldsymbol{B}[i', j] = 0, \boldsymbol{B}[i', j'] = 1$. This is the prohibited pattern. Thus, there is no column permutation in which *both* rows i and i' are formed by a sequence of elements 1 followed by a sequence of elements 0. □

A.6.5 Lemma L-8

Lemma L-8. If binary matrix $\boldsymbol{B}$ is equivalent to its maximal, then $\boldsymbol{B}$ is lonesum.

Proof. Let matrix $\widetilde{\boldsymbol{B}}$ be generated by permuting the $\boldsymbol{B}$-columns in decreasing column sum. By lemma L-7, $\widetilde{\boldsymbol{B}}$ is the maximal of $\boldsymbol{B}$ and is uniquely determined by its row sums. But $\widetilde{\boldsymbol{B}}$ was generated using the column sums of $\boldsymbol{B}$. Thus $\boldsymbol{B}$ is uniquely determined by its row sums and column sums and thereby $\boldsymbol{B}$ is lonesum. □

Remark A.1. *The cyclical property of Lemmas **L-5**–**L-8** implies that the following statements are equivalent: (1) $\boldsymbol{B}$ is lonesum; (2) $\boldsymbol{B}$ has no 2×2 sub-matrix with the prohibited patterns (52); (3) $\boldsymbol{B}$ is equivalent to its maximal. This fact is exploited in the next lemma.*

A.6.6 Lemma L-9

Lemma L-9. Let there be a binary matrix $\boldsymbol{B}$ where $\boldsymbol{B}[i, j]; i \in \{1, \dots, N_r\}, j \in \{1, \dots, N_c\}$. If $\boldsymbol{B}$ is lonesum then items 1 and 2 below hold.

1. For any $j, j' \in \{1, \dots, N_c\}$, we have that:

   $$\boldsymbol{B}[i, j] \leq \boldsymbol{B}[i, j'] \text{ for all } i \in \{1, \dots, N_r\} \text{ or } \boldsymbol{B}[i, j] \geq \boldsymbol{B}[i, j'] \text{ for all } i \in \{1, \dots, N_r\}.$$

2. For any $i, i' \in \{1, \dots, N_r\}$, we have that:

   $$\boldsymbol{B}[i, j] \leq \boldsymbol{B}[i', j] \text{ for all } j \in \{1, \dots, N_c\} \text{ or } \boldsymbol{B}[i, j] \geq \boldsymbol{B}[i', j] \text{ for all } j \in \{1, \dots, N_c\}.$$

Proof. We use proof by contradiction.

Suppose that Condition 1 does not hold. The negation of Condition 1 is that: For some $j, j' \in \{1, \ldots, N_c\}$, there exists some $i, i' \in \{1, \ldots, N_r\}$ such that $\boldsymbol{B}[i, j] > \boldsymbol{B}[i, j']$ and $\boldsymbol{B}[i', j] < \boldsymbol{B}[i', j']$. Thus it must be the case that $\boldsymbol{B}[i, j] = \boldsymbol{B}[i', j'] = 1$ and $\boldsymbol{B}[i, j'] = \boldsymbol{B}[i', j] = 0$. Thus the 2×2 sub-matrix of $\boldsymbol{B}$ generated by rows i, i' and columns j, j' is given by:

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

which is a prohibited pattern and, according to Lemma **L-5**, $\boldsymbol{B}$ cannot be lonesum.

The proof of Condition 2 follows the same rationale of the proof of Condition 1. Suppose that Condition 2 does not hold. The negation of Condition 2 states that: For some $i, i' \in \{1, \ldots, N_r\}$, there exists some $j, j' \in \{1, \ldots, N_c\}$ such that $\boldsymbol{B}[i, j] > \boldsymbol{B}[i', j]$ and $\boldsymbol{B}[i, j'] < \boldsymbol{B}[i', j']$.

Thus it must be the case that $\boldsymbol{B}[i, j] = \boldsymbol{B}[i', j'] = 1$ and $\boldsymbol{B}[i, j'] = \boldsymbol{B}[i', j] = 0$, which generates a prohibited pattern (see Item 1).

□

A.6.7 Lemma L-10

Lemma L-10. Consider a lonesum binary matrix $\boldsymbol{B}$. Let $\tilde{\boldsymbol{B}}$ be matrix generated by ordering the rows of $\boldsymbol{B}$ in increasing row-sum and ordering the columns of $\boldsymbol{B}$ in decreasing column-sum. Then $\tilde{\boldsymbol{B}}$ is lower triangular and:

1. Any row in $\tilde{\boldsymbol{B}}$ that has both elements 1 and 0 must be a sequence of elements 1 followed by a sequence of elements 0.

2. Any column of $\tilde{\boldsymbol{B}}$ that has both elements 1 and 0 must be a sequence of elements 0 followed by a sequence of elements 1.

Proof. $\boldsymbol{B}$ is lonesum, then by Remark A.1, $\boldsymbol{B}$ is equivalent to its maximal. The maximal of $\boldsymbol{B}$ consists of the matrix $\tilde{\boldsymbol{B}}$ of the same dimension as $\boldsymbol{B}$ whose rows share the same sum of the rows in $\boldsymbol{B}$ and those rows consist of a sequence of elements 1 followed by a sequence of elements 0. Moreover, $\tilde{\boldsymbol{B}}$ can be obtained by ordering the columns of $\boldsymbol{B}$ in decreasing column-sum. Note also that if $\boldsymbol{B}$ is lonesum, its transpose, that is, $\boldsymbol{B}'$ is also lonesum and the maximal of $\boldsymbol{B}'$ can be obtained by ordering the columns of $\boldsymbol{B}'$ in decreasing order. A consequence of this fact is that if the rows of $\boldsymbol{B}$ is ordered in increasing row-sum, then each column consists of sequence of elements 0 followed by a sequence of elements 1. □

A.6.8 Lemma L-11

Lemma L-11. Let a binary matrix $\tilde{\boldsymbol{B}}$ be lower triangular, that is,

1. Any row in $\tilde{\boldsymbol{B}}$ that has both elements 1 and 0 must be a sequence of elements 1 followed by a sequence of elements 0.

2. Any column of $\tilde{\boldsymbol{B}}$ that has both elements 1 and 0 must be a sequence of elements 0 followed by a sequence of elements 1.

Also, let $r(i) = \sum_{j=1}^{N_c} \tilde{\boldsymbol{B}}[i,j]$ (row-sum) and $c(j) = \sum_{i=1}^{N_r} \tilde{\boldsymbol{B}}[i,j]$ (column-sum) where $i \in \{1,\ldots,N_r\}$ and $j \in \{1,\ldots,N_c\}$. If $\tilde{\boldsymbol{B}}$ has strictly positive column and row sums, then it must be the case that:

$$\tilde{\boldsymbol{B}}[i,j] = 0 \Leftrightarrow \left(\sum_{j'=1}^{N_c} \mathbf{1}\Big[c(j) \leq c(j')\Big]\right) > r(i)$$
$$\tilde{\boldsymbol{B}}[i,j] = 1 \Leftrightarrow \left(\sum_{j'=1}^{N_c} \mathbf{1}\Big[c(j) \leq c(j')\Big]\right) \leq r(i)$$

Proof. For any given row i, there is a $\tilde{j} \in \{1,\ldots,N_c\}$ such that $r(i) = \tilde{j}$. By the lower triangular property of $\tilde{\boldsymbol{B}}$, we must have that:

$$c(1) \geq \cdots \geq c(\tilde{j}-1) \geq c(\tilde{j}) > c(\tilde{j}+1) \geq \ldots \geq c(N_c).$$

Thus we can write $r(i)$ as:

$$r(i) = \tilde{j} = \sum_{j'=1}^{N_c} \mathbf{1}\Big[c(\tilde{j}) \leq c(j')\Big], \tag{A.3}$$

and it must be case that the Inequalities (A.4)–(A.5) below hold:

$$r(i) = \tilde{j} < \sum_{j'=1}^{N_c} \mathbf{1}\Big[c(j) \le c(j')\Big], \text{ for all } j \text{ such that } j > \tilde{j}. \tag{A.4}$$

$$r(i) = \tilde{j} \ge \sum_{j'=1}^{N_c} \mathbf{1}\Big[c(j) \le c(j')\Big], \text{ for all } j \text{ such that } j \le \tilde{j}. \tag{A.5}$$

Thus we have that:

$$\tilde{\boldsymbol{B}}[i,j] = 0 \Leftrightarrow j > \tilde{j} \Leftrightarrow \sum_{j'=1}^{N_c} \mathbf{1}\Big[c(j) \le c(j')\Big] > r(i) = \tilde{j},$$

where the first implication comes from the row property of the lower triangular matrix $\tilde{\boldsymbol{B}}$ and the definition of $\tilde{j}$. The second implication arises from Inequality (A.4).

Also:

$$\tilde{\boldsymbol{B}}[i,j] = 1 \Leftrightarrow j \le \tilde{j} \Leftrightarrow \sum_{j'=1}^{N_c} \mathbf{1}\Big[c(j) \le c(j')\Big] \le r(i) = \tilde{j},$$

where the first implication comes from the row property of the lower triangular matrix $\tilde{\boldsymbol{B}}$ and the definition of $\tilde{j}$. The second implication arises from Inequality (A.5).

□

A.6.9 Lemma L-12

Lemma L-12 gives a formula that characterizes each element $\boldsymbol{B}[i,j]$ of a lonemsum matrix in terms of its row-sum $\sum_{j'=1}^{N_c} \boldsymbol{B}[i,j']$ and the column-sum $\sum_{i'=1}^{N_r} \boldsymbol{B}[i',j']$; $\quad i' \in \{1,\cdots,N_r\}$, $j' \in \{1,\ldots,N_c\}$.

Lemma L-12. If a binary matrix $\boldsymbol{B}$ is lonesum with strictly positive row and column sums, then each element $\boldsymbol{B}[i,j]; i \in \{1,\ldots,N_r\}, j \in \{1,\ldots,N_c\}$ can be expressed as:

$$\boldsymbol{B}[i,j] = \mathbf{1}\Bigg[\underbrace{\left(\sum_{j'=1}^{N_c} \mathbf{1}\left[\sum_{i'=1}^{N_r} \boldsymbol{B}[i',j] \le \sum_{i'=1}^{N_r} \boldsymbol{B}[i',j']\right]\right)}_{\text{Number of columns whose sum is bigger than column-sum of } \boldsymbol{B}[\cdot,j]} \le \underbrace{\left(\sum_{j'=1}^{N_c} \boldsymbol{B}[i,j']\right)}_{\text{row-sum of } \boldsymbol{B}_t[i,\cdot]}\Bigg]. \tag{A.6}$$

Proof. Let $\tilde{\boldsymbol{B}}$ be matrix generated by ordering the rows of $\boldsymbol{B}$ in increasing row-sum and ordering the columns of $\boldsymbol{B}$ in decreasing column-sum. Thus by Lemma L-10, $\tilde{\boldsymbol{B}}$ is lower triangular such that:

1. Any row in $\tilde{\boldsymbol{B}}$ that has both elements 1 and 0 must be a sequence of elements 1 followed by a sequence of elements 0.

2. Any column of $\tilde{\boldsymbol{B}}$ that has both elements 1 and 0 must be a sequence of elements 0 followed by a sequence of elements 1.

Thus by Lemma L-11, it must the the case that the following inequalities hold:

$$\tilde{\boldsymbol{B}}[i,j] = 0 \Leftrightarrow \left(\sum_{j'=1}^{N_c} \mathbf{1}\Big[c(j) \leq c(j')\Big]\right) > r(i)$$

$$\tilde{\boldsymbol{B}}[i,j] = 1 \Leftrightarrow \left(\sum_{j'=1}^{N_c} \mathbf{1}\Big[c(j) \leq c(j')\Big]\right) \leq r(i)$$

Therefore we can express $\tilde{\boldsymbol{B}}[i,j]$ by the following expression:

$$\tilde{\boldsymbol{B}}[i,j] = \mathbf{1}\left[\left(\sum_{j'=1}^{N_c} \mathbf{1}\Big[c(j) \leq c(j')\Big]\right) \leq \Big(r(i)\Big)\right] \tag{A.7}$$

Equation (A.7) only depends on the column and row sums of $\tilde{\boldsymbol{B}}$. But any row-permutation of $\tilde{\boldsymbol{B}}$ does not change its column sum. Moreover, any column-permutation of $\tilde{\boldsymbol{B}}$ does not change its row sum. Thus Equation (A.7) also holds for any row or column permutation of $\tilde{\boldsymbol{B}}$. In other words, Equation (A.7) holds for any matrix that is equivalent to $\tilde{\boldsymbol{B}}$. In particular, Equation (A.7) holds for $\tilde{\boldsymbol{B}}$. The proof is completed by acknowledging that (A.7) is the same as Equation (A.6). □

A.6.10 Lemma L-13

Lemma L-13. Let $\boldsymbol{B}$ be a binary matrix where $\boldsymbol{B}[i,j]; i \in \{1,\ldots,N_r\}, j \in \{1,\ldots,N_c\}$ and $\sigma_1,\ldots,\sigma_{N_c}$ be a sequence of strictly positive numbers. If $\boldsymbol{B}$ is lonesum then:

$$\begin{aligned}
&\mathbf{1}\left[\left(\sum_{j'=1}^{N_c} \mathbf{1}\left[\sum_{i'=1}^{N_r} \boldsymbol{B}[i',j] \leq \sum_{i'=1}^{N_r} \boldsymbol{B}[i',j']\right]\right) \leq \left(\sum_{j'=1}^{N_c} \boldsymbol{B}[i,j']\right)\right] \\
&= \mathbf{1}\left[\left(\sum_{j'=1}^{N_c} \sigma_{j'} \cdot \mathbf{1}\left[\sum_{i'=1}^{N_r} \boldsymbol{B}[i',j] \leq \sum_{i'=1}^{N_r} \boldsymbol{B}[i',j']\right]\right) \leq \left(\sum_{j'=1}^{N_c} \sigma_{j'} \cdot \boldsymbol{B}[i,j']\right)\right].
\end{aligned} \tag{A.8}$$

for any $(i, j) \in \{1, \ldots, N_r\} \times \{1, \ldots, N_c\}$.

Proof. By Lemma **L-9**, we have that for any columns $j, j'\{1, \ldots, N_c\}$, $\boldsymbol{B}[i', j] \leq \boldsymbol{B}[i', j']$ or $\boldsymbol{B}[i', j] \geq \boldsymbol{B}[i', j']$ for all rows $i' \in \{1, \ldots, N_r\}$. As a shorthand notation, let $c(j) = \sum_{i=1}^{N_r} \boldsymbol{B}[i, j]$ be the column sum. In this notation, Equation (A.8) can be rewritten as:

$$\mathbf{1}\left[\left(\sum_{j'=1}^{N_c} \mathbf{1}\left[c(j) \leq c(j')\right]\right) \leq \left(\sum_{j'=1}^{N_c} \boldsymbol{B}[i, j']\right)\right] = \mathbf{1}\left[\left(\sum_{j'=1}^{N_c} \sigma_{j'} \mathbf{1}\left[c(j) \leq c(j')\right]\right) \leq \left(\sum_{j'=1}^{N_c} \sigma_{j'} \boldsymbol{B}[i, j']\right)\right].$$

We want to prove that:

$$\begin{aligned}
&\mathbf{1}\left[\left(\sum_{j'=1}^{N_c} \mathbf{1}\left[c(j) \leq c(j')\right]\right) \leq \left(\sum_{j'=1}^{N_c} \boldsymbol{B}[i, j']\right)\right] = 1 \Rightarrow \\
&\mathbf{1}\left[\left(\sum_{j'=1}^{N_c} \sigma_{j'} \mathbf{1}\left[c(j) \leq c(j')\right]\right) \leq \left(\sum_{j'=1}^{N_c} \sigma_{j'} \boldsymbol{B}[i, j']\right)\right] = 1 \\
\text{and } &\mathbf{1}\left[\left(\sum_{j'=1}^{N_c} \mathbf{1}\left[c(j) \leq c(j')\right]\right) \leq \left(\sum_{j'=1}^{N_c} \boldsymbol{B}[i, j']\right)\right] = 0 \Rightarrow \\
&\mathbf{1}\left[\left(\sum_{j'=1}^{N_c} \sigma_{j'} \mathbf{1}\left[c(j) \leq c(j')\right]\right) \leq \left(\sum_{j'=1}^{N_c} \sigma_{j'} \boldsymbol{B}[i, j']\right)\right] = 0.
\end{aligned}$$

Consider the first case:

$$\mathbf{1}\left[\left(\sum_{j'=1}^{N_c} \mathbf{1}\left[c(j) \leq c(j')\right]\right) \leq \left(\sum_{j'=1}^{N_c} \boldsymbol{B}[i, j']\right)\right] = 1 \Rightarrow \boldsymbol{B}[i, j] = 1 \text{ by Lemma } \textbf{L-12}.$$

$$\begin{aligned}
\text{Then, } &\mathbf{1}\left[c(j) \leq c(j')\right] = 1 \Rightarrow \boldsymbol{B}[i', j] \leq \boldsymbol{B}[i', j'] \ \forall\ i' \in \{1, \ldots, N_r\}. \\
&\text{But, } \boldsymbol{B}[i, j] = 1 \text{ therefore it must be that } \boldsymbol{B}[i, j'] = 1. \\
\text{Thus, } &\mathbf{1}\left[c(j) \leq c(j')\right] = 1 \Rightarrow \boldsymbol{B}[i, j'] = 1. \\
\Rightarrow &\mathbf{1}\left[c(j) \leq c(j')\right] \leq \boldsymbol{B}[i, j'] \ \forall\ j' \in \{1, \ldots, N_c\}. \\
\Rightarrow &\sigma_{j'} \mathbf{1}\left[c(j) \leq c(j')\right] \leq \sigma_{j'} \boldsymbol{B}[i, j'] \ \forall\ j' \in \{1, \ldots, N_c\}. \\
\Rightarrow &\mathbf{1}\left[\sum_{j'=1}^{N_c} \sigma_{j'} \mathbf{1}\left[c(j) \leq c(j')\right] \leq \sum_{j'=1}^{N_c} \sigma_{j'} \boldsymbol{B}[i, j']\right] = 1.
\end{aligned}$$

Consider the second case:

$$\mathbf{1}\left[\left(\sum_{j'=1}^{N_c}\mathbf{1}\left[c(j)\leq c(j')\right]\right)\leq\left(\sum_{j'=1}^{N_c}\boldsymbol{B}[i,j']\right)\right]=0\Rightarrow\boldsymbol{B}[i,j]=0 \text{ by Lemma L-12}$$

But if $\boldsymbol{B}[i,j']=1$ and $\boldsymbol{B}[i,j]=0$ then $\boldsymbol{B}[i',j]\leq\boldsymbol{B}[i',j']\ \forall\ i'\in\{1,\ldots,N_r\}$

Thus, $\boldsymbol{B}[i,j']=1\Rightarrow\mathbf{1}\left[c(j)\leq c(j')\right]=1$

$$\Rightarrow\mathbf{1}\left[c(j)\leq c(j')\right]\geq\boldsymbol{B}[i,j']\ \forall\ j'\in\{1,\ldots,N_c\}$$

$$\Rightarrow\sigma_{j'}\mathbf{1}\left[c(j)\leq c(j')\right]\geq\sigma_{j'}\boldsymbol{B}[i,j']\ \forall\ j'\in\{1,\ldots,N_c\}$$

But $\left(\sum_{j'=1}^{N_c}\mathbf{1}\left[c(j)\leq c(j')\right]\right)>\left(\sum_{j'=1}^{N_c}\boldsymbol{B}[i,j']\right)$

So $\exists\ j'$ such that $\mathbf{1}\left[c(j)\leq c(j')\right]>\boldsymbol{B}[i,j']$

$$\Rightarrow\sum_{j'=1}^{N_c}\sigma_{j'}\mathbf{1}\left[c(j)\leq c(j')\right]>\sum_{j'=1}^{N_c}\sigma_{j'}\boldsymbol{B}[i,j']$$

$$\Rightarrow\mathbf{1}\left[\sum_{j'=1}^{N_c}\sigma_{j'}\mathbf{1}\left[c(j)\leq c(j')\right]\leq\sum_{j'=1}^{N_c}\sigma_{j'}\boldsymbol{B}[i,j']\right]=0$$

□

A.6.11 Lemma L-14

Lemma L-14. Suppose $\boldsymbol{B}$ is a binary matrix where $\boldsymbol{B}[i,j]; i\in\{1,\ldots,N_r\}, j\in\{1,\ldots,N_c\}$. Define a sequence of strictly positive numbers $\zeta_1,\ldots,\zeta_{N_r}$. If $\boldsymbol{B}$ is lonesum, then:

$$\mathbf{1}\left[\sum_{i'=1}^{N_r}\boldsymbol{B}[i',j]\leq\sum_{i'=1}^{N_r}\boldsymbol{B}[i',j']\right]=\mathbf{1}\left[\sum_{i'=1}^{N_r}\zeta_{i'}\boldsymbol{B}[i',j]\leq\sum_{i'=1}^{N_r}\zeta_{i'}\boldsymbol{B}[i',j']\right]. \tag{A.9}$$

Proof. By Lemma L-9, we have that for any columns $j,j'\in\{1,\ldots,N_c\}$, $\boldsymbol{B}[i',j]\leq\boldsymbol{B}[i',j']$

or $\boldsymbol{B}[i',j] \geq \boldsymbol{B}[i',j']$ for all rows $i' \in \{1,\ldots,N_r\}$. Thus:

$$\begin{aligned}
\text{Suppose } \mathbf{1}\left[\sum_{i'=1}^{N_r}\boldsymbol{B}[i',j] \leq \sum_{i'=1}^{N_r}\boldsymbol{B}[i',j']\right] &= 1 \\
&\Rightarrow \boldsymbol{B}[i',j] \leq \boldsymbol{B}[i',j'] \text{ for all } i' \in \{1,\ldots,N_c\}, \\
&\Rightarrow \zeta_{i'}\boldsymbol{B}[i',j] \leq \zeta_{i'}\boldsymbol{B}[i',j'] \text{ for all } i' \in \{1,\ldots,N_c\}, \\
&\Rightarrow \mathbf{1}\left[\sum_{i'=1}^{N_r}\zeta_{i'}\boldsymbol{B}[i',j] \leq \sum_{i'=1}^{N_r}\zeta_{i'}\boldsymbol{B}[i',j']\right] = 1. \\
\text{Now suppose } \mathbf{1}\left[\sum_{i'=1}^{N_r}\boldsymbol{B}[i',j] \leq \sum_{i'=1}^{N_r}\boldsymbol{B}[i',j']\right] &= 0 \\
&\Rightarrow \boldsymbol{B}[i',j] \geq \boldsymbol{B}[i',j'] \text{ for all } i' \in \{1,\ldots,N_c\} \text{ and } \exists i; \boldsymbol{B}[i,j] > \boldsymbol{B}[i,j'] \\
&\Rightarrow \zeta_{i'}\boldsymbol{B}[i',j] \geq \zeta_{i'}\boldsymbol{B}[i',j'] \text{ for all } i' \in \{1,\ldots,N_c\} \text{ and } \exists i; \zeta_i\boldsymbol{B}[i,j] > \zeta_i\boldsymbol{B}[i,j'] \\
&\Rightarrow \mathbf{1}\left[\sum_{i'=1}^{N_r}\zeta_{i'}\boldsymbol{B}[i',j] \leq \sum_{i'=1}^{N_r}\zeta_{i'}\boldsymbol{B}[i',j']\right] = 0.
\end{aligned}$$

□

A.6.12 Proof of Theorem T-3

Proof. The equivalence proof of Theorem **T-3** must show that items (i)–(iv) cyclically imply each other. We use Lemmas **L-4**–**L-14** to do so. We divide the proof into a few steps:

1. The first step explores the lonesum properties of the binary matrices $\boldsymbol{B}_t$ generated by the response matrix $\boldsymbol{R}$. We use Lemmas **L-5**–**L-8** to show that (i) ⇔ (ii).

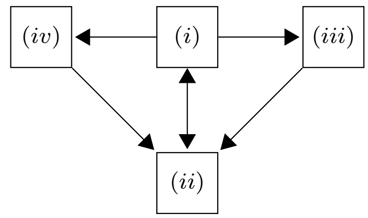
2. We use Lemma **L-9** to prove (i) ⇒ (iii) ⇒ (ii).

3. We use Lemmas **L-12**–**L-14** to prove (i) ⇒ (iv).

4. The last step of our proof is to show that (iv) ⇒ (ii).

The proof strategy can be represented by the following graph:

Figure A.1: **Schematics of the Proof of Theorem T-3**



(i) ⇔ (ii): The direct implication is a consequence of Lemma **L-5**. If each $\boldsymbol{B}_t$ is lonesum, then **no** 2×2 sub-matrix of $\boldsymbol{B}_t$ takes the form of the forbidden patterns (52). This means that no 2×2 sub-matrix of $\boldsymbol{B}_t$ takes the form:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \text{nor} \quad \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

But $\boldsymbol{B}_t = \mathbf{1}[R = t]$. Thus **no** 2×2 sub-matrix of $\boldsymbol{R}$ takes the forms:

$$\begin{pmatrix} t & t' \\ t'' & t \end{pmatrix} \text{ or } \begin{pmatrix} t' & t \\ t & t'' \end{pmatrix}, \text{where } t' \neq t \text{ and } t'' \neq t. \tag{A.10}$$

The reverse implication is a consequence of the equivalence between lonesum matrices and the forbidden patterns that was proved by the cyclical implication of the Lemmas **L-5** ⇒ **L-6** ⇒ **L-8** ⇒ **L-5**. Specifically, if (A.10) holds, then **no** 2×2 sub-matrices of $\boldsymbol{B}_t; t \in \text{supp}(T)$ takes the form of the forbidden patterns (52). Thus each $\boldsymbol{B}_t; t \in \text{supp}(T)$ is lonesum.

(i) ⇒ (iii) ⇒ (ii) : If each $\boldsymbol{B}_t; t \in \text{supp}(T)$ is lonesum, then, according to Lemma **L-9**, any two row-indexes i, i' of $\boldsymbol{B}_t$ must satisfy:

$$\boldsymbol{B}_t[i,j] \leq \boldsymbol{B}_t[i',j] \text{ or } \boldsymbol{B}_t[i,j] \geq \boldsymbol{B}_t[i',j] \text{ for all column-indexes } j. \tag{A.11}$$

But $\boldsymbol{B}_t = \mathbf{1}[\boldsymbol{R} = t]$. Rewriting Equation (A.11) in terms of the response matrix $\boldsymbol{R}$ generates:

$$\big(\mathbf{1}[T = t]|Z = z_i, \boldsymbol{S} = \boldsymbol{s}_j\big) \leq \big(\mathbf{1}[T = t]|Z = z_{i'}, \boldsymbol{S} = \boldsymbol{s}_j\big) \text{ or} \tag{A.12}$$

$$\big(\mathbf{1}[T = t]|Z = z_i, \boldsymbol{S} = \boldsymbol{s}_j\big) \geq \big(\mathbf{1}[T = t]|Z = z_{i'}, \boldsymbol{S} = \boldsymbol{s}_j\big) \text{ occurs} \tag{A.13}$$

for $\forall\, \boldsymbol{s}_j \in \text{supp}(\boldsymbol{S})$ and any $z_i, z_{i'} \in \text{supp}(Z)$.

Since $\boldsymbol{S}$ is a balancing score for $\boldsymbol{V}$, for every $\boldsymbol{v} \in \text{supp}(\boldsymbol{V})$, there exists a unique $\boldsymbol{s}_j \in \text{supp}(\boldsymbol{S})$ such that $f_{\boldsymbol{S}}(\boldsymbol{v}) = \boldsymbol{s}_j$ (see Section 3.1). Thereby, for any $z_i \in supp(Z)$ and $t \in \text{supp}(T)$ we have that:

$$\big(\mathbf{1}[T = t]|Z = z_i, \boldsymbol{V} = \boldsymbol{v}\big) = \big(\mathbf{1}[T = t]|Z = z_i, \boldsymbol{S} = f_{\boldsymbol{S}}(\boldsymbol{v})\big) = \big(\mathbf{1}[T = t]|Z = z_i, \boldsymbol{S} = \boldsymbol{s}_j\big) = \mathbf{B}_t[i, j]. \tag{A.14}$$

Using Equation (A.14) in (A.12) and (A.13) allows us to write that for any $z_i, z_{i'} \in \text{supp}(Z)$ it must be the case that:

$$\big(\mathbf{1}[T = t]|Z = z_i, \boldsymbol{V} = \boldsymbol{v}\big) \leq \big(\mathbf{1}[T = t]|Z = z_{i'}, \boldsymbol{V} = \boldsymbol{v}\big) \text{ for all values } \boldsymbol{v} \in \text{supp}(\boldsymbol{V}), \tag{A.15}$$

$$\text{or } \big(\mathbf{1}[T = t]|Z = z_i, \boldsymbol{V} = \boldsymbol{v}\big) \geq \big(\mathbf{1}[T = t]|Z = z_{i'}, \boldsymbol{V} = \boldsymbol{v})\big) \text{ for all values } \boldsymbol{v} \in \text{supp}(\boldsymbol{V}). \tag{A.16}$$

For each agent $\omega \in \Omega$, that is a unique value $\boldsymbol{v} \in \text{supp}(\boldsymbol{V})$ such that $\boldsymbol{V}_\omega = \boldsymbol{v}$. Therefore we can express the indicator function for the choice conditional on $\boldsymbol{V}$ and Z, that is, $\mathbf{1}[T = t|\boldsymbol{V} = \boldsymbol{v}, Z = z]$, as the indicator function of the counterfactual choice $\mathbf{1}[T_\omega(z) = t]$ for $\boldsymbol{V}_\omega = \boldsymbol{v}$. Thus we can restate Equations (A.15)–(A.16) as (A.17)–(A.18) below.

$$\mathbf{1}[T_\omega(z_i) = t] \leq \mathbf{1}[T_\omega(z_{i'}) = t] \text{ for all } \omega \in \Omega, \tag{A.17}$$

or

$$\text{or } \mathbf{1}[T_\omega(z_i) = t] \geq \mathbf{1}[T_\omega(z_{i'}) = t] \text{ for all } \omega \in \Omega. \tag{A.18}$$

Now suppose that for $z, z' \in \text{supp}(Z)$ and $t, t' \in \text{supp}(T)$ such that $t \neq t'$, there exists some $\omega \in \Omega$ such that $T_\omega(z') = t'$ and $T_\omega(z) = t$. Thus as the instrument changes from

z' to z, agent ω is induced to choose t. If (A.17)–(A.18) hold, then it **cannot** be the case that there exists an agent $w' \in \Omega$ that is induced to choose t as the instrument change from z to z'. In other words, it cannot be the case that $T_{\omega'}(z) = t''$ and $T_{\omega'}(z) = t$, such that $t \neq t'' \in \text{supp}(T)$. Let ω be associated with response-type $\boldsymbol{s}$ and ω' with $\boldsymbol{s}'$. Thus the following pattern cannot occur:

$$\begin{bmatrix} (T|Z=z, \boldsymbol{S}=\boldsymbol{s}) & (T|Z=z, \boldsymbol{S}=\boldsymbol{s}') \\ (T|Z=z', \boldsymbol{S}=\boldsymbol{s}) & (T|Z=z', \boldsymbol{S}=\boldsymbol{s}') \end{bmatrix} = \begin{bmatrix} t & t'' \\ t' & t \end{bmatrix}, \textit{where } t' \neq t \textit{ and } t'' \neq t. \tag{A.19}$$

Equation (A.19) implies item (ii).

(i) ⇒ (iv): If $\boldsymbol{B}_t; t \in \text{supp}(T)$ is lonesum, then, by Lemma **L-12**, each element of $\boldsymbol{B}_t[i,j]$ can be expressed as:

$$\boldsymbol{B}_t[i,j] = \mathbf{1}\left[\left(\sum_{j'=1}^{N_{\boldsymbol{S}}} \mathbf{1}\left[\sum_{i'=1}^{N_Z} \boldsymbol{B}_t[i',j] \leq \sum_{i'=1}^{N_Z} \boldsymbol{B}_t[i',j']\right]\right) \leq \left(\sum_{j'=1}^{N_{\boldsymbol{S}}} \boldsymbol{B}_t[i,j']\right)\right] \text{ for all } t \in \text{supp}(T). \tag{A.20}$$

But $\text{P}(\boldsymbol{S} = \boldsymbol{s}_j) > 0$ for all $j \in \{1, \ldots, N_{\boldsymbol{S}}\}$. Thus by Lemma **L-13**, the following equality also holds:

$$\mathbf{1}\left[\left(\sum_{j'=1}^{N_{\boldsymbol{S}}} \mathbf{1}\left[\sum_{i'=1}^{N_Z} \boldsymbol{B}_t[i',j] \leq \sum_{i'=1}^{N_Z} \boldsymbol{B}_t[i',j']\right]\right) \leq \left[\sum_{j'=1}^{N_{\boldsymbol{S}}} \boldsymbol{B}_t[i,j']\right] = \right.$$
$$= \mathbf{1}\left[\left(\sum_{j'=1}^{N_{\boldsymbol{S}}} P(\boldsymbol{S}=\boldsymbol{s}_{j'}) \cdot \mathbf{1}\left[\sum_{i'=1}^{N_Z} \boldsymbol{B}_t[i',j] \leq \sum_{i'=1}^{N_Z} \boldsymbol{B}_t[i',j']\right]\right) \leq \left(\sum_{j'=1}^{N_{\boldsymbol{S}}} P(\boldsymbol{S}=\boldsymbol{s}_{j'}) \cdot \boldsymbol{B}_t[i,j']\right)\right]. \tag{A.21}$$

Since $\text{P}(Z = z_i) > 0$ for all $i \in \{1, \ldots, N_Z\}$, by Lemma **L-14**, the following equality also holds:

$$\mathbf{1}\left[\sum_{i'=1}^{N_Z} \boldsymbol{B}_t[i',j] \leq \sum_{i'=1}^{N_Z} \boldsymbol{B}_t[i',j']\right] = \mathbf{1}\left[\sum_{i'=1}^{N_Z} \text{P}(Z=z_{i'})\boldsymbol{B}_t[i',j] \leq \sum_{i'=1}^{N_Z} \text{P}(Z=z_{i'})\boldsymbol{B}_t[i',j']\right]. \tag{A.22}$$

Note that Equations (15) and (17) can be represented in terms of $\boldsymbol{B}_t$ as:

$$\mathrm{P}(T=t|Z=z_i)=\sum_{j=1}^{N_{\boldsymbol{S}}}\mathrm{P}(\boldsymbol{S}=\boldsymbol{s}_j)\boldsymbol{B}_t[i,j], \tag{A.23}$$

$$\mathrm{P}(T=t|\boldsymbol{S}=\boldsymbol{s}_j)=\sum_{i=1}^{N_Z}\mathrm{P}(Z=z_i)\boldsymbol{B}_t[i,j]. \tag{A.24}$$

If we substitute Equation (A.24) into (A.22), we obtain:

$$\mathbf{1}\Big[\sum_{i'=1}^{N_Z}\boldsymbol{B}_t[i',j]\leq\sum_{i'=1}^{N_Z}\boldsymbol{B}_t[i',j']\Big]=\mathbf{1}\Big[\mathrm{P}(T=t|\boldsymbol{S}=\boldsymbol{s}_j)\leq\mathrm{P}(T=t|\boldsymbol{S}=\boldsymbol{s}_{j'})\Big]. \tag{A.25}$$

If we substitute Equation (A.23) into (A.21), we can rewrite (A.20) as:

$$\boldsymbol{B}_t[i,j]=\mathbf{1}\Big[\sum_{j'=1}^{N_{\boldsymbol{S}}}\mathrm{P}(\boldsymbol{S}=\boldsymbol{s}_{j'})\cdot\mathbf{1}\big[\mathrm{P}(T=t|\boldsymbol{S}=\boldsymbol{s}_j)\leq\mathrm{P}(T=t|\boldsymbol{S}=\boldsymbol{s}_{j'})\big]\leq\mathrm{P}(T=t|Z=z_i)\Big]. \tag{A.26}$$

Thus if we define:

$$\tau(z_i,t)=\mathrm{P}(T=t|Z=z_i)$$

$$\varphi(\boldsymbol{s}_j,t)=-\sum_{j'=1}^{N_{\boldsymbol{S}}}\mathrm{P}(\boldsymbol{S}=\boldsymbol{s}_{j'})\cdot\mathbf{1}\big[\mathrm{P}(T=t|\boldsymbol{S}=\boldsymbol{s}_j)\leq\mathrm{P}(T=t|\boldsymbol{S}=\boldsymbol{s}_{j'})\big],$$

and use the fact that $\boldsymbol{B}_t[i,j]=\mathbf{1}[T=t|Z=z_i,\boldsymbol{S}=\boldsymbol{s}_j]$, we can rewrite Equation (A.26) as:

$$\mathbf{1}[T=t|Z=z_i,\boldsymbol{S}=\boldsymbol{s}_j]=\mathbf{1}\big[\varphi(\boldsymbol{s}_j,t)+\tau(z_i,t)\geq 0\big]. \tag{A.27}$$

Item (iv) of **T-3** is obtained using the fact that $\boldsymbol{S}$ is a balancing score for $\boldsymbol{V}$.

(iv) $\Rightarrow$ (ii): It suffices to show that if Equation (A.27) characterizes choice, then the prohibited pattern of condition (iii) cannot arise. Select an arbitrary 2×2 sub-matrix of the response matrix:

$$\boldsymbol{R} = \begin{pmatrix} (T|\boldsymbol{S}=\boldsymbol{s}, Z=z) & (T|\boldsymbol{S}=\boldsymbol{s}', Z=z) \\ (T|\boldsymbol{S}=\boldsymbol{s}, Z=z') & (T|\boldsymbol{S}=\boldsymbol{s}', Z=z') \end{pmatrix}$$
$$= \begin{pmatrix} \underset{t\in \text{supp}(T)}{\text{argmax}} \big(\varphi(\boldsymbol{s},t)+\tau(z,t)\big) & \underset{t\in \text{supp}(T)}{\text{argmax}} \big(\varphi(\boldsymbol{s}',t)+\tau(z,t)\big) \\ \underset{t\in \text{supp}(T)}{\text{argmax}} \big(\varphi(\boldsymbol{s},t)+\tau(z',t)\big) & \underset{t\in \text{supp}(T)}{\text{argmax}} \big(\varphi(\boldsymbol{s}',t)+\tau(z',t)\big) \end{pmatrix}. \tag{A.28}$$

In this notation, we must prove that if $\mathbf{1}[T=t|\boldsymbol{S}=\boldsymbol{s}, Z=z] = \mathbf{1}[T=t|\boldsymbol{S}=\boldsymbol{s}', Z=z'] = 1$ then it must be the case the case that $\mathbf{1}[T=t|\boldsymbol{S}=\boldsymbol{s}, Z=z'] = 1$ or $\mathbf{1}[T=t|\boldsymbol{S}=\boldsymbol{s}', Z=z] = 1$.

$$\begin{aligned}
\mathbf{1}[T=t|\boldsymbol{S}=\boldsymbol{s}, Z=z] = 1 &\Rightarrow \varphi(\boldsymbol{s},t)+\tau(z,t) \geq 0 \text{ and } \varphi(\boldsymbol{s},t')+\tau(z,t') < 0 \quad \forall\, t' \in \text{supp}(T)\setminus\{t\} \\
\mathbf{1}[T=t|\boldsymbol{S}=\boldsymbol{s}', Z=z'] = 1 &\Rightarrow \varphi(\boldsymbol{s}',t)+\tau(z',t) \geq 0 \text{ and } \varphi(\boldsymbol{s}',t')+\tau(z',t') < 0 \quad \forall\, t' \in \text{supp}(T)\setminus\{t\} \\
&\Rightarrow \varphi(\boldsymbol{s},t)+\tau(z',t) \geq 0 \text{ or } \varphi(\boldsymbol{s}',t')+\tau(z,t) \geq 0 \\
&\Rightarrow \mathbf{1}[T=t|\boldsymbol{S}=\boldsymbol{s}, Z=z'] = 1 \text{ or } \mathbf{1}[T=t|\boldsymbol{S}=\boldsymbol{s}', Z=z] = 1.^{4}
\end{aligned}$$

□

A.7 Proof of Theorem T-4

Proof. It suffices to show that the assumptions of Theorem T-4 imply any of the conditions of the Equivalence Theorem T-3. We show that the separability of Theorem T-4 implies condition (iii) of Theorem T-3. $\boldsymbol{S}$ is a balancing score for $\boldsymbol{V}$, thus for any $\boldsymbol{v} \in \text{supp}(\boldsymbol{V})$ there is a unique $\boldsymbol{s} \in \text{supp}(\boldsymbol{S})$ such that $\boldsymbol{s} = f_{\boldsymbol{S}}(\boldsymbol{v})$. Without loss of generality, we can rewrite the separability condition of Theorem T-4 as:

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

[4] To see this, suppose that $\varphi(\boldsymbol{s},t)+\tau(z',t) < 0$ and $\varphi(\boldsymbol{s}',t)+\tau(z,t) < 0$, then $\varphi(\boldsymbol{s},t)+\tau(z,t)+\varphi(\boldsymbol{s}',t)+\tau(z',t) < 0$, which contradicts the hypothesis.

Let $\boldsymbol{s} \in \operatorname{supp}(\boldsymbol{S}), z \in \operatorname{supp}(Z)$ and $t \in \operatorname{supp}(T)$ such that $t = \operatorname{argmax}_{t' \in \operatorname{supp}(T)} u(\boldsymbol{s}, t') + h(z, t')$. Then it must be the case that $t = \operatorname{argmax}_{t' \in \operatorname{supp}(T)} \Psi(t, z, \boldsymbol{s})$. Thus, assuming no ties in utility outcomes, $\Psi(t, z, \boldsymbol{s}) > \Psi(t', z, \boldsymbol{s}) \forall t' \in \operatorname{supp}(T) \setminus \{t\}$ and therefore $u(\boldsymbol{s}, t) + h(z, t) > 0$ and $u(\boldsymbol{s}, t') + h(z, t') \leq 0 \ \forall \, t' \in \operatorname{supp}(T) \setminus \{t\}$. Thus we obtain:

$$t = \operatorname{argmax}_{t' \in \operatorname{supp}(T)} u(\boldsymbol{s}, t') + h(z, t') \Leftrightarrow u(\boldsymbol{s}, t) + h(z, t) > 0. \tag{A.29}$$

Now for condition (iii) of **T-3** to hold, we need to prove the following statement:

$$\text{Let } \boldsymbol{s}, \boldsymbol{s}' \in \operatorname{supp}(\boldsymbol{S}), z, z' \in \operatorname{supp}(Z), \text{ and } t \in \operatorname{supp}(T).$$

$$\text{If } t = \operatorname{argmax}_{t' \in \operatorname{supp}(T)} u(\boldsymbol{s}, t') + h(z, t') \text{ and } t = \operatorname{argmax}_{t' \in \operatorname{supp}(T)} u(\boldsymbol{s}', t') + h(z', t'), \tag{A.30}$$

$$\text{then } t = \operatorname{argmax}_{t' \in \operatorname{supp}(T)} u(\boldsymbol{s}, t') + h(z', t') \text{ or } t = \operatorname{argmax}_{t' \in \operatorname{supp}(T)} u(\boldsymbol{s}', t') + h(z, t'). \tag{A.31}$$

But, according to Equation (A.29), Equation (A.30) implies that $u(\boldsymbol{s}, t) + h(z, t) > 0$ and $u(\boldsymbol{s}', t) + h(z', t) > 0$. This implies that $u(\boldsymbol{s}, t) + h(z', t) > 0$ or $u(\boldsymbol{s}', t) + h(z, t) > 0$ or both. Therefore, according to Equation (A.29), it must be the case that:

$$t = \operatorname{argmax}_{t' \in \operatorname{supp}(T)} u(\boldsymbol{s}, t') + h(z', t') \text{ or } t = \operatorname{argmax}_{t' \in \operatorname{supp}(T)} u(\boldsymbol{s}', t') + h(z, t')$$

as desired. □

A.8 Proof of Theorem T-5

The proof of Theorem **T-5** is based on Lemma **L-15** stated below.

A.8.1 Lemma L-15

Lemma L-15. Binary matrix $\boldsymbol{B}$ is lonesum $\Leftrightarrow \iota_c' \Big(\big(\boldsymbol{B}'(\iota_r \iota_c' - \boldsymbol{B})\big) \odot \big(\boldsymbol{B}'(\iota_r \iota_c' - \boldsymbol{B})\big)' \Big) \iota_c = 0$, where ι_c and ι_r are vectors of elements 1 of column and row dimension of $\boldsymbol{B}$ respectively.

Proof. From Remark A.1, $\boldsymbol{B}$ is lonesum if and only if **no** 2×2 sub-matrix is equal to the prohibited patterns (52). Each row i of any two columns j, j' of $\boldsymbol{B}$, say $(\boldsymbol{B}[i, j] \ , \ \boldsymbol{B}[i, j'])$

must be of the following four types $\{(0,0), (0,1), (1,0), (1,1)\}$. But $\boldsymbol{B}[i,j]\cdot(1-\boldsymbol{B}[i,j']) = 1$ if and only if $(\boldsymbol{B}[i,j], \boldsymbol{B}[i,j']) = (1,0)$ and $(1-\boldsymbol{B}[i,j])\cdot\boldsymbol{B}[i,j'] = 1$ if and only if $(\boldsymbol{B}[i,j], \boldsymbol{B}[i,j']) = (0,1)$. This fact is illustrated in the table below:

Row Type	Any Columns j,j' $\boldsymbol{B}[i,j]$	$\boldsymbol{B}[i,j']$	$(\boldsymbol{B}[i,j], \boldsymbol{B}[i,j']) = (1,0)$ $\boldsymbol{B}[i,j]\cdot(1-\boldsymbol{B}[i,j'])$	$(\boldsymbol{B}[i,j], \boldsymbol{B}[i,j']) = (0,1)$ $(1-\boldsymbol{B}[i,j])\cdot\boldsymbol{B}[i,j']$
Type 1	0	0	0	0
Type 2	0	1	0	1
Type 3	1	0	1	0
Type 4	1	1	0	0

Thus the vector multiplication $\xi_{(0,1)}(j,j') = \left(\iota_c - \boldsymbol{B}[\cdot,j]\right)\cdot\boldsymbol{B}[\cdot,j']$ gives the number of rows equal to $(0,1)$ in the sub-matrix of $\boldsymbol{B}$ that consists of columns j and j'. In the same fashion, $\xi_{(1,0)}(j,j') = \left(\iota_c - \boldsymbol{B}[\cdot,j']\right)\cdot\boldsymbol{B}[\cdot,j]$ gives the number of rows equal to $(1,0)$ in the sub-matrix of $\boldsymbol{B}$ that consists of columns j and j'. If $\xi_{(1,0)}(j,j') > 0$ and $\xi_{(0,1)}(j,j') > 0$ then there exists at least one $(0,1)$-row and at least one $(1,0)$-row in the sub-matrix of $\boldsymbol{B}$ that consists of columns j and j'. This would show the presence of a prohibited pattern in $\boldsymbol{B}$. Thus, $\boldsymbol{B}$ is lonesum if and only if $\xi_{(0,1)}(j,j')\cdot\xi_{(1,0)}(j,j') = 0$ for all pairs $(j,j') \in \{1,\ldots,c\}\times\{1,\ldots,c\}$. Stated otherwise,

$$\boldsymbol{B} \text{ is lonesum } \Leftrightarrow \sum_{j=1}^{c}\sum_{j'=1}^{c}\xi_{(1,0)}(j,j')\cdot\xi_{(0,1)}(j,j') = 0. \tag{A.32}$$

The lemma is obtained by rewriting Equation (A.32) in matrix form. □

A.8.2 Proof of Theorem T-5

Proof. We first prove that if each binary matrix $\boldsymbol{B}_t; t \in \text{supp}(T)$ is lonesum then $\boldsymbol{M}$ is also lonesum. According to Lemma **L-5**, it suffices to show that the prohibited patterns (52) cannot arise in any 2×2 submatrix in $\boldsymbol{M}$. Recall that the from definition (57), we have that:

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

Let a generic 2×2 sub-matrix of $\boldsymbol{M}$ above be represented in by matrix (A.33):

$$\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \tag{A.33}$$

We investigate all possible configurations that matrix (A.33) may take:

1. If all elements $b_{11}, b_{12}, b_{21}, b_{22}$ of matrix (A.33) belong to some $\boldsymbol{B}_t$ then the prohibited pattern does not arise because each $\boldsymbol{B}_t$ is lonesum.

2. If none of the elements $b_{11}, b_{12}, b_{21}, b_{22}$ of matrix (A.33) belong to any of the binary matrices in $\boldsymbol{B}_t$, then matrix (A.33) takes one of the four possibilities below:

$$\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \in \left\{ \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \right\}.$$

3. If b_{11}, b_{12} of matrix (A.33) belong to some $\boldsymbol{B}_t$ but b_{21}, b_{22} do not belong to any of the $\boldsymbol{B}_t; t \in \text{supp}(T)$, then it must be the case that $b_{21} = b_{22} = 1$, which prevents the prohibited pattern from arising. On the other hand, if b_{21}, b_{22} belong to some $\boldsymbol{B}_t$ but b_{11}, b_{12} do not, then it must be the case that $b_{11} = b_{12} = 0$, which prevents the prohibited pattern.

4. If b_{11}, b_{21} of matrix (A.33) belongs to some $\boldsymbol{B}_t$ but b_{12}, b_{22} do not belong to any of the $\boldsymbol{B}_t; t \in \text{supp}(T)$, then it must be the case that $b_{12} = b_{22} = 0$, which prevents the prohibited pattern from arising. On the other hand, if b_{12}, b_{22} belong to some $\boldsymbol{B}_t$

but b_{11}, b_{21} do not, then it must be the case that $b_{11} = b_{21} = 1$, which prevents the prohibited pattern.

5. If b_{11} or b_{21} of matrix (A.33) belongs to one or two matrices $\boldsymbol{B}_t; t \in \text{supp}(T)$ and b_{12}, b_{21} do not belong to any of the matrices $\boldsymbol{B}_t; t \in \text{supp}(T)$, then it must be the case that $b_{12} = 0$ and $b_{21} = 1$, which prevents the prohibited pattern from arising.

There are no other possibilities besides the ones listed above. Thus we can conclude that no 2×2 submatrix in $\boldsymbol{M}$ takes the prohibited pattern (52) and, by Lemma **L-5**, $\boldsymbol{M}$ is lonesum. According to Lemma **L-15** $\boldsymbol{M}$ is lonesum if and only if $\iota_c'\Big(\big(\boldsymbol{M}'(\iota_r\iota_c' - \boldsymbol{M})\big) \odot \big(\boldsymbol{M}'(\iota_r\iota_c' - \boldsymbol{M})\big)'\Big)\iota_c = 0$, where ι_c and ι_r are vectors of elements 1 of column and row dimension of $\boldsymbol{M}$ respectively, which completes the proof.

□

A.9 Proof of Theorem T-6

We generate the expressions that identify $P(\boldsymbol{S} \in \Sigma_t(i))$ and $E(\kappa(Y(t))|\boldsymbol{S} \in \Sigma_t(i))$. Those parameters can be rewritten in the following matrix form using the notation of Section 4:

$$P(S \in \Sigma_t(i)) = \boldsymbol{b}_t(i)\boldsymbol{P_S} \tag{A.34}$$

$$\text{and } E(\kappa(Y(t))|S \in \Sigma_t(i)) = \frac{\boldsymbol{b}_t(i)\boldsymbol{Q_S}(t)}{\boldsymbol{b}_t(i)\boldsymbol{P_S}} \tag{A.35}$$

Thus the theorem requires us to identify the terms $\boldsymbol{b}_t(i)\boldsymbol{P_S}$ and $\boldsymbol{b}_t(i)\boldsymbol{Q_S}(t)$. In other words, we aim to show that the terms $\boldsymbol{b}_t(i)\boldsymbol{P_S}$ and $\boldsymbol{b}_t(i)\boldsymbol{Q_S}(t)$ can be expressed in terms of observables. To do so, we can rely on Equations (18)–(19) and the generalized solution of linear system described in Lemma **L-2**:

$$\boldsymbol{P}_Z(t) = \boldsymbol{B}_t\boldsymbol{P_S} \Rightarrow \boldsymbol{b}_t(i)\boldsymbol{P_S} = \boldsymbol{b}_t(i)\Big(\boldsymbol{B}_t^+\boldsymbol{P}_Z(t) + (\boldsymbol{I} - \boldsymbol{B}_t^+\boldsymbol{B}_t)\boldsymbol{\lambda}_P\Big), \tag{A.36}$$

$$\boldsymbol{Q}_Z(t) = \boldsymbol{B}_t\boldsymbol{Q_S}(t) \Rightarrow \boldsymbol{b}_t(i)\boldsymbol{Q_S}(t) = \boldsymbol{b}_t(i)\Big(\boldsymbol{B}_t^+\boldsymbol{Q}_Z(t) + (\boldsymbol{I} - \boldsymbol{B}_t^+\boldsymbol{B}_t)\boldsymbol{\lambda}_Q\Big); \tag{A.37}$$

where $\boldsymbol{\lambda}_Q, \boldsymbol{\lambda}_P$ are arbitrary real valued vectors of $N_{\boldsymbol{S}}$ dimension. Equation (A.36) shows that $\boldsymbol{b}_t(i)\boldsymbol{P_S}$ can be expressed by the sum of two terms:

1. Term $\boldsymbol{b}_t(i)\boldsymbol{B}_t^{+}\boldsymbol{P}_Z(t)$ which can be determined by the data and is identified.

2. Term $\boldsymbol{b}_t(i)(\boldsymbol{I}-\boldsymbol{B}_t^{+}\boldsymbol{B}_t)\boldsymbol{\lambda}_P$ is a source of non-identification as $\boldsymbol{\lambda}_P$ is unknown and can take any value in $\mathbb{R}^{N_{\boldsymbol{S}}}$.

Thus a necessary and sufficient condition for $\boldsymbol{b}_t(i)\boldsymbol{P}_{\boldsymbol{S}}$ to be identified is that $\boldsymbol{b}_t(i)(\boldsymbol{I}-\boldsymbol{B}_t^{+}\boldsymbol{B}_t)$ be equal to a vector of zeros. The same requirement applies to the identification of $\boldsymbol{b}_t(i)\boldsymbol{Q}_{\boldsymbol{S}}(t)$. Thus, to prove the theorem, it suffices to demonstrate that $\boldsymbol{b}_t(i)(\boldsymbol{I}-\boldsymbol{B}_t^{+}\boldsymbol{B}_t)=\boldsymbol{0}$. Otherwise stated, we need to show that $\boldsymbol{b}_t(i)\boldsymbol{B}_t^{+}\boldsymbol{B}_t=\boldsymbol{b}_t(i)$. We prove this condition in several steps. We first prove two lemmas that are useful to prove that $\boldsymbol{b}_t(i)\boldsymbol{B}_t^{+}\boldsymbol{B}_t=\boldsymbol{b}_t(i)$.

Lemma L-16. Let $\boldsymbol{B}_t$ be the binary matrix associated with a response matrix $\boldsymbol{R}$ for which unordered monotonicity **A-3** holds. If vectors $\boldsymbol{B}_t[\cdot,\boldsymbol{s}]$ and $\boldsymbol{B}_t[\cdot,\boldsymbol{s}']$ associated with response types $\boldsymbol{s},\boldsymbol{s}'\in\text{supp}(S)$ have the same sum, then these vectors must be identical.

Proof. Suppose $\boldsymbol{B}_t[\cdot,\boldsymbol{s}]$ and $\boldsymbol{B}_t[\cdot,\boldsymbol{s}']$ have the same sum but are not identical. Then there must be at least two row indexes j,j' such that $\boldsymbol{B}_t[j,\boldsymbol{s}]=1$, $\boldsymbol{B}_t[j,\boldsymbol{s}']=0$ and $\boldsymbol{B}_t[j',\boldsymbol{s}]=0$, $\boldsymbol{B}_t[j',\boldsymbol{s}']=1$. Then the 2×2 sub-matrix generated by rows j,j' and columns $\boldsymbol{s},\boldsymbol{s}'$ of $\boldsymbol{B}_t$ constitute a prohibited pattern of Remark *6.3* and therefore $\boldsymbol{B}_t$ is not lonesum, which contradicts unordered monotonicity **A-3** according to Item (i) of Theorem **T-3**. □

Remark A.2. *Lemma **L-16** can be equivalently stated as: if unordered monotonicity **A-3** holds and $\boldsymbol{s},\boldsymbol{s}'\in\Sigma_t(i)$, then $\boldsymbol{B}_t[\cdot,\boldsymbol{s}]=\boldsymbol{B}_t[\cdot,\boldsymbol{s}']$.*

Now recall that $\Sigma_t(i); i\in\{1,\ldots,N_Z\}$ is the set of response types $\boldsymbol{s}\in\text{supp}(S)$ whose sum of the associated vector $\boldsymbol{B}_t[\cdot,\boldsymbol{s}]$ is n. According to **L-16**, each response type $\boldsymbol{s}\in\Sigma_t(i)$ has the same binary vector $\boldsymbol{B}_t[\cdot,\boldsymbol{s}]$ in $\boldsymbol{B}_t$.

Let $\boldsymbol{C}_t(i)=\boldsymbol{B}_t[\cdot,\boldsymbol{s}];\boldsymbol{s}\in\Sigma_t(i)$ denote this vector. We now define two matrices that are useful for our analysis. Let $\boldsymbol{C}_t=[\boldsymbol{C}_t(1),\ldots,\boldsymbol{C}_t(n_Z)]$ be the matrix that consists of all unique non-zero vectors in $\boldsymbol{B}_t$. Thus, columns in matrix $\boldsymbol{C}_t$ have different sums that may range from 1 to N_Z. As a consequence, $\boldsymbol{C}_t$ has at most N_Z columns and its row dimension is also N_Z.

Moreover, $\boldsymbol{B}_t$ is lonesum which implies that $\boldsymbol{C}_t$ is also lonesum. Thus $\boldsymbol{C}_t$ is equivalent to a maximal matrix under permutation of its columns and rows. This implies that $\boldsymbol{C}_t$ has full column rank, thereby $\boldsymbol{C}_t'\boldsymbol{C}_t$ has full rank and the inverse $(\boldsymbol{C}_t'\boldsymbol{C}_t)^{-1}$ exists.

Let $\boldsymbol{D}_t$ be the matrix that stacks the non-zeros row-vectors $b_t(1),\ldots,b_t(N_Z)$, namely $\boldsymbol{D}_t = [b_t(1)',\ldots,b_t(N_Z)']'$. $\boldsymbol{D}_t$ has $N_{\boldsymbol{S}}$ columns and has at most N_Z rows. The sum of each column in $\boldsymbol{D}_t$ is equal to one or zero. The sum of each row in $\boldsymbol{D}_t$ is equal or bigger than one and its rows are orthogonal, that is $b_t(i)\cdot b_t(i') = 0$ for any $i,i' \in \{1,\ldots,N_Z\}$. As a consequence, $\boldsymbol{D}_t$ has full row-rank and thereby the inverse $(\boldsymbol{D}_t'\boldsymbol{D}_t)^{-1}$ exists.

Remark A.3. The binary matrix $\boldsymbol{B}_t$ can be conveniently decomposed by the matrix multiplication $\boldsymbol{B}_t = \boldsymbol{C}_t \cdot \boldsymbol{D}_t$.

The example below illustrates the decomposition.

Example **E-1**. Consider binary matrix $\boldsymbol{B}_{t_a}$ of Table 7 that is associated with choice t_a of the response matrix $\boldsymbol{R}$ in Table 3. The decomposition $\boldsymbol{B}_{t_a} = \boldsymbol{C}_{t_a}\cdot\boldsymbol{D}_{t_a}$ is given by:

$$\boldsymbol{B}_{t_a} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \Rightarrow \boldsymbol{B}_{t_a} = \boldsymbol{C}_{t_a}\cdot\boldsymbol{D}_{t_a},$$

$$\text{where } \boldsymbol{C}_{t_a} = \Big[\boldsymbol{C}_{t_a}(1), \boldsymbol{C}_{t_a}(2), \boldsymbol{C}_{t_a}(3)\Big] = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix},$$

$$\text{and } \boldsymbol{D}_{t_a} = \begin{bmatrix} \boldsymbol{b}_{t_a}(1) \\ \boldsymbol{b}_{t_a}(2) \\ \boldsymbol{b}_{t_a}(3) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

The next lemma expresses the pseudo-inverse matrix $\boldsymbol{B}_t^+$ in terms of $\boldsymbol{C}_t$ and $\boldsymbol{D}_t$.

Lemma L-17. Let $\boldsymbol{B}_t$ be the binary matrix associated with a response matrix $\boldsymbol{R}$ for which unordered monotonicity **A-3** holds. Then the Moore-Penrose pseudoinverse $\boldsymbol{B}_t^+$ is given by:

$$\boldsymbol{B}_t^+ = \boldsymbol{D}_t'\big(\boldsymbol{D}_t\boldsymbol{D}_t'\big)^{-1}(\boldsymbol{C}_t'\boldsymbol{C}_t)^{-1}\boldsymbol{C}_t'.$$

Proof. Matrix $\boldsymbol{B}^+$ is defined by four properties: (1) $\boldsymbol{B}\boldsymbol{B}^+\boldsymbol{B} = \boldsymbol{B}$; (2) $\boldsymbol{B}^+\boldsymbol{B}\boldsymbol{B}^+ = \boldsymbol{B}^+$; (3) $\boldsymbol{B}^+\boldsymbol{B}$ is symmetric and (4) $\boldsymbol{B}\boldsymbol{B}^+$ is symmetric. Matrix $\boldsymbol{B}^+$ is also unique. Thus it suffices to show that $\boldsymbol{B}_t^+ = \boldsymbol{D}_t'\big(\boldsymbol{D}_t\boldsymbol{D}_t'\big)^{-1}(\boldsymbol{C}_t'\boldsymbol{C}_t)^{-1}\boldsymbol{C}_t'$, where $\boldsymbol{B}_t = \boldsymbol{C}_t\boldsymbol{D}_t$ satisfies the properties above.

$$
\begin{aligned}
(1)\ \boldsymbol{B}_t \cdot \boldsymbol{B}_t^+ \cdot \boldsymbol{B}_t \quad &= \boldsymbol{C}_t\boldsymbol{D}_t \cdot \boldsymbol{D}_t'\big(\boldsymbol{D}_t\boldsymbol{D}_t'\big)^{-1}(\boldsymbol{C}_t'\boldsymbol{C}_t)^{-1}\boldsymbol{C}_t' \cdot \boldsymbol{C}_t\boldsymbol{D}_t \\
&= \boldsymbol{C}_t\boldsymbol{D}_t = \boldsymbol{B}_t \\
(2)\ \boldsymbol{B}_t^+ \cdot \boldsymbol{B}_t \cdot \boldsymbol{B}_t^+ \quad &= \boldsymbol{D}_t'\big(\boldsymbol{D}_t\boldsymbol{D}_t'\big)^{-1}(\boldsymbol{C}_t'\boldsymbol{C}_t)^{-1}\boldsymbol{C}_t' \cdot \boldsymbol{C}_t\boldsymbol{D}_t \cdot \boldsymbol{D}_t'\big(\boldsymbol{D}_t\boldsymbol{D}_t'\big)^{-1}(\boldsymbol{C}_t'\boldsymbol{C}_t)^{-1}\boldsymbol{C}_t' \\
&= \boldsymbol{D}_t'\big(\boldsymbol{D}_t\boldsymbol{D}_t'\big)^{-1}(\boldsymbol{C}_t'\boldsymbol{C}_t)^{-1}\boldsymbol{C}_t' = \boldsymbol{B}_t^+ \\
(3)\ \boldsymbol{B}_t^+ \cdot \boldsymbol{B}_t \quad &= \boldsymbol{D}_t'\big(\boldsymbol{D}_t\boldsymbol{D}_t'\big)^{-1}(\boldsymbol{C}_t'\boldsymbol{C}_t)^{-1}\boldsymbol{C}_t' \cdot \boldsymbol{C}_t\boldsymbol{D}_t \\
&= \boldsymbol{D}_t'\big(\boldsymbol{D}_t\boldsymbol{D}_t'\big)^{-1}\boldsymbol{D}_t \text{ which is symmetric} \\
(4)\ \boldsymbol{B}_t \cdot \boldsymbol{B}_t^+ \quad &= \boldsymbol{C}_t\boldsymbol{D}_t \cdot \boldsymbol{D}_t'\big(\boldsymbol{D}_t\boldsymbol{D}_t'\big)^{-1}(\boldsymbol{C}_t'\boldsymbol{C}_t)^{-1}\boldsymbol{C}_t \\
&= \boldsymbol{C}_t(\boldsymbol{C}_t'\boldsymbol{C}_t)^{-1}\boldsymbol{C}_t \text{ which is symmetric.}
\end{aligned}
$$

□

We are now equipped to prove Theorem **T-6** and show that $\boldsymbol{b}_t(i)\boldsymbol{B}_t^+\boldsymbol{B}_t = \boldsymbol{b}_t(i)$.

Proof. First note that $\boldsymbol{b}_t(i)$ consists of a row in the matrix $\boldsymbol{D}_t$ thus $\boldsymbol{b}_t(i)$ can be expressed in terms of matrix $\boldsymbol{D}_t$ as $\boldsymbol{b}_t(i) = \boldsymbol{e}\boldsymbol{D}_t$, where $\boldsymbol{e}$ is a vector that has the element 1 in the position that the row $\boldsymbol{b}_t(i)$ takes in the matrix $\boldsymbol{D}_t$ and zeros in the remaining elements. Thus we have to show that $\boldsymbol{e}\boldsymbol{D}_t\boldsymbol{B}_t^+\boldsymbol{B}_t = \boldsymbol{e}\boldsymbol{D}_t$. We prove the more general statement that $\boldsymbol{D}_t\boldsymbol{B}_t^+\boldsymbol{B}_t = \boldsymbol{D}_t$.

$$\boldsymbol{D}_t \cdot \boldsymbol{B}_t^+ \cdot \boldsymbol{B}_t = \boldsymbol{D}_t \cdot \boldsymbol{D}_t'\big(\boldsymbol{D}_t\boldsymbol{D}_t'\big)^{-1}(\boldsymbol{C}_t'\boldsymbol{C}_t)^{-1}\boldsymbol{C}_t' \cdot \boldsymbol{C}_t\boldsymbol{D}_t = \boldsymbol{D}_t, \tag{A.38}$$

where the first equality in (A.38) relies on the result of Lemma **L-17** and Remark A.3. Equation (A.38) implies that $\boldsymbol{b}_t(i)\boldsymbol{B}_t^+\boldsymbol{B}_t = \boldsymbol{b}_t(i)$, The fact that $\boldsymbol{b}_t(i)\boldsymbol{B}_t^+\boldsymbol{B}_t = \boldsymbol{b}_t(i)$ implies that $\boldsymbol{b}_t(i)\boldsymbol{P}_{\boldsymbol{S}}$ and $\boldsymbol{b}_t(i)\boldsymbol{Q}_{\boldsymbol{S}}(t)$ can be identified by:

$$\boldsymbol{b}_t(i)\boldsymbol{P}_{\boldsymbol{S}} = \boldsymbol{b}_t(i)\boldsymbol{B}_t^+\boldsymbol{P}_Z(t), \tag{A.39}$$

$$\boldsymbol{b}_t(i)\boldsymbol{Q}_{\boldsymbol{S}}(t) = \boldsymbol{b}_t(i)\boldsymbol{B}_t^+\boldsymbol{Q}_Z(t). \tag{A.40}$$

We now use the equations above to express Equations (A.34)–(A.35) of the beginning of this proof as identified quantities:

$$\begin{aligned} P(S \in \Sigma_t(i)) &= \boldsymbol{b}_t(i)\boldsymbol{P_S} &= \boldsymbol{b}_t(i)\boldsymbol{B}_t^+\boldsymbol{P}_Z(t), \quad \text{by (A.39)} \\ \text{and } E(\kappa(Y(t))|\boldsymbol{S} \in \Sigma_t(i)) &= \frac{\boldsymbol{b}_t(i)\boldsymbol{Q_S}(t)}{\boldsymbol{b}_t(i)\boldsymbol{P_S}} &= \frac{\boldsymbol{b}_t(i)\boldsymbol{B}_t^+\boldsymbol{Q}_Z(t)}{\boldsymbol{b}_t(i)\boldsymbol{B}_t^+\boldsymbol{P}_Z(t)}, \quad \text{by (A.39) and (A.40).} \end{aligned}$$

□

A.10 Applying Theorem T-6 to Response Matrix of Table 3

Table 3 presents the following response matrix:

$$\boldsymbol{R} = \begin{bmatrix} t_a & t_a & t_a & t_b & t_b & t_c & t_c \\ t_a & t_a & t_a & t_a & t_b & t_a & t_c \\ t_a & t_b & t_c & t_b & t_b & t_c & t_c \end{bmatrix}$$

According to Theorem T-6, we have that:

$$E(Y(t_a)|\boldsymbol{S} \in \{\boldsymbol{s}_4, \boldsymbol{s}_6\}) = E(Y(t_a)|\boldsymbol{S} \in \Sigma_{t_a}(1)) = \frac{\boldsymbol{b}_{t_a}(1)\boldsymbol{B}_t^+\boldsymbol{Q}_Z(t_a)}{\boldsymbol{b}_{t_a}(1)\boldsymbol{B}_t^+\boldsymbol{P}_Z(t_a)}; \boldsymbol{b}_{t_a}(1) = [0,0,0,1,0,1,0], \tag{A.41}$$

$$E(Y(t_a)|\boldsymbol{S} \in \{\boldsymbol{s}_2, \boldsymbol{s}_3\}) = E(Y(t_a)|\boldsymbol{S} \in \Sigma_{t_a}(2)) = \frac{\boldsymbol{b}_{t_a}(2)\boldsymbol{B}_t^+\boldsymbol{Q}_Z(t_a)}{\boldsymbol{b}_{t_a}(2)\boldsymbol{B}_t^+\boldsymbol{P}_Z(t_a)}; \boldsymbol{b}_{t_a}(2) = [0,1,1,0,0,0,0], \tag{A.42}$$

$$E(Y(t_a)|\boldsymbol{S} = \boldsymbol{s}_1) = E(Y(t_a)|\boldsymbol{S} \in \Sigma_{t_a}(3)) = \frac{\boldsymbol{b}_{t_a}(3)\boldsymbol{B}_t^+\boldsymbol{Q}_Z(t_a)}{\boldsymbol{b}_{t_a}(3)\boldsymbol{B}_t^+\boldsymbol{P}_Z(t_a)}; \boldsymbol{b}_{t_a}(3) = [1,0,0,0,0,0,0]. \tag{A.43}$$

The observed parameters are:

$$\boldsymbol{Q}_Z(t_a) = [E(Y \cdot \mathbf{1}[T = t_1]|Z = z_{no}), E(Y \cdot \mathbf{1}[T = t_1]|Z = z_a), E(Y \cdot \mathbf{1}[T = t_1]|Z = z_{bc})], \tag{A.44}$$

$$\boldsymbol{P}_Z(t_a) = [E(\mathbf{1}[T = t_1]|Z = z_{no}), E(\mathbf{1}[T = t_1]|Z = z_a), E(\mathbf{1}[T = t_1]|Z = z_{bc})]. \tag{A.45}$$

The binary matrix $\boldsymbol{B}_{t_a} = \mathbf{1}[\boldsymbol{R} = t_1]$ is played in Table 7. This binary matrix generates the following generalized inverse matrix:

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

Matrix $\boldsymbol{B}_{t_a}^{+}$ multiplied by $\boldsymbol{b}_{t_a}(i); i = 1, 2, 3$ is given by:

$$\boldsymbol{b}_{t_a}(1) \cdot \boldsymbol{B}_{t_a}^{+} = [-1, 1, 0],$$

$$\boldsymbol{b}_{t_a}(2) \cdot \boldsymbol{B}_{t_a}^{+} = [1, 0, -1],$$

$$\boldsymbol{b}_{t_a}(3) \cdot \boldsymbol{B}_{t_a}^{+} = [0, 0, 1].$$

Matrix $\boldsymbol{B}_{t_a}^{+}$ applied to equations (A.41)–(A.43) generates the following identifying equations:

$$\begin{aligned} E(Y(t_a)|\boldsymbol{S} \in \{\boldsymbol{s}_4, \boldsymbol{s}_6\}) &= \frac{E(Y \cdot \mathbf{1}[T = t_a]|Z = z_a) - E(Y \cdot \mathbf{1}[T = t_a]|Z = z_{no})}{P(T = t_a|Z = z_a) - P(T = t_a|Z = z_{no})}, \\ E(Y(t_a)|\boldsymbol{S} \in \{\boldsymbol{s}_2, \boldsymbol{s}_3\}) &= \frac{E(Y \cdot \mathbf{1}[T = t_a]|Z = z_{no}) - E(Y \cdot \mathbf{1}[T = t_a]|Z = z_{bc})}{P(T = t_a|Z = z_{no}) - P(T = t_a|Z = z_{bc})}, \\ E(Y(t_a)|\boldsymbol{S} = \boldsymbol{s}_1) &= \frac{E(Y \cdot \mathbf{1}[T = t_a]|Z = z_{bc})}{P(T = t_a|Z = z_{bc}))}. \end{aligned}$$

A.11 Proof of Corollary C-2

Proof. The set of t-Always-takers is denoted by $\Sigma_t(N_Z)$ and consists of a single response type in $\boldsymbol{R}$ whose elements are all t. From Theorem **T-6**, $P(S \in \Sigma_t(N_Z)) = \boldsymbol{b}_t(N_Z)\boldsymbol{B}_t^{+}\boldsymbol{P}_Z(t)$.

The response-type set t-Switchers is given by t-Switchers $\in \cup_{i=1}^{N_Z-1}\Sigma_t(i)$, thus:

$$\begin{aligned} P(\boldsymbol{S} \in \cup_{i=1}^{N_Z-1}\Sigma_t(i)) &= \sum_{i=1}^{N_Z-1} P(S = \Sigma_t(i)) \\ &= \sum_{i=1}^{N_Z-1} \boldsymbol{b}_t(i)\boldsymbol{B}_t^{+}\boldsymbol{P}_Z(t) \text{ by Theorem } \textbf{T-6} \\ &= \left(\sum_{i=1}^{N_Z-1} \boldsymbol{b}_t(i)\right) \boldsymbol{B}_t^{+}\boldsymbol{P}_Z(t). \end{aligned}$$

The response variable support supp($\boldsymbol{S}$) can be partitioned as:

supp(S) $=$ t-Always-takers $\cup$ t-Switchers $\cup$ t-Never-takers. Thus:

$$\begin{aligned} 1 &= P(\boldsymbol{S} \in \text{supp}(\boldsymbol{S})) \\ &= P(\boldsymbol{S} \in t\text{-Always-takers} \cup t\text{-Switchers} \cup t\text{-Never-takers}) \\ &= P(\boldsymbol{S} \in t\text{-Always-takers}) + P(\boldsymbol{S} \in t\text{-Switchers}) + P(\boldsymbol{S} \in t\text{-Never-takers}). \end{aligned}$$

□

A.12 Proof of Corollary C-3

Proof. The equation for $E(Y(t)|t\text{-Always-takers})$ is a direct application of Theorem **T-6** to the response-type set t-Always-takers $= \Sigma_t(N_Z)$. Applying Theorem **T-6** to $E(Y(t)|t\text{-Switchers})$ we obtain:

$$\begin{aligned} E(Y(t)|t\text{-Switchers}) &= \sum_{i=1}^{N_Z-1} E(Y(t)|\boldsymbol{S} \in \Sigma_t(i)) \cdot \frac{P(\boldsymbol{S} \in \Sigma_t(i))}{P(\boldsymbol{S} \in \cup_{i=1}^{N_Z-1}\Sigma_t(i))} \\ E(Y(t)|t\text{-Switchers}) &= \sum_{i=1}^{N_Z-1} \left(\frac{\boldsymbol{b}_t(i)\boldsymbol{B}_t^{+}\boldsymbol{Q}_Z(t)}{\boldsymbol{b}_t(i)\boldsymbol{B}_t^{+}\boldsymbol{P}_Z(t)}\right) \frac{\boldsymbol{b}_t(i)\boldsymbol{B}_t^{+}\boldsymbol{P}_Z(t)}{\left(\sum_{i=1}^{N_Z-1} \boldsymbol{b}_t(i)\right) \boldsymbol{B}_t^{+}\boldsymbol{P}_Z(t)} \\ &= \sum_{i=1}^{N_Z-1} \frac{\boldsymbol{b}_t(i)\boldsymbol{B}_t^{+}\boldsymbol{Q}_Z(t)}{\left(\sum_{i=1}^{N_Z-1} \boldsymbol{b}_t(i)\right) \boldsymbol{B}_t^{+}\boldsymbol{P}_Z(t)} \\ &= \frac{\left(\sum_{i=1}^{N_Z-1} \boldsymbol{b}_t(i)\right) \boldsymbol{B}_t^{+}\boldsymbol{Q}_Z(t)}{\left(\sum_{i=1}^{N_Z-1} \boldsymbol{b}_t(i)\right) \boldsymbol{B}_t^{+}\boldsymbol{P}_Z(t)}. \end{aligned}$$

□

A.13 Proof of Theorem T-7

Proof. Let the support of T be $\{t_1, \ldots, t_{N_T}\}$ and the support of Z be $\text{supp}(Z) = \{z_1, \ldots, z_{N_Z}\}$. Let the response matrix be given by $\boldsymbol{R}$, which has N_Z rows. Indicator matrix $\boldsymbol{B}_t = \mathbf{1}[\boldsymbol{R} = t]$; $t \in \text{supp}(T)$ has the same row and column dimensions of $\boldsymbol{R}$. Its elements are equal to 1 if the associated element in $\boldsymbol{R}$ is t and 0 otherwise.

Indicator Matrix $\boldsymbol{B}_T$ is generated by stacking matrices $\boldsymbol{B}_t$; $t \in \text{supp}(T)$, that is:

$$\boldsymbol{B}_T = \begin{bmatrix} \boldsymbol{B}_{t_1} \\ \vdots \\ \boldsymbol{B}_{t_{N_T}} \end{bmatrix}.$$

Thus $\boldsymbol{B}_T$ consists of N_T binary sub-matrices $\boldsymbol{B}_t$, each of them having N_Z rows. Our goal is to determine the maximum of $\text{rank}(\boldsymbol{B}_T)$ in terms of N_Z and N_T. The rank of $\boldsymbol{B}_T$ is given by the number of linearly independent rows or columns in $\boldsymbol{B}_T$. Thus, $\text{rank}(\boldsymbol{B}_T)$ must be less or equal than its row dimension, that is, $N_T N_Z$. We can reduce this number further.

First, consider the case where $P(T = t|Z = z) > 0$ for all $z \in \text{supp}(Z)$ and $t \in \textit{supp}(T)$. Thus, each possible value $t \in \textit{supp}(T)$ of the treatment choice must appear at least once in each row of $\boldsymbol{R}$

We investigate the sub-matrix of $\boldsymbol{B}_T$ generated by the first row of each matrix $\boldsymbol{B}_t$ for $t \in \{t_1, \ldots, t_{N_T}\}$, namely:

$$\begin{bmatrix} \boldsymbol{B}_{t_1}[1, \cdot] \\ \boldsymbol{B}_{t_2}[1, \cdot] \\ \vdots \\ \boldsymbol{B}_{t_{N_T}}[1, \cdot] \end{bmatrix}. \tag{A.46}$$

Sub-matrix (A.46) is generated from the first row of $\boldsymbol{R}$, that is $\boldsymbol{R}[1, \cdot]$. Each element of the row $\boldsymbol{R}[1, \cdot]$ takes a value in $\{t_1, \ldots, t_{N_T}\}$. Thus, each column of sub-matrix (A.46) has one and only one element that is equal to 1 while all other elements are zero. Thus, since each treatment value in $\{t_1, \ldots, t_{N_T}\}$ appears at least once in row $\boldsymbol{R}[1,\cdot]$, then the sub-matrix (A.46) has N_T linearly independent rows. Moreover, we have that the sum of each column in sub-matrix (A.46) is one, thus:

$$[1 \, , \, \ldots \, , \, 1] = \sum_{j=1}^{N_T} \boldsymbol{B}_{t_j}[1,\cdot]. \tag{A.47}$$

Note that Equation (A.47) would also hold even if some rows of sub-matrix (A.46) were all zeros. However, all zero rows are ruled out by the assumption $P(T = t|Z = z) > 0$.

Now consider the sub-matrix of $\boldsymbol{B}_T$, generated by the second row of each matrix $\boldsymbol{B}_t$ for $t \in \{t_1, \ldots, t_{N_T}\}$:

$$\begin{bmatrix} \boldsymbol{B}_{t_1}[2,\cdot] \\ \boldsymbol{B}_{t_2}[2,\cdot] \\ \vdots \\ \boldsymbol{B}_{t_{N_T}}[2,\cdot] \end{bmatrix}. \tag{A.48}$$

Sub-matrix (A.48) is generated by N_T linearly independent rows since each value of the treatment choice, i.e. $\{t_1, \ldots, t_{N_T}\}$, appears at least once in the second row of $\boldsymbol{R}$. Moreover, we also have that

$$[1 \, , \, \ldots \, , \, 1] = \sum_{j=1}^{N_T} \boldsymbol{B}_{t_j}[2,\cdot]. \tag{A.49}$$

Equation (A.49) would hold even if some rows of sub-matrix (A.48) were all zeros.

We can use Equation (A.47) to express, for example, the last row of sub-matrix (A.48) as a linear combination of rows in sub-matrix (A.46) and the remaining rows of sub-matrix (A.48):

$$
\begin{aligned}
[1\ ,\ \ldots\ ,\ 1] &= \sum_{j=1}^{N_T} \boldsymbol{B}_{t_j}[2,\cdot] \\
\Rightarrow \boldsymbol{B}_{t_{N_T}}[2,\cdot] &= [1\ ,\ \ldots\ ,\ 1] - \sum_{j=1}^{N_T-1} \boldsymbol{B}_{t_j}[2,\cdot] \\
\Rightarrow \boldsymbol{B}_{t_{N_T}}[2,\cdot] &= \sum_{j=1}^{N_T} \boldsymbol{B}_{t_j}[1,\cdot] - \sum_{j=1}^{N_T-1} \boldsymbol{B}_{t_j}[2,\cdot]. \qquad \text{(A.50)}
\end{aligned}
$$

We can iterate this approach for each i^{th} row of $\{\boldsymbol{B}_{t_1},\ldots,\boldsymbol{B}_{t_{N_T}}\}$. We can then express the i^{th} row of $\boldsymbol{B}_{t_{N_T}}$, such that $i > 1$, as a linear combination of the first rows $\boldsymbol{B}_t[1,\cdot]; t \in \{t_1,\ldots,t_{N_T}\}$ and the i^{th} rows of $\{\boldsymbol{B}_{t_1},\ldots,\boldsymbol{B}_{t_{N_T-1}}\}$. Thus there are $N_Z - 1$ rows in $\boldsymbol{B}_{t_{N_T}}$ that are not linearly independent of the remaining rows of $\boldsymbol{B}_T$. As a consequence, the number of linearly independent rows in $\boldsymbol{B}_T$ is at most $N_Z N_T - (N_Z - 1)$ and therefore $\text{rank}(\boldsymbol{B}_T) \leq N_Z N_T - (N_Z - 1) = 1 + N_Z(N_T - 1)$.

Now suppose that $P(T = t_1|Z = z) = 0$ for some $z \in \text{supp}(Z)$. Equation (A.47) still holds. But these rows are linearly dependent, thereby we have that:

$$
\text{rank}(\boldsymbol{B}_T) \leq 1 + N_Z(N_T - 1) - \sum_{i=1}^{N_Z} \mathbf{1}[P(T = t_1|Z = z_i) = 0]
$$

.

Now suppose that $P(T = t_{\tilde{j}}|Z = z_2) = 0$ for some $\tilde{j} \in \{1,\ldots,N_T\}$. We must have that $P(T = t_{j'}|Z = z_2) > 0$ for some $j' \in \{1,\ldots,N_T\}$ Therefore we can rewrite Equation (A.50) associated with the second rows of each $\boldsymbol{B}_t; t \in \text{supp}(T)$ as:

$$
\begin{aligned}
\Rightarrow \boldsymbol{B}_{t_{j'}}[2,\cdot] &= \sum_{j=1}^{N_T} \boldsymbol{B}_{t_j}[1,\cdot] - \sum_{j \in \{1,\ldots,N_T\}\setminus\{j'\}} \boldsymbol{B}_{t_j}[2,\cdot], \\
&= \sum_{j=1}^{N_T} \boldsymbol{B}_{t_j}[1,\cdot] - \sum_{j \in \{1,\ldots,N_T\}\setminus\{j',\tilde{i}\}} \boldsymbol{B}_{t_j}[2,\cdot].
\end{aligned}
$$

But $\boldsymbol{B}_{t_{\tilde{i}}}[2,\cdot]$ is a row of zeros and is linearly dependent. Therefore it reduces the maximum rank of $\text{rank}(\boldsymbol{B}_T)$. We can apply this analysis for all $z \in \{z_2,\ldots,z_{N_Z}\}$ and for all rows of zeros in $\boldsymbol{B}_t; t \in \text{supp}(T)$. This generates the following restriction:

$$\text{rank}(\boldsymbol{B}_T) \leq 1 + N_Z(N_T - 1) - \sum_{j=1}^{N_T} \sum_{i=1}^{N_Z} \mathbf{1}[P(T = t_j | Z = z_i) = 0]$$

. □

A.14 Additional Identification Results for Strata Probabilities

In this appendix, we present additional results on identification that do not appear in the main text of the paper.

One desirable property of a monotone response matrix is that it consists of all potential response types that are consistent with a monotone property. We term matrix $\boldsymbol{R}$ *complete* if $\boldsymbol{R}$ is an unordered monotone response matrix such that the inclusion of any additional response type to $\boldsymbol{R}$ would violate monotonicity. As demonstrated in Appendix D, the selection of possible response types is not unique. A range of possible complete response types exist for any given N_Z and N_T. Completeness does not necessarily imply that the number of response types is $1 + (N_T - 1)N_Z$. The number of response types in complete response matrices may be bigger. The next theorem demonstrates how to exploit this completeness criteria to identify response-type probabilities.

Theorem T-8. *Consider the IV model (1)–(3), where Z takes values in $\{z_1 \ldots, z_{N_T}\}$ and T takes values in $\{t_1, \ldots, t_{N_T}\}$. Let $\boldsymbol{R}$ be a complete unordered monotone response matrix and let $z \in \text{supp}(Z)$ and $t \in \text{supp}(T)$ such that $P(T = t'|Z = z) \geq P(T = t'|Z = z')$ for $z' \in \text{supp}(Z) \setminus \{z\}$ and $t' \in \text{supp}(T) \setminus \{t\}$ then:*

Response-type probabilities $P(\boldsymbol{S} = \boldsymbol{s})$ for all $\boldsymbol{s} \in \text{supp}(\boldsymbol{S})$ are identified,

and the response-matrix $\boldsymbol{R}$ has the following properties:

1. *$\boldsymbol{R}$ is uniquely determined.*
2. *$\boldsymbol{R}$ consists of $1 + (N_T - 1)N_Z$ response types.*
3. *$\boldsymbol{R}$ generates $\boldsymbol{B}_T$ such that $\text{rank}(\boldsymbol{B}_T) = 1 + (N_T - 1)N_Z$.*

Proof. Our proof exploits the properties of complete matrices. To do so, it is helpful to define some useful notation. We denote the set of possible response types by Φ, which consists of all $N_T^{N_Z}$ possible N_Z-dimensional response-type vectors defined as:

$$\Phi = \{[\tau_1, \ldots, \tau_{N_Z}]' \text{ with elements } \tau_i \text{ such that } \tau_i \in \{t_1, \ldots, t_{N_T}\} \text{ for all } i \in \{1, \ldots, N_Z\}\}. \quad \text{(A.51)}$$

The support of the response variable $\boldsymbol{S}$ is a subset of Φ and can be represented as $\text{supp}(\boldsymbol{S}) = \{\boldsymbol{s} \in \Phi; P(\boldsymbol{S} = \boldsymbol{s}) > 0\}$. Φ represents all possible response types. $\boldsymbol{S}$ are the ones generated the data. Example **E-2** illustrates set Φ for the case of binary instruments with binary treatment choices.

Example **E-2**. Let treatment choice T take only binary values in $\text{supp}(T) = \{t_0, t_1\}$. $Z \in \{z_0, z_1\}$ so the cardinality of the support of Z is $N_Z = |\,\text{supp}(Z)| = 2$. Then set Φ is defined by:

$$\Phi = \{[\tau_1, \tau_2]' \text{ with elements } \tau_i \text{ such that } \tau_i \in \{t_0, t_1\} \text{ for } i \in \{1, 2\}\}.$$

Elements of the set Φ are the 2-dimensional vectors $[\tau_1, \tau_2]'$, such that τ_1, the first element of the vector, can take one of the only two (scalar) values, t_1 or t_0. The second element of the vector $[\tau_1, \tau_2]'$ is τ_2, which also takes one of the only two (scalar) values, t_1 or t_0. Thus, we can enumerate the elements of set Φ as following:

$$\Phi = \left\{ \begin{bmatrix} t_0 \\ t_0 \end{bmatrix}, \begin{bmatrix} t_0 \\ t_1 \end{bmatrix}, \begin{bmatrix} t_1 \\ t_0 \end{bmatrix}, \begin{bmatrix} t_1 \\ t_1 \end{bmatrix} \right\}.$$

We can associate the first element (first row) of each vector in Φ to the counterfactual choice of an agent ω when the instrument is set to z_0, that is, $T_\omega(z_0)$, and the second element (second row) of each vector in Φ to the counterfactual choice of an agent ω when the instrument is set to z_1, that is, $T_\omega(z_1)$. If we assume the standard LATE monotonicity relationship that $\mathbf{1}[T(z_1) = t_1] \geq \mathbf{1}[T(z_0) = t_1]$, we can generate the support of the response variable $\boldsymbol{S}$ as a subset of set Φ that eliminates the response type $[t_1\ ,\ t_0]'$, that is:

$$\text{supp}(\boldsymbol{S}) = \left\{ \begin{bmatrix} t_0 \\ t_0 \end{bmatrix} , \begin{bmatrix} t_0 \\ t_1 \end{bmatrix} , \begin{bmatrix} t_1 \\ t_1 \end{bmatrix} \right\} \subset \Phi.$$

A.14.1 Definition of Complete Response Matrices

The response matrix $\boldsymbol{R}$ consists of the matrix that includes all of the response types in $\text{supp}(\boldsymbol{S})$. Response matrix $\boldsymbol{R}$ is complete if:

1. Response matrix $\boldsymbol{R}$ is unordered monotone;

2. The matrix $\widetilde{\boldsymbol{R}} = [\boldsymbol{R}, \boldsymbol{s}]$ is not an unordered monotone response for each $\boldsymbol{s} \in \Phi \backslash \text{supp}(\boldsymbol{S})$.

We term a response type $\boldsymbol{s} \in \Phi$ admissible relative to an unordered monotone response $\boldsymbol{R}$ if:

1. Response type $\boldsymbol{s}$ is not in $\boldsymbol{R}$, that is, $\boldsymbol{s} \notin \text{supp}(\boldsymbol{S})$;

2. The matrix $\widetilde{\boldsymbol{R}} = [\boldsymbol{R}\ ,\ \boldsymbol{s}]$ is still an unordered monotone response.

Thus, a response matrix $\boldsymbol{R}$ is complete if no response type $\boldsymbol{s}$ in $\Phi \setminus \text{supp}(\boldsymbol{S})$ is admissible.

A.14.2 Number of Treatment Values is Each Row of the Response Matrix

Let $r(i,t)$ be the number of elements t in the i^{th} row of $\boldsymbol{R}$. Thus the number $r(i,t)$ can be obtained by the sum of the binary element in the i^{th} row of matrix $\boldsymbol{B}_t$ across columns $j \in \{1, \ldots, N_{\boldsymbol{S}}\}$:

$$r(i,t) = \sum_{j=1}^{N_{\boldsymbol{S}}} \boldsymbol{B}_t[i,j],$$

where $\boldsymbol{B}_t$ denotes the indicator matrix $B_t = \mathbf{1}[\boldsymbol{R} = t]$.

We use $m(t)$ to denote the *row-index* i in $\boldsymbol{R}$ whose number of elements t is maximum, that is:

$$m(t) = \underset{i \in \{1, \ldots, N_Z\}}{\text{argmax}}\ r(i,t). \tag{A.52}$$

Thus, $m(t) \in \{1, \ldots, N_Z\}$ for all $t \in \text{supp}(T)$ and more than one treatment can have its maximum sum at the same row, i.e., $m(t) = m(t'); t, t' \in \text{supp}(T)$. Nevertheless, it cannot be that all treatment values $t \in \text{supp}(T)$ have their maximum sum at the same row in $\boldsymbol{R}$ (see Lemma **L-20** below).

Let $\boldsymbol{s}(i)$ denotes the treatment status in the i^{th} row of a response type $\boldsymbol{s} \in \Phi$.

A.14.3 Overview of the Properties of Unordered Monotone and Complete Matrices

In order to prove **T-8** below, we rely on properties of unordered monotone responses and also complete response matrices. Specifically, Lemmas **L-18**–**L-21** investigate properties of unordered monotone response matrices while Lemmas **L-22**–**L-24** focus on properties of unordered monotone response matrices that are *complete*. We list the associated Lemmas and their consequences to facilitate the understanding of the theorem.

- Properties of unordered monotone response matrices:
  1. Lemma **L-18** connects the row-sums of treatment values with the possible values that elements of unordered monotone response matrices can take.
  2. Lemma **L-19** states a property of response types of unordered monotone response matrices when the same row gives the maximum sum of treatment values for more than one treatment status.
  3. Lemma **L-20** states a condition of row-sums of treatment values across response types.
  4. Lemma **L-21** is an auxiliary lemma used in the main proof.
- Properties of *complete* unordered monotone response matrices:
  1. Lemma **L-22** states that if $\boldsymbol{R}$ is complete then $\boldsymbol{R}$ must have N_T response types whose elements within a response type are the same for each value t in the support of T.

2. Lemma **L-23** characterizes the non-zero binary vectors in $\boldsymbol{B}_t; t \in \text{supp}(T)$ for complete response matrices $\boldsymbol{R}$.

3. Lemma **L-24** states that the non-zero binary vectors in each $\boldsymbol{B}_t; t \in \text{supp}(T)$ have full rank if $\boldsymbol{R}$ is complete.

A.14.4 Comments of the use of Lemmas and our Main Proof

Our main proof is constructive and relies on each of the Lemmas **L-18**–**L-21**. Lemma **L-22** is used as a starting point in the construction of the response matrix that complies with the assumptions of Theorem **T-8**. The most important property of complete matrices exploited in our main proof is stated in Lemma **L-23**. Namely if $\boldsymbol{R}$ is complete then each $\boldsymbol{B}_t; t \in \text{supp}(T)$ must have N_Z distinct non-zero binary vectors.

We use Lemma **L-19** to characterize an essential property of unordered monotone response matrices. The purpose of Lemma **L-21** is used to justify a simplified notation that facilitates our exposition of the proof. Lemma **L-24** gives a rank condition of complete matrices. We use this condition to generate the rank condition of Corollary **C-1**, which identifies response-type probabilities.

Lemma **L-18** is a restriction of unordered response matrices within response types. Lemma **L-20** is a restriction of unordered response matrices across response types. Both restrictions are useful tools for proofs based on contradictions. Lemma **L-18** and Lemma **L-20** are used throughout our analysis.

A.14.5 Lemma L-18

Lemmas **L-18**–**L-20** describe useful properties of unordered monotone response matrices that are helpful in proving the main theorem.

Lemma L-18. If response matrix $\boldsymbol{R}$ is unordered monotone, and $\boldsymbol{s}$ is a response type that belongs to $\boldsymbol{R}$, and if $\boldsymbol{s}(i) = t$, and $r(i', t) \geq r(i, t)$, then $\boldsymbol{s}(i') = t$ for all $i' \in \{1, \ldots, N_Z\}$. In particular, if $\boldsymbol{s}(i) = t$, for some $i \in \{1, \ldots, N_Z\}$ then $\boldsymbol{s}(m(t)) = t$.

Proof. Suppose that $\boldsymbol{s}(i) = t$, $r(i', t) \geq r(i, t)$ but $\boldsymbol{s}(i') \neq t$. Then there must be a column j in $\boldsymbol{B}_t$ such that $\boldsymbol{B}_t[i, j] = 0$ and $\boldsymbol{B}_t[i', j] = 1$. Let j' be the column that represents response type $\boldsymbol{s}$, that is, $\boldsymbol{s} = \boldsymbol{R}[\cdot, j']$. But $\boldsymbol{s}(i) = t$, implies that $\boldsymbol{B}_t[i, j'] = 1$ and $\boldsymbol{s}(i') \neq t$ implies that $\boldsymbol{B}_t[i', j'] = 0$. Now consider the 2×2 sub-matrix in $\boldsymbol{R}$ defined by rows i, i' and columns j, j' :

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

which takes the forbidden patterns of Equation (52). Therefore $\boldsymbol{B}_t$ cannot be lonesum. As a consequence, $\boldsymbol{R}$ cannot be an unordered monotone response, which contradicts the statement of the lemma. In particular, if $\boldsymbol{s}(i) = t$, for some $i \in \{1, \ldots, N_Z\}$ then $\boldsymbol{s}(m(t)) = t$ because $m(t)$ is the row-index that provides the maximum sum of elements t among all $r(i', t); i' \in \{1, \ldots, N_Z\}$. □

A.14.6 Lemma L-19

Lemma L-19. If response matrix $\boldsymbol{R}$ is unordered monotone and $m(t) = m(t')$ for some $t, t' \in \text{supp}(T)$ such that $t \neq t'$, then no response type can take both values t and t'. Namely, there is no column $j \in \{1, \ldots, N_{\boldsymbol{S}}\}$ such that $\boldsymbol{R}[i, j] = t$ and $\boldsymbol{R}[i', j] = t'$ for any two rows $i, i' \in \{1, \ldots, N_Z\}$.

Proof. Suppose there exists a column $j \in \{1, \ldots, N_{\boldsymbol{S}}\}$ such that $\boldsymbol{R}[i, j] = t$ and $\boldsymbol{R}[i', j] = t'$ for some rows $i, i' \in \{1, \ldots, N_Z\}$. Then, by Lemma **L-18**, we must have that $\boldsymbol{R}[m(t), j] = t$ and $\boldsymbol{R}[m(t'), j] = t'$, which is impossible since by hypothesis $m(t') = m(t)$. □

A.14.7 Lemma L-20

Lemma L-20. If response matrix $\boldsymbol{R}$ is unordered monotone then for any two rows $i, i' \in \{1, \ldots, N_Z\}$, it cannot be the case that $r(i', t) \geq r(i, t)$ for all $t \in \textit{supp}(T)$. In particular, it cannot be that case that $m(t)$ is equal to some $i \in \{1, \ldots, N_Z\}$ for all $t \in \textit{supp}(T)$.

Proof. $\boldsymbol{R}$ is a response matrix, so columns and rows must differ in at least one element. According to Lemma **L-18**, if $r(i', t) \geq r(i, t)$ and $\boldsymbol{s}(i) = t$, then $\boldsymbol{s}(i') = t$. But the number of elements in a row is $N_{\boldsymbol{S}}$. Thus if $r(i', t) \geq r(i, t)$ for all $t \in supp(T)$, it must be the case that $r(i', t) = r(i, t)$ for all $t \in supp(T)$. But for the rows to differ, it must be the case that exists some columns $j, j' \in \{1, \ldots, N_{\boldsymbol{S}}\}$ such that

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

for at least one treatment $t \in \text{supp}(T)$, which characterizes the forbidden pattern of Equation (52) and violates the assumed monotonicity of $\boldsymbol{R}$. □

A.14.8 Lemma L-21

Lemma L-21. Let $\boldsymbol{R}$ be an unordered monotone response matrix, such that $\text{supp}(T) = \{t_1, t_2, \ldots, t_{N_T}\}$ and $\text{supp}(Z) = \{z_1, z_2, \ldots, z_{N_Z}\}$. Let the first row of $\boldsymbol{R}$ be the row where the sum of elements t_1 is maximum, that is, $m(t_1) = 1$. Also let the last row of $\boldsymbol{R}$ be the row that generates that the largest row-sum of treatment values $t_2, \ldots, t_{N_T}$, that is, $m(t_2) = m(t_3) = \ldots = m(t_{N_T}) = N_Z$. Then the last row of $\boldsymbol{R}$ is also the row that generates the minimum sum of elements t_1.

Proof. Suppose not. So it must be the case that the row-index that gives the minimum row-sum of elements t_1 is i such that $i < N_Z$. Thus we have that $r(i, t_1) \leq r(N_Z, t_1)$. Now take the sub-matrix of $\boldsymbol{R}$ generated by rows i and N_Z which is also an unordered monotone matrix. But in this case we would have that $r(N_Z, t)$ would take the maximum value for all $t \in \text{supp}(T)$, which is impossible due to Lemma **L-20**. □

A.14.9 Lemma L-22

Lemmas **L-22**–**L-24** describe useful properties of unordered monotone response matrices that are *complete*.

Lemma L-22. If an an unordered monotone response matrix $\boldsymbol{R}$ is complete then each N_Z-dimensional vector of elements $t \in \text{supp}(T)$, that is $\iota_{N_Z} \cdot t$ belongs to $\boldsymbol{R}$.

Proof. According to condition (iii) of Theorem T-3, $\boldsymbol{R}$ is an unordered response if and only if the following forbidden 2×2 sub-matrices do not belong to $\boldsymbol{R}$:

$$\begin{pmatrix} t & t' \\ t'' & t \end{pmatrix} \text{ or } \begin{pmatrix} t' & t \\ t & t'' \end{pmatrix}, \text{where } t' \neq t \text{ and } t'' \neq t.$$

Treatment status varies within vectors of both forbidden sub-matrices. But the treatment status does not vary in response type $\iota_{N_Z} \cdot t$. Thus if $\boldsymbol{R}$ is an unordered response matrix and $\iota_{N_Z} \cdot t$ does not belong to $\boldsymbol{R}$, then $\iota_{N_Z} \cdot t$ is admissible. But $\boldsymbol{R}$ is complete and does not allow any further admissible response types. Thereby it must be the case that $\iota_{N_Z} \cdot t$ already belongs to $\boldsymbol{R}$. □

A.14.10 Lemma L-23

Recall that $\Sigma(t)$ is the set of non-zero vectors in $\boldsymbol{B}_t$ of $\boldsymbol{R}$. We can state the following property of $\Sigma(t)$ for complete matrices:

Lemma L-23. Let $\boldsymbol{R}$ be a complete unordered monotone response matrix. Then each set $\Sigma(t); t \in \text{supp}(T)$ has N_Z distinct non-zero binary vectors of dimension N_Z. Moreover, let $\boldsymbol{C}_t$ be the matrix generated by the binary vectors in $\Sigma(t)$. Then $\boldsymbol{C}_t$ is a lonesum matrix that is equivalent to a N_Z-dimensional square lower triangular matrix. In particular, each set $\Sigma(t); t \in \text{supp}(T)$ has N_Z distinct non-zero binary vectors of dimension N_Z whose column-sum is $1, 2, \ldots, N_Z$.

Proof. Without loss of generality, suppose that the rows $\{1, \ldots, N_Z\}$ of $\boldsymbol{R}$ are ordered in increasing values of the sum of the treatment statuses t in $\text{supp}(T)$, that is,

$$r(1,t) \leq r(2,t) \leq \ldots \leq r(N_Z, t). \tag{A.53}$$

Response matrix $\boldsymbol{R}$ is unordered monotone, so $\boldsymbol{B}_t$ is lonesum. According to Lemma **L-4**, any selection of vectors in $\boldsymbol{B}_t$ is lonesum. In particular, $\boldsymbol{C}_t$ is the selection of all unique non-zeros vectors in $\boldsymbol{B}_t$. Therefore, $\boldsymbol{C}_t$ is lonesum.

According to Lemma **L-18**, for any $\boldsymbol{s} \in \text{supp}(\boldsymbol{S})$, if $r(i', t) \geq r(i, t)$ and $\boldsymbol{s}(i) = t$, then $\boldsymbol{s}(i') = t$. This implies that if $r(i', t) \geq r(i, t)$,

$$\boldsymbol{B}_t[i, j] = 1 \Rightarrow \boldsymbol{B}_t[i', j] = 1 \text{ for any } j \in \{1, \ldots, N_{\boldsymbol{S}}\}. \tag{A.54}$$

But $\boldsymbol{C}_t$ is a sub-matrix of $\boldsymbol{B}_t$, thus implication A.54 also holds for elements in that submatrix. Moreover, under the ranking of row-sums (A.53), $\boldsymbol{C}_t$ is a lower triangular matrix.

To finish the lemma we need to prove that if $\boldsymbol{R}$ is complete, then $\boldsymbol{C}_t$ has N_Z binary vectors. Suppose not. Specifically, suppose that the binary vector that takes values 1 for rows-indexes equal to or greater than i and zero otherwise does not belong to $\boldsymbol{C}_t$. This condition on $\boldsymbol{C}_t$ translates to the following assertion on response matrix $\boldsymbol{R}$:

$$\begin{aligned}&\text{For row-index } i \in \{1, \ldots, N_Z\}, \text{ there is no response type } \boldsymbol{s} \text{ of } \boldsymbol{R} \text{ that satisfies:}\\ &(1)\ \boldsymbol{s}(i) = t, \text{ and } (2)\ \boldsymbol{s}(i') = t \text{ for all } i' > i, \text{ and } (3)\ \boldsymbol{s}(i') \neq t \text{ for all } i' < i.\end{aligned} \tag{A.55}$$

Our goal is to show that the condition in (A.55) cannot occur when $\boldsymbol{R}$ is complete. Indeed, if $\boldsymbol{R}$ is complete then it has to be the case that the response type described in (A.55) is not admissible. Then it must be the case that response type $\boldsymbol{s}$ defined in (A.55) generates the forbidden pattern (52). But $\boldsymbol{s}$ cannot generate a forbidden pattern in $\boldsymbol{B}_t$. Indeed $\boldsymbol{s}(i) = t$ and $\boldsymbol{s}(i') = t$ for all $i' > i$, in accordance with Lemma **L-18**.

So it must be the case that the inclusion of a response type $\boldsymbol{s}$ generates a forbidden pattern for each treatment t' that differs from t. In particular the response type $\boldsymbol{s}$ that takes value $\boldsymbol{s}(i') = t'$ for $i' < i$ and $\boldsymbol{s}(i'') = t$ for $i'' \geq i$ is not admissible. Thus, there must exist a column j such that the following forbidden pattern occurs:

$$\begin{pmatrix} \mathbf{1}[\boldsymbol{R}[i',j]=t'] & \mathbf{1}[\boldsymbol{s}(i')=t'] \\ \mathbf{1}[\boldsymbol{R}[i'',j]=t'] & \mathbf{1}[\boldsymbol{s}(i'')=t'] \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{A.56}$$

But, according to Lemma **L-22**, if $\boldsymbol{R}$ is complete, then treatment t' must appear at least once in each row. In particular, t' must appear at row i'. Thus there must be a column j' such that $\boldsymbol{R}[i',j']=t'$. But $\boldsymbol{R}$ is a monotone response, so $\boldsymbol{B}_{t'}$ is lonesum and therefore we must have that:

$$\begin{pmatrix} \mathbf{1}[\boldsymbol{R}[i',j]=t'] & \mathbf{1}[\boldsymbol{R}[i',j']=t'] \\ \mathbf{1}[\boldsymbol{R}[i'',j]=t'] & \mathbf{1}[\boldsymbol{R}[i'',j']=t'] \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, \tag{A.57}$$

where we apply **L-22** to get the upper righthand element and **L-18** to get the lower righthand element.

Equation (A.57) implies that $r(i',t') < r(i,t')$. This condition holds for each $t' \in \text{supp}(T)\setminus\{t\}$. But we also have that $r(i',t) \leq r(i,t)$. Thus the i^{th} row-sum is less than equal than i^{th} row-sum for all treatment statuses. This condition is impossible due to Lemma **L-20**. □

Lemma **L-22** is a particular consequence of Lemma **L-23** above. Nevertheless, we have stated and proved Lemma **L-22** because it is used in Lemma **L-23**. Indeed, Lemma **L-22** allows us to use the fact that if $\boldsymbol{R}$ is complete then each treatment choice $t \in \text{supp}(T)$ must appear at least once in each row of response matrix $\boldsymbol{R}$.

A.14.11 Lemma L-24

Lemma L-24. Let $\boldsymbol{R}$ be a complete unordered monotone response matrix. Then it follows that $rank(\boldsymbol{B_t}) = N_Z$ for all $t \in \text{supp}(T)$.

Proof. By **L-23**, $\boldsymbol{C}_t$ is equivalent to a lower triangular $N_Z \times N_Z$ squared matrix whose column sums are $1, 2, \cdots, N_Z$. Thus, $\boldsymbol{C_t}$ must have full rank, that is, $\text{rank}(\boldsymbol{C}_t) = N_Z$. Since

$\boldsymbol{B}_t$ includes the columns of C_t, then $\text{rank}(\boldsymbol{B}_t) \geq \text{rank}(\boldsymbol{C}_t) = N_Z$. But the row-dimension of $\boldsymbol{B}_t$ is N_Z and thereby $\text{rank}(\boldsymbol{B}_t) \leq N_Z$. Therefore $\text{rank}(\boldsymbol{B}_t) = N_Z$. □

A.14.12 The Proof of Theorem T-8

Proof. We employ the notation used throughout this paper, which defines the support of treatment choice T and instrumental variable Z as $\text{supp}(T) = \{t_1, t_2, \ldots, t_{N_T}\}$ and $\text{supp}(Z) = \{z_1, z_2, \ldots, z_{N_Z}\}$. The theorem assumes that there is $z \in \text{supp}(Z)$ such that $P(T = t'|Z = z) \geq P(T = t'|Z = z')$ for $z' \in \text{supp}(Z) \setminus \{z\}$. But $\boldsymbol{R}$ is unordered monotone. Thus, the row in $\boldsymbol{R}$ associated with the value z of instrumental variable Z is also the row that gives the maximum sum of elements t'. But $P(T = t'|Z = z) \geq P(T = t'|Z = z')$ for all $t' \in \text{supp}(T) \setminus \{t\}$. Thus we have that $m(t') = m(t'')$ for all $t', t'' \in \text{supp}(T) \setminus \{t\}$. By Lemma L-20, it must be the case that $m(t) \neq m(t')$.

Without loss of generality, we introduce the following useful notation:

(a) $m(t_2) = m(t_3) = \ldots = m(t_{N_T}) = N_Z$;

(b) $m(t_1) = 1$;

(c) $r(1, t_1) \geq r(2, t_1) \geq \ldots \geq r(N_Z, t_1)$.

Item (a) specifies that the value z of the instrumental variable Z stated in the premise of the theorem is placed as the last row of $\boldsymbol{R}$. Item (a) complies with the theorem's assumption, stating that the last row has the maximum sum of treatment status $t_2, \ldots, t_{N_T}$. Item (b) states that the treatment choice t_1 is the one that does not takes the maximum sum in row N_Z, in compliance with Lemma L-20. Instead t_1 has its maximum sum at the first row of $\boldsymbol{R}$. Item (c) states that the rows of $\boldsymbol{R}$ are ordered by decreasing sum of elements t_1. For this to happen, it is necessary that the last row of $\boldsymbol{R}$ must be the row whose sum of elements t_1 is minimum, which holds according to Lemma L-21. This specification facilitates the proof of our claim and is assumed without loss of generality.

Our proof is constructive and is based on the following steps:

1. Lemma **L-22** implies that the vector $\iota_{N_Z} \cdot t_1$ belongs to $\boldsymbol{R}$. Our count of response types in $\boldsymbol{R}$ is now 1.

2. Lemma **L-22** also implies that the vector $\iota_{N_Z} \cdot t, t \in \{t_2, \ldots, t_{N_T}\}$ must also belong to $\boldsymbol{R}$. Our count of response types in $\boldsymbol{R}$ is now $1 + (N_T - 1)$.

3. According to Lemma **L-19**, there is no response type $\boldsymbol{s}$ in $\boldsymbol{R}$ that takes two values of treatment choices in $\{t_2, \ldots, t_{N_Z}\}$. Thus any response type that takes more than one value of treatment choices in $\text{supp}(T)$ must take only two values, one being t_1 and another being some $t' \in \{t_2, \ldots, t_{N_Z}\}$.

4. According to Lemma **L-18**, and the previous remark, if a response type of an unordered response matrix $\boldsymbol{R}$ takes more than one value of the treatment choice, then it must be the case that the response type is of the type described below:

$$\boldsymbol{s} = [t_1 \ , \ t_1 \ , \ \ldots \ , \ t_1 \ , \ t', \ \ldots \ , \ t']' \text{ where } t' \in \{t_2, \ldots, t_{N_T}\}$$

.

5. But according to Lemma **L-23**, each $\Sigma(t'); t' \in \{t_2, \ldots, t_{N_T}\}$ has N_Z distinct non-zero binary vectors of dimension N_Z whose column sum is $1, 2, \ldots, N_Z$. In particular, each $\Sigma(t'); t' \in \{t_2, \ldots, t_{N_T}\}$ must have a binary vector $\xi \in \Sigma(t'); t' \in \{t_2, \ldots, t_{N_T}\}$ whose sum is 1. According to item 3 above, the response type that generates this binary vector must be:

$$\boldsymbol{s} = [t_1 \ , \ t_1 \ , \ \ldots \ , \ t_1, \ t']' \text{ where } t' \in \{t_2, \ldots, t_{N_T}\}$$

. This adds $N_T - 1$ additional response types to $\boldsymbol{R}$. Our total count is $1 + 2 \cdot (N_T - 1)$.

6. Each $\Sigma(t'); t' \in \{t_2, \ldots, t_{N_T}\}$ must have a binary vector $\xi \in \Sigma(t'); t' \in \{t_2, \ldots, t_{N_T}\}$ whose column sum is 2. According to item 3 above, the response type that generates this binary vector must be:

$$\boldsymbol{s} = [t_1 \ , \ t_1 \ , \ \ldots \ , \ t_1, \ t', \ t']' \text{ where } t' \in \{t_2, \ldots, t_{N_T}\}$$

. This adds $N_T - 1$ additional response types to $\boldsymbol{R}$. Our total count is $1 + 3(N_T - 1)$.

7. If we iterate the process we have that the total number of response types in $\boldsymbol{R}$ is $N_{\boldsymbol{S}} = 1 + N_Z(N_T - 1)$. In addition, we have that $\Sigma(t_1)$ is the square upper triangular binary matrix as required by Lemma **L-23**.

8. $\boldsymbol{R}$ is complete as it exhausts all possible response types of the the type described in item 3 above.

It remains to prove the rank of $\boldsymbol{B}_T$ is equal to $1 + N_Z(N_T - 1)$. Binary matrix $\boldsymbol{B}_T$ is defined by stacked $\boldsymbol{B}_t$ matrices such that $t \in \{t_1, \ldots, t_{N_Z}\}$. That is,

$$\boldsymbol{B}_T = [\boldsymbol{B}'_{t_1}\ ,\ \boldsymbol{B}'_{t_2}\ ,\ \ldots\ ,\ \boldsymbol{B}'_{t_{N_T}}].'$$

Consider the sub-matrix of $\boldsymbol{B}_T$ defined by

$$\tilde{\boldsymbol{B}} = \begin{bmatrix} \boldsymbol{B}_{t_2} \\ \vdots \\ \boldsymbol{B}_{t_{N_T}} \end{bmatrix}. \tag{A.58}$$

But each response type $\boldsymbol{s} \in \text{supp}(\boldsymbol{S})$ of $\boldsymbol{R}$ is one of the types:

$$\boldsymbol{s} = [t\ ,\ t\ ,\ \ldots\ ,\ t]'; t \in \{t_1, \ldots, t_{N_T}\},$$

$$\text{or } \boldsymbol{s} = [t_1\ ,\ t_1\ ,\ \ldots\ ,\ t_1\ ,\ t',\ \ldots\ ,\ t']' \text{ where } t' \in \{t_2, \ldots, t_{N_T}\}.$$

Thus, by Lemma **L-19**, if $B_{t'}[i, j] = 1$ for some column $j \in \{1, \ldots, N_{\boldsymbol{S}}\}$ and for some $t' \in \{t_2, \ldots, t_{N_T}\}$, then we have that $B_{t''}[i', j] = 0$ for any $t'' \in \{t_2, \ldots, t_{N_T}\}$ such that $t'' \neq t'$ and for any $i' \in \{1, \ldots, N_Z\}$. Thus $\text{rank}(\tilde{\boldsymbol{B}}) = \sum_{k=2}^{N_T} \text{rank}(\boldsymbol{B}_{t_k})$. But by Lemma **L-24**, the column rank of each $\boldsymbol{B}_t$ is equal to N_Z. Thereby $\text{rank}(\tilde{\boldsymbol{B}}) = (N_T - 1)N_Z$. Now the column in $\tilde{\boldsymbol{B}}$ associated with the response type $\iota_{N_Z} \cdot t_1$ is zero. Thus $\text{rank}(\boldsymbol{B}_T) \geq 1 + \text{rank}(\tilde{\boldsymbol{B}}) = 1 + (N_T - 1)N_Z$. But, according to Theorem **T-7** in the main paper, the rank of a matrix

$\boldsymbol{B}_T$ cannot be bigger than $1 + (N_T - 1)N_Z$. Thus, $\text{rank}(\boldsymbol{B}_T) = 1 + (N_T - 1)N_Z$. But $1 + (N_T - 1)N_Z$ is also the number of columns in $\boldsymbol{R}$, so according to Corollary **C-1**, the response-type probabilities are identified. □

A.15 An Alternative Form for T-8

Alternative conditions can be used to identify $P_{\boldsymbol{S}}$. We state one set of conditions as Theorem **T-8**′.

Theorem T-8. ′(Alt. Version of **T-8**) *Let* $\boldsymbol{R}$ *be a complete unordered monotone response matrix such that each value* $z \in \{z_1, \ldots, z_N\}$ *maximizes the propensity score* $P(T = t|Z = z)$ *for a single treatment choice* $t \in \{t_1, \ldots, t_N\}$*. Now let* $t, \tilde{t} \in \text{supp}(T)$ *and* $z', z'', z''' \in \text{supp}(T)$ *such that:*

$$P(T = t|Z = z') \geq P(T = t|Z = z'') \text{ and}$$
$$P(T = \tilde{t}|Z = z''') \geq P(T = \tilde{t}|Z = z'')$$
$$\Rightarrow P(T = \tilde{t}|Z = z') \geq P(T = \tilde{t}|Z = z'').$$

Then we have that response-type probabilities $P(\boldsymbol{S} = \boldsymbol{s})$ *for all* $\boldsymbol{s} \in \text{supp}(\boldsymbol{S})$ *are identified and the response-matrix* $\boldsymbol{R}$ *has the following properties:*

1. $\boldsymbol{R}$ *is uniquely determined.*

2. $\boldsymbol{R}$ *consists of* $1 + (N_T - 1) \cdot N_Z$ *response types.*

3. $\boldsymbol{R}$ *is such that* $\text{rank}(\boldsymbol{B}_T)$ *is also equal to* $1 + (N_T - 1) \cdot N_Z$.

Proof. See Heckman and Pinto (2015). □

A.16 Another Alternative Form of T-8

Theorem T-8. ″(Alt. Version of **T-8**) *Consider the the IV model* (1)–(3) *where* Z *takes values in* $\{z_1 \ldots, z_{N_T}\}$ *and* T *takes values in* $\{t_1, \ldots, t_{N_T}\}$*. Assume the following conditions:*

(i) *Response matrix* $\boldsymbol{R}$ *is unordered monotone.*

(ii) $1 > P(T = t|Z = z) > 0$ *for all* $z \in \text{supp}(Z), t \in \text{supp}(T)$.

(iii) $N_{\boldsymbol{S}} = 1 + (N_T - 1)N_Z$.

(iv) *Response matrix* $\boldsymbol{R}$ *is* complete, *i.e., the inclusion of any additional response type to* $\boldsymbol{R}$ *would violate monotonicity.*

If these four conditions hold, then response-type probabilities $P(\boldsymbol{S} = \boldsymbol{s}); \boldsymbol{s} \in \text{supp}(\boldsymbol{S})$ *are identified.*

Proof. See Heckman and Pinto (2015). □

The condition requiring uniform directions of response to choices outside of t to variation in instruments is quite strong. The general lesson is that additional restrictions beyond the standard IV conditions and monotonicity **A-3** are required to identify $P_{\boldsymbol{S}}$ but not for mean counterfactuals.

□

A.17 Additional Identification Results for Counterfactual Outcomes (T-9,C-5)

Let $\mathscr{T} \subset \text{supp}(T)$ be a subset of treatment choice values. We use $E(Y(\mathscr{T})|\boldsymbol{S} = \boldsymbol{s})$ to denote the weighted average of counterfactual outcomes $E(Y(t)|\boldsymbol{S} = \boldsymbol{s})$ across $t \in \mathscr{T}$:

$$E(Y(\mathscr{T})|\boldsymbol{S} = \boldsymbol{s}) = \sum_{t \in \mathscr{T}} E(Y(t)|\boldsymbol{S} = \boldsymbol{s}) \frac{P(T = t|\boldsymbol{S} = \boldsymbol{s})}{P(T \in \mathscr{T}|\boldsymbol{S} = \boldsymbol{s})}, \tag{A.59}$$

$$\text{where} \quad P(T \in \mathscr{T}|\boldsymbol{S} = \boldsymbol{s}) = \sum_{t \in \mathscr{T}} P(T = t|\boldsymbol{S} = \boldsymbol{s}). \tag{A.60}$$

A subset of particular interest is $\bar{t} = \text{supp}\, T \setminus \{t\}$, which stands for the set of all treatment choices except t. $T \in \bar{t}$ stands for the event of not choosing t. Let $\mathscr{S} \subset \text{supp}\, \boldsymbol{S}$ be a subset of response types. We use $E(Y(\mathscr{T})|\boldsymbol{S} \in \mathscr{S})$ to denote the weighted average of $E(Y(\mathscr{T})|\boldsymbol{S} = \boldsymbol{s})$

across the response types $\boldsymbol{s} \in \mathscr{S}$:

$$E(Y(\mathscr{T})|\boldsymbol{S} \in \mathscr{S}) = \sum_{\boldsymbol{s} \in \mathscr{S}} E(Y(\mathscr{T})|\boldsymbol{S} = \boldsymbol{s}) \frac{P(\boldsymbol{S} = \boldsymbol{s})}{P(\boldsymbol{S} \in \mathscr{S})} \tag{A.61}$$

Response-type subsets of interest are t-Switchers for which $\Sigma_t(i); i \in \{1, \ldots, N_Z\}$. Under this notation, we state the following identification result:

Theorem T-9. *Consider the the IV model (1)–(3) in which unordered monotonicity* ***A-3*** *holds. Let $t \in \text{supp}(T)$ and $i \in \{1, \ldots, N_Z - 1\}$ such that if $\Sigma_t(i) \cap \Sigma_{t'}(i') \neq \emptyset$, for some $t' \in \text{supp}(T) \setminus t$ and $i' \in \{1, \ldots, N_Z - 1\}$, then $\Sigma_{t'}(i') \subset \Sigma_t(i)$. Under these conditions, $E(Y(t) - Y(\bar{t})|\boldsymbol{S} \in \Sigma_t(i))$ is identified by:*

$$E(Y(t) - Y(\bar{t})|\boldsymbol{S} \in \Sigma_t(i)) = \frac{\boldsymbol{b}_t(i)\boldsymbol{B}_t^{+}\boldsymbol{Q}_Z(t)}{\boldsymbol{b}_t(i)\boldsymbol{B}_t^{+}\boldsymbol{P}_Z(t)} - \frac{\boldsymbol{b}_t(i)\sum_{t' \in \bar{t}} \left(\boldsymbol{B}_{t'}^{+}\boldsymbol{Q}_Z(t')\right) \odot \left(\boldsymbol{B}_{t'}'\boldsymbol{Pr_Z}\right)}{\boldsymbol{b}_t(i)\sum_{t' \in \bar{t}} \left(\boldsymbol{B}_{t'}^{+}\boldsymbol{P}_Z(t')\right) \odot \left(\boldsymbol{B}_{t'}'\boldsymbol{Pr_Z}\right)}$$

where $\boldsymbol{Pr_Z} = [P(Z = z_1), \ldots, P(Z = z_{N_Z})]'$.

Proof. See Section A.18 in this Appendix. □

Theorem **T-9** considers a response-type set $\Sigma_t(i)$ such that $n \in \{1, \ldots, N_Z - 1\}$. This implies that $\Sigma_t(i)$ is partition set of t-Switchers and each response type in $\Sigma_t(i)$ must contain choice t but also choices other than t. **T-9** elicits a coarse property of the set $\Sigma_t(i)$. If a set $\Sigma_{t'}(i')$ for $t' \neq t$ shares any of the response types in $\Sigma_t(i)$ then $\Sigma_t(i)$ must contain $\Sigma_{t'}(i')$. The set $\Sigma_{t_a}(2)$ of the response matrix in Table 3 provides an example of this condition because $\Sigma_{t_a}(2) = \Sigma_{t_b}(1) \cup \Sigma_{t_c}(1)$ where $\Sigma_{t_a}(2) = \{s_2, s_3\}$, $\Sigma_{t_b}(1) = \{s_2\}$, and $\Sigma_{t_c}(1) = \{s_3\}$. **T-9** renders the following expression (see Appendix A.19 for derivation):

$$E(Y(t_a)|t_a \in \Sigma_{t_a}(2)) = \frac{E(Y \cdot \mathbf{1}[T = t_a]|Z = z_{no}) - E(Y \cdot \mathbf{1}[T = t_a]|Z = z_{bc})}{P(T = t_a|Z = z_{no}) - P(T = t_a|Z = z_{bc})},$$

$$E(Y(\bar{t}_a)|\boldsymbol{S} \in \Sigma_{t_a}(2)) =$$

$$\left(\frac{\left(E(Y \cdot \mathbf{1}[T = t_c]|Z = z_{bc}) - E(Y \cdot \mathbf{1}[T = t_c]|Z = z_{no})\right) + \left(E(Y \cdot \mathbf{1}[T = t_b]|Z = z_{bc}) - E(Y \cdot \mathbf{1}[T = t_b]|Z = z_{no})\right)}{\left(P(T = t_c|Z = z_{bc}) - P(T = t_c|Z = z_{no})\right) + \left(P(T = t_b|Z = z_{bc}) - P(T = t_b|Z = z_{no})\right)}\right)$$

$$\text{also equivalent to } E(Y(\bar{t}_a)|\boldsymbol{S} \in \Sigma_{t_a}(2)) = \frac{\left(E(Y \cdot \mathbf{1}[T \neq t_a]|Z = z_{bc}) - E(Y \cdot \mathbf{1}[T \neq t_a]|Z = z_{no})\right)}{\left(P(T \neq t_a|Z = z_{bc}) - P(T \neq t_a|Z = z_{no})\right)}.$$

Corollary **C-5** simply extends **T-9** to each partition set $\Sigma_t(i)$ of t-Switchers.

Corollary C-5. *Consider the the IV model (1)–(3) in which unordered monotonicity* ***A-3*** *holds. Let $t \in \text{supp}(T)$ such that for any $t' \in \text{supp}(T) \setminus t$, and for any $i' \in \{1, \ldots, N_Z - 1\}$ there exists $i \in \{0, 1, \ldots, N_Z - 1\}$ such that $\Sigma_{t'}(i') \subset \Sigma_t(i)$, then $E(Y(t) - Y(\bar{t})|t$-Switchers)* *is identified by the following equation:*[5]

$$E(Y(t) - Y(\bar{t})|t\text{-Switchers}) = \sum_{i=1}^{N_Z-1}\left(\frac{\boldsymbol{b}_t(i)\boldsymbol{B}_t^+\boldsymbol{Q}_Z(t)}{\boldsymbol{b}_t(i)\boldsymbol{B}_t^+\boldsymbol{P}_Z(t)} - \frac{\boldsymbol{b}_t(i)\sum_{t'\in\bar{t}}\left(\boldsymbol{B}_{t'}^+\boldsymbol{Q}_Z(t')\right)\odot\left(\boldsymbol{B}_{t'}'\boldsymbol{Pr_Z}\right)}{\boldsymbol{b}_t(i)\sum_{t'\in\bar{t}}\left(\boldsymbol{B}_{t'}^+\boldsymbol{P}_Z(t')\right)\odot\left(\boldsymbol{B}_{t'}'\boldsymbol{Pr_Z}\right)}\right)\zeta_n,$$
$$\text{where } \zeta_n = \frac{\boldsymbol{b}_t(i)\boldsymbol{B}_t^+\boldsymbol{P}_Z(t)}{\left(\sum_{i=1}^{N_Z-1}\boldsymbol{b}_t(i)\right)\boldsymbol{B}_t^+\boldsymbol{P}_Z(t)}.$$

Proof. See Section A.20 in this Appendix. ☐

See Appendix A.21 for calculations of $E(Y(t_a) - Y(\bar{t}_a)|t_a\text{-Switchers})$ associated with the response matrix $\boldsymbol{R}$ of Table 3.

A.18 Proof of Theorem T-9

The expression in Theorem **T-9** can be disaggregated into two components given in Equations (A.62) and (A.63):

$$E(Y(t)|\boldsymbol{S}\in\Sigma_t(i)) = \frac{\boldsymbol{b}_t(i)\boldsymbol{B}_t^+\boldsymbol{Q}_Z(t)}{\boldsymbol{b}_t(i)\boldsymbol{B}_t^+\boldsymbol{P}_Z(t)}, \tag{A.62}$$

$$E(Y(\bar{t})|\boldsymbol{S}\in\Sigma_t(i)) = \frac{\boldsymbol{b}_t(i)\sum_{t'\in\bar{t}}\left(\boldsymbol{B}_{t'}^+\boldsymbol{Q}_Z(t')\right)\odot\left(\boldsymbol{B}_{t'}'\boldsymbol{Pr_Z}\right)}{\boldsymbol{b}_t(i)\sum_{t'\in\bar{t}}\left(\boldsymbol{B}_{t'}^+\boldsymbol{P}_Z(t')\right)\odot\left(\boldsymbol{B}_{t'}'\boldsymbol{Pr_Z}\right)}. \tag{A.63}$$

Theorem **T-6** identifies $E(Y(t)|\boldsymbol{S}\in\Sigma_t(i))$ of Equation (A.62). Thus it suffices to demonstrate that $E(Y(\bar{t})|\boldsymbol{S}\in\Sigma_t(i))$ can be expressed as described in Equation (A.63).

We now revisit Equations (A.59)–(A.61) that define $E(Y(\bar{t})|\boldsymbol{S}\in\Sigma_t(i))$ as:

$$E(Y(\bar{t})|\boldsymbol{S}\in\Sigma_t(i)) = \sum_{\boldsymbol{s}\in\Sigma_t(i)} E(Y(\bar{t})|\boldsymbol{S}=\boldsymbol{s})\frac{P(\boldsymbol{S}=\boldsymbol{s})}{P(\boldsymbol{S}\in\Sigma_t(i))} \tag{A.64}$$

$$\text{where } E(Y(\bar{t})|\boldsymbol{S}=\boldsymbol{s}) = \sum_{t'\in\bar{t}} E(Y(t')|\boldsymbol{S}=\boldsymbol{s})\frac{P(T=t'|\boldsymbol{S}=\boldsymbol{s})}{P(T\in\bar{t}|\boldsymbol{S}=\boldsymbol{s})} \text{ and } \bar{t}\equiv\text{supp}(T)\setminus\{t\}. \tag{A.65}$$

Equations (A.64)–(A.65) can be concatenated into the following equation:

$$E(Y(\bar{t})|\boldsymbol{S}\in\Sigma_t(i)) = \frac{\sum_{\boldsymbol{s}\in\Sigma_t(i)}\left(\sum_{t'\in\bar{t}} E(Y(t')|\boldsymbol{S}=\boldsymbol{s})\frac{P(T=t'|\boldsymbol{S}=\boldsymbol{s})}{P(T\in\bar{t}|\boldsymbol{S}=\boldsymbol{s})}\right)P(\boldsymbol{S}=\boldsymbol{s})}{P(\boldsymbol{S}\in\Sigma_t(i))}. \tag{A.66}$$

[5] Here we adopt $\kappa(Y) = Y$, therefore $\boldsymbol{Q}_Z(t)$ is given by: $\boldsymbol{Q}_Z(t) = [E(Y|T=t, Z=z_1),\ldots,E(Y|T=t, Z=z_{N_Z})]'\odot\boldsymbol{P}_Z(t)$.

Thus, to prove the theorem, it suffices to show that the following equation holds:

$$\sum_{\boldsymbol{s}\in\Sigma_t(i)} \frac{\sum_{t'\in\bar{t}} E(Y(t')|\boldsymbol{S}=\boldsymbol{s})P(T=t'|\boldsymbol{S}=\boldsymbol{s})P(\boldsymbol{S}=\boldsymbol{s})}{P(\boldsymbol{S}\in\Sigma_t(i))P(T\in\bar{t}|\boldsymbol{S}=\boldsymbol{s})} = \left[\frac{\boldsymbol{b}_t(i)\sum_{t'\in\bar{t}}\left(\boldsymbol{B}_{t'}^{+}\boldsymbol{Q}_Z(t')\right)\odot\left(\boldsymbol{B}_{t'}'\boldsymbol{Pr_Z}\right)}{\boldsymbol{b}_t(i)\sum_{t'\in\bar{t}}\left(\boldsymbol{B}_{t'}^{+}\boldsymbol{P}_Z(t')\right)\odot\left(\boldsymbol{B}_{t'}'\boldsymbol{Pr_Z}\right)}\right]. \tag{A.67}$$

The proof is divided in five steps. We first prove four lemmas (**L-25**–**L-28**) that will then be used to demonstrate Equation (A.67).

Lemma **L-25** below investigates probabilities $P(T=t|\boldsymbol{S}=\boldsymbol{s}); \boldsymbol{s}\in\Sigma_t(i)$:

Lemma L-25. Consider the IV model (1)–(3) in which unordered monotonicity **A-3** holds. Let $t\in\text{supp}(T)$ and $\boldsymbol{s},\boldsymbol{s}'\in\text{supp}(\boldsymbol{S})$ such that $\boldsymbol{s},\boldsymbol{s}'\in\Sigma_t(i)$ for some $i\in\{1,\ldots,N_Z-1\}$, then the following equalities for probabilities $P(T=t|\boldsymbol{S}=\boldsymbol{s}); \boldsymbol{s}\in\Sigma_t(i)$ hold:

$$P(T=t|\boldsymbol{S}=\boldsymbol{s}) = P(T=t|\boldsymbol{S}=\boldsymbol{s}') \text{ for any } \boldsymbol{s},\boldsymbol{s}'\in\Sigma_t(i) \tag{A.68}$$

$$P(T=t|\boldsymbol{S}=\boldsymbol{s}) = P(T=t|\boldsymbol{S}\in\Sigma_t(i)) \text{ for all } \boldsymbol{s}\in\Sigma_t(i) \tag{A.69}$$

$$P(T=t|\boldsymbol{S}\in\Sigma_t(i))P(\boldsymbol{S}\in\Sigma_t(i)) = \boldsymbol{b}_t(i)\Big(\boldsymbol{B}_t^{+}\boldsymbol{P}_Z(t)\Big)\odot\Big(\boldsymbol{B}_t'\boldsymbol{Pr_Z}\Big), \tag{A.70}$$

where $\odot$ means the Hadamard (element-wise) multiplication and $\boldsymbol{Pr_Z}$ and $\boldsymbol{b}_t(i)$ are previously defined notations: $\boldsymbol{Pr_Z} = [P(Z=z_1),\ldots,P(Z=z_{N_Z})]'$ is the vector of instrumental variable probabilities and $\boldsymbol{b}_t(i) = \big[\mathbf{1}[s_1\in\Sigma_t(i)],\ldots,\mathbf{1}[s_{N_{\boldsymbol{S}}}\in\Sigma_t(i)]\big]$ is the binary row-vector that indicates if each $\boldsymbol{s}\in\text{supp}(\boldsymbol{S})$ belongs to $\Sigma_t(i)$.

Proof. Probability $P(T=t|\boldsymbol{S}=\boldsymbol{s})$ can be written as:

$$P(T=t|\boldsymbol{S}=\boldsymbol{s}) = \sum_{z\in\text{supp}(Z)} P(T=t|\boldsymbol{S}=\boldsymbol{s},Z=z)P(Z=z|\boldsymbol{S}=\boldsymbol{s}),$$

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

$$\Rightarrow P(T=t|\boldsymbol{S}=\boldsymbol{s}) = \boldsymbol{B}_t[\cdot,s]'\boldsymbol{Pr_Z}. \tag{A.71}$$

The second equality simply restates Equation (17) of the main paper. It uses the fact that $Z \perp\!\!\!\perp S$ as shown in **L-1**. The third equality (Equation (A.71)) relies on the definition of $\boldsymbol{B}_t$. In our notation $\boldsymbol{B}_t[\cdot,\boldsymbol{s}]'$ means the transpose of the column-vector $\boldsymbol{B}_t[\cdot,\boldsymbol{s}]$ and $\boldsymbol{B}_t[\cdot,\boldsymbol{s}']$ means the column-vector of $\boldsymbol{B}_t$ associated with response type $\boldsymbol{s}'$. Lemma **L-16** states that if unordered monotonicity **A-3** holds and $\boldsymbol{s},\boldsymbol{s}'\in\Sigma_t(i)$, then $\boldsymbol{B}_t[\cdot,\boldsymbol{s}] = \boldsymbol{B}_t[\cdot,\boldsymbol{s}']$ (see also

Remark A.2). We combine Equation (A.71) ($P(T=t|\boldsymbol{S}=\boldsymbol{s}) = \boldsymbol{B}_t[\cdot,\boldsymbol{s}]'\boldsymbol{Pr_Z}$) and Lemma **L-16** to prove Equation (A.68) of the lemma:

$$P(T=t|\boldsymbol{S}=\boldsymbol{s}) = \boldsymbol{B}_t[\cdot,\boldsymbol{s}]'\boldsymbol{Pr_Z} = \boldsymbol{B}_t[\cdot,\boldsymbol{s}']'\boldsymbol{Pr_Z} = P(T=t|\boldsymbol{S}=\boldsymbol{s}').$$

Equation (A.68) implies Equation (A.69) as shown below:

$$\begin{aligned} P(T=t|\boldsymbol{S}\in\Sigma_t(i)) &= \sum_{\boldsymbol{s}'\in\Sigma_t(i)} P(T=t|\boldsymbol{S}=\boldsymbol{s}')\frac{P(\boldsymbol{S}=\boldsymbol{s}')}{P(\boldsymbol{S}\in\Sigma_t(i))} \\ &= P(T=t|\boldsymbol{S}=\boldsymbol{s})\sum_{\boldsymbol{s}'\in\Sigma_t(i)}\frac{P(\boldsymbol{S}=\boldsymbol{s}')}{P(\boldsymbol{S}\in\Sigma_t(i))} \text{ for any } \boldsymbol{s}\in\Sigma_t(i). \\ &= P(T=t|\boldsymbol{S}=\boldsymbol{s}). \end{aligned}$$

Equation (A.70) comes from expressing $P(T=t|\boldsymbol{S}\in\Sigma_t(i))$ in matrix notation and applying the results just stated. From Theorem **T-6** we have that $P(\boldsymbol{S}\in\Sigma_t(i))$ is identified by: $P(\boldsymbol{S}\in\Sigma_t(i)) = \boldsymbol{b}_t(i)\boldsymbol{B}_t^{+}\boldsymbol{P}_Z(t)$. Note that $\boldsymbol{b}_t(i)$ is a row-vector of dimension $1\times N_{\boldsymbol{S}}$ and $\boldsymbol{B}_t^{+}\boldsymbol{P}_Z(t)$ is a vector of dimension $N_{\boldsymbol{S}}\times 1$. Let ξ represent the vector $\xi = \boldsymbol{B}_t^{+}\boldsymbol{P}_Z(t)$. In this notation, we can write $P(T=t|\boldsymbol{S}\in\Sigma_t(i))\cdot P(\boldsymbol{S}\in\Sigma_t(i))$ as:

$$\begin{aligned} &P(\boldsymbol{S}\in\Sigma_t(i))\cdot P(T=t|\boldsymbol{S}\in\Sigma_t(i)) \\ &= \Big(\sum_{\boldsymbol{s}\in\text{supp}(\boldsymbol{S})} \boldsymbol{b}_t(i)[1,\boldsymbol{s}]\cdot\xi[\boldsymbol{s},1]\Big)\cdot P(T=t|\boldsymbol{S}\in\Sigma_t(i)) \\ &= \sum_{\boldsymbol{s}\in\text{supp}(\boldsymbol{S})} \boldsymbol{b}_t(i)[1,\boldsymbol{s}]\cdot\xi[\boldsymbol{s},1]\cdot P(T=t|\boldsymbol{S}\in\Sigma_t(i)) \\ &= \sum_{\boldsymbol{s}\in\Sigma_t(i)} \boldsymbol{b}_t(i)[1,\boldsymbol{s}]\cdot\xi[\boldsymbol{s},1]\cdot P(T=t|\boldsymbol{S}\in\Sigma_t(i)) + \sum_{\boldsymbol{s}\notin\Sigma_t(i)} \boldsymbol{b}_t(i)[1,\boldsymbol{s}]\cdot\xi[\boldsymbol{s},1]\cdot P(T=t|\boldsymbol{S}\in\Sigma_t(i)) \\ &= \sum_{\boldsymbol{s}\in\Sigma_t(i)} \boldsymbol{b}_t(i)[1,\boldsymbol{s}]\cdot\xi[\boldsymbol{s},1]\cdot P(T=t|\boldsymbol{S}=\boldsymbol{s}) + \sum_{\boldsymbol{s}\notin\Sigma_t(i)} \boldsymbol{b}_t(i)[1,\boldsymbol{s}]\cdot\xi[\boldsymbol{s},1]\cdot P(T=t|\boldsymbol{S}=\boldsymbol{s}) \\ &= \sum_{\boldsymbol{s}\in\text{supp}(\boldsymbol{S})} \boldsymbol{b}_t(i)[1,\boldsymbol{s}]\cdot\xi[\boldsymbol{s},1]\cdot P(T=t|\boldsymbol{S}=\boldsymbol{s}) \\ &= \sum_{\boldsymbol{s}\in\text{supp}(\boldsymbol{S})} \boldsymbol{b}_t(i)[1,\boldsymbol{s}]\cdot\xi[\boldsymbol{s},1]\cdot (\boldsymbol{B}_t[\cdot,\boldsymbol{s}]'\boldsymbol{Pr_Z}) \\ &= \boldsymbol{b}_t(i)\Big(\xi\odot(\boldsymbol{B}_t'\boldsymbol{Pr_Z})\Big) \\ &= \boldsymbol{b}_t(i)\Big(\boldsymbol{B}_t^{+}\boldsymbol{P}_Z(t)\Big)\odot\Big(\boldsymbol{B}_t'\boldsymbol{Pr_Z}\Big). \end{aligned}$$

The first equality simply states $P(\boldsymbol{S}\in\Sigma_t(i))$ as a summation. The second equality includes the value $P(T=t|\boldsymbol{S}\in\Sigma_t(i))$ inside the summation. The third equality splits the summation into two terms according to the following partition of the support of $\boldsymbol{S}:\Sigma_t(i)$ and $\text{supp}(\boldsymbol{S})\setminus\Sigma_t(i)$. The fourth equality replaces $P(T=t|\boldsymbol{S}\in\Sigma_t(i))$ by $P(T=t|\boldsymbol{S}=\boldsymbol{s})$ in each term of the sum. The reasons for the replacement differ in each term of the summation. The replacement in the first term of the fourth equality is due to Equation (A.69) which states that $P(T=t|\boldsymbol{S}=\boldsymbol{s}) = P(T=t|\boldsymbol{S}\in\Sigma_t(i))$ for all $\boldsymbol{s}\in\Sigma_t(i)$. The replacement in the second term of the fourth equality is due to the fact that $\boldsymbol{b}_t(i)[1,\boldsymbol{s}]=0$ for all $\boldsymbol{s}\notin\Sigma_t(i)$. The fifth equality regroups the summation terms. The sixth equality uses the Equation (A.71)

$(P(T=t|\boldsymbol{S}=\boldsymbol{s}) = \boldsymbol{B}_t[\cdot,\boldsymbol{s}]'\boldsymbol{Pr_Z}.)$ The seventh equality expresses the summation in matrix form. The eighth equality uses the definition of the vector $\xi = \boldsymbol{B}_t^{+}\boldsymbol{P}_Z(t)$. Note that if $\Sigma_t(i)=\emptyset$ then $\boldsymbol{b}_t(i)$ is a row-vector of elements zero and $P(T=t|\boldsymbol{S}\in\Sigma_t(i))=0$.

□

Our proof benefits from a convenient partition of the response-type set $\Sigma_t(i)$ explored in the lemma below.

Lemma L-26. Consider the IV model (1)–(3). Let $t = \text{supp}(T)\setminus\{t\}$ and $i\in\{0,1,\ldots,N_Z\}$, then, for any $t'\in\bar{t}\equiv\text{supp}(T)\setminus\{t\}$, we can always partition the response-type set $\Sigma_t(i)$ as:

$$\Sigma_t(i) = \cup_{i'=0}^{N_Z}\Big(\Sigma_t(i)\cap\Sigma_{t'}(i')\Big), \text{ and } \boldsymbol{b}_t(i) = \sum_{i'=0}^{N_Z}\boldsymbol{b}_{t'}(i')\odot\boldsymbol{b}_t(i). \tag{A.72}$$

$$\text{Moreover, if } \Sigma_t(i)\cap\Sigma_{t'}(i')\neq\emptyset \Rightarrow \Sigma_t(i)\cap\Sigma_{t'}(i') = \Sigma_{t'}(i') \text{ holds for some } t'\in\bar{t}, \tag{A.73}$$

$$\text{then } \mathbf{1}[\Sigma_t(i)\cap\Sigma_{t'}(i')\neq\emptyset]\cdot\boldsymbol{b}_{t'}(i') = \boldsymbol{b}_{t'}(i')\odot\boldsymbol{b}_t(i). \tag{A.74}$$

Proof. The partition in (A.72) uses the fact that $\cup_{i'=0}^{N_Z}\Sigma_{t'}(i')$ is a partition of supp$(\boldsymbol{S})$ for any $t'\in\text{supp}(T)$. The equation in (A.72) is a direct consequence of the definition of $\boldsymbol{b}_t(i)$ which is a binary row-vector that indicates is a response type belongs to set $\Sigma_t(i)$. Property (A.73) restates the theorem assumption. Suppose $\Sigma_t(i)\cap\Sigma_{t'}(i')\neq\emptyset$, then, according to the Assumption (A.73), $\Sigma_t(i)\cap\Sigma_{t'}(i') = \Sigma_{t'}(i')$. This implies that if $\boldsymbol{b}_{t'}(i')[1,\boldsymbol{s}]=1$, then it must be that $\boldsymbol{b}_t(i)[1,\boldsymbol{s}]=1$ for any $\boldsymbol{s}\in\text{supp}(\boldsymbol{S})$. Therefore $\boldsymbol{b}_{t'}(i')\odot\boldsymbol{b}_t(i) = 1\cdot\boldsymbol{b}_{t'}(i')$. Instead, if $\Sigma_t(i)\cap\Sigma_{t'}(i')=\emptyset$, then $\boldsymbol{b}_{t'}(i')\odot\boldsymbol{b}_t(i)$ is a row-vector of zero elements, which can be written as $\boldsymbol{b}_{t'}(i')\odot\boldsymbol{b}_t(i) = 0\cdot\boldsymbol{b}_{t'}(i')$. Thus the equation $\mathbf{1}[\Sigma_t(i)\cap\Sigma_{t'}(i')\neq\emptyset]\cdot\boldsymbol{b}_{t'}(i') = \boldsymbol{b}_{t'}(i')\odot\boldsymbol{b}_t(i)$ holds regardless if $\Sigma_t(i)\cap\Sigma_{t'}(i')$ is empty or not.

□

Lemma **L-27** below investigates probabilities $P(T\in\bar{t}|\boldsymbol{S}=\boldsymbol{s});\boldsymbol{s}\in\Sigma_t(i)$. It can be understood as a counterpart of Lemma **L-25**, which focuses on probabilities $P(T=t|\boldsymbol{S}=\boldsymbol{s});\boldsymbol{s}\in\Sigma_t(i)$.

Lemma L-27. Consider the IV model (1)–(3) in which unordered monotonicity **A-3** holds. Let $\bar{t}=\text{supp}(T)\setminus\{t\}$ and $\boldsymbol{s},\boldsymbol{s}'\in\text{supp}(\boldsymbol{S})$ such that $\boldsymbol{s},\boldsymbol{s}'\in\Sigma_t(i)$ for some $i\in\{1,\ldots,N_Z-1\}$, then the following equalities for probabilities $P(T=t|\boldsymbol{S}=\boldsymbol{s});\boldsymbol{s}\in\Sigma_t(i)$ hold:

$$P(T\in\bar{t}|\boldsymbol{S}=\boldsymbol{s}) = P(T\in\bar{t}|\boldsymbol{S}=\boldsymbol{s}') \text{ for any } \boldsymbol{s},\boldsymbol{s}'\in\Sigma_t(i) \tag{A.75}$$

$$P(T\in\bar{t}|\boldsymbol{S}=\boldsymbol{s}) = P(T\in\bar{t}|\boldsymbol{S}\in\Sigma_t(i)) \text{ for all } \boldsymbol{s}\in\Sigma_t(i). \tag{A.76}$$

Moreover, if $\Sigma_t(i) \cap \Sigma_{t'}(i') \neq \emptyset \Rightarrow \Sigma_t(i) \cap \Sigma_{t'}(i') = \Sigma_{t'}(i')$ holds for all $t' \in \bar{t}$, and $i' \in \{1, \dots, N_Z\}$,

$$\text{then } P(T \in \bar{t}|\boldsymbol{S} \in \Sigma_t(i))P(\boldsymbol{S} \in \Sigma_t(i)) = \boldsymbol{b}_t(i) \sum_{t' \in \bar{t}} \left(\boldsymbol{B}_{t'}^{+}\boldsymbol{P}_Z(t')\right) \odot \left(\boldsymbol{B}_{t'}'\boldsymbol{Pr_Z}\right). \tag{A.77}$$

where $\odot$, $\boldsymbol{Pr_Z}$ and $\boldsymbol{b}_t(i)$ follow our previous notation.

Proof. From Equation (A.75) comes from (A.68) of Lemma **L-25**:

$$P(T \in \bar{t}|\boldsymbol{S} = \boldsymbol{s}) = \Big(1 - P(T = t|\boldsymbol{S} = \boldsymbol{s})\Big) = \Big(1 - P(T = t|\boldsymbol{S} = \boldsymbol{s}')\Big) = P(T \in \bar{t}|\boldsymbol{S} = \boldsymbol{s}') \text{ for all } \boldsymbol{s}, \boldsymbol{s}' \in \Sigma_t(i).$$

From Equation (A.76) comes from (A.69) of Lemma **L-25**:

$$P(T \in \bar{t}|\boldsymbol{S} \in \Sigma_t(i)) = \Big(1 - P(T = t|\boldsymbol{S} \in \Sigma_t(i))\Big) = \Big(1 - P(T = t|\boldsymbol{S} = \boldsymbol{s})\Big) = P(T \in \bar{t}|\boldsymbol{S} = \boldsymbol{s}) \text{ for any } \boldsymbol{s} \in \Sigma_t(i).$$

We now rewrite probability $P(T \in \bar{t}|\boldsymbol{S} \in \Sigma_t(i))$ in (A.76) as:

$$\begin{aligned}
P(T \in \bar{t}|\boldsymbol{S} \in \Sigma_t(i)) &= \sum_{t' \in \bar{t}} P(T = t'|\boldsymbol{S} \in \Sigma_t(i)) \\
&= \sum_{t' \in \bar{t}} \sum_{i'=0}^{N_Z} \mathbf{1}[\Sigma_t(i) \cap \Sigma_{t'}(i') \neq \emptyset] P(T = t'|\boldsymbol{S} \in \Sigma_t(i) \cap \Sigma_{t'}(i')) \frac{P(\boldsymbol{S} \in \Sigma_t(i) \cap \Sigma_{t'}(i'))}{P(\boldsymbol{S} \in \Sigma_t(i))} \\
&= \sum_{t' \in \bar{t}} \sum_{i'=0}^{N_Z} \mathbf{1}[\Sigma_t(i) \cap \Sigma_{t'}(i') \neq \emptyset] P(T = t'|\boldsymbol{S} \in \Sigma_{t'}(i')) \frac{P(\boldsymbol{S} \in \Sigma_{t'}(i'))}{P(\boldsymbol{S} \in \Sigma_t(i))} \\
\Rightarrow P(T \in \bar{t}|\boldsymbol{S} \in \Sigma_t(i))P(\boldsymbol{S} \in \Sigma_t(i)) &= \sum_{t' \in \bar{t}} \sum_{i'=0}^{N_Z} \mathbf{1}[\Sigma_t(i) \cap \Sigma_{t'}(i') \neq \emptyset] P(T = t'|\boldsymbol{S} \in \Sigma_{t'}(i')) P(\boldsymbol{S} \in \Sigma_{t'}(i')) \\
&= \sum_{t' \in \bar{t}} \sum_{i'=0}^{N_Z} \mathbf{1}[\Sigma_t(i) \cap \Sigma_{t'}(i') \neq \emptyset] \boldsymbol{b}_{t'}(i') \Big(\boldsymbol{B}_t^{+}\boldsymbol{P}_Z(t)\Big) \odot \Big(\boldsymbol{B}_t'\boldsymbol{Pr_Z}\Big). \\
&= \sum_{t' \in \bar{t}} \left(\sum_{i'=0}^{N_Z} \mathbf{1}[\Sigma_t(i) \cap \Sigma_{t'}(i') \neq \emptyset] \boldsymbol{b}_{t'}(i')\right) (\boldsymbol{B}_{t'}^{+}\boldsymbol{P}_Z(t')) \odot (\boldsymbol{B}_t'\boldsymbol{Pr_Z}) \\
&= \sum_{t' \in \bar{t}} \left(\sum_{i'=0}^{N_Z} \boldsymbol{b}_{t'}(i') \odot \boldsymbol{b}_t(i)\right) (\boldsymbol{B}_{t'}^{+}\boldsymbol{P}_Z(t')) \odot (\boldsymbol{B}_t'\boldsymbol{Pr_Z}) \\
&= \sum_{t' \in \bar{t}} \boldsymbol{b}_t(i) (\boldsymbol{B}_{t'}^{+}\boldsymbol{P}_Z(t')) \odot (\boldsymbol{B}_t'\boldsymbol{Pr_Z}) \\
\therefore P(T \in \bar{t}|\boldsymbol{S} \in \Sigma_t(i))P(\boldsymbol{S} \in \Sigma_t(i)) &= \boldsymbol{b}_t(i) \sum_{t' \in \bar{t}} (\boldsymbol{B}_{t'}^{+}\boldsymbol{P}_Z(t')) \odot (\boldsymbol{B}_t'\boldsymbol{Pr_Z})
\end{aligned}$$

The second equality applies the law of iterated expectations over the partition suggested by Equation (A.70) of Lemma A.72. The equality introduces the binary indicator $\mathbf{1}[\Sigma_t(i) \cap \Sigma_{t'}(i') \neq \emptyset]$ that assures a summation within the response types in $\Sigma_t(i)$. Nevertheless, if a response-type set Σ is empty, that is $\Sigma = \emptyset$ then $P(\boldsymbol{S} \in \Sigma) = 0$. The third equality uses the assumption that $\Sigma_t(i) \cap \Sigma_{t'}(i') = \Sigma_{t'}(i')$ whenever $\Sigma_t(i) \cap \Sigma_{t'}(i') \neq \emptyset$. The fourth equality generates the term $P(T \in \bar{t}|\boldsymbol{S} \in \Sigma_t(i))P(\boldsymbol{S} \in \Sigma_t(i))$, which is our object of analysis as stated in the lemma. The fifth equality applies the result of Equation (A.70) of Lemma **L-25**, namely $P(T = t'|\boldsymbol{S} \in \Sigma_{t'}(i'))P(\boldsymbol{S} \in \Sigma_{t'}(i')) = \boldsymbol{b}_{t'}(i')(\boldsymbol{B}_t^{+}\boldsymbol{P}_Z(t)) \odot (\boldsymbol{B}_t'\boldsymbol{Pr_Z})$. The sixth equality isolates the common term $(\boldsymbol{B}_t^{+}\boldsymbol{P}_Z(t)) \odot (\boldsymbol{B}_t'\boldsymbol{Pr_Z})$ of the summation. The seventh equality uses the result stated by Equation (A.74) of Lemma **L-26**, that is, $\mathbf{1}[\Sigma_t(i) \cap \Sigma_{t'}(i') \neq$

$\emptyset] \cdot \boldsymbol{b}_{t'}(i') = \boldsymbol{b}_{t'}(i') \odot \boldsymbol{b}_t(i)$. The eighth equality uses the result stated by Equation (A.72) of Lemma **L-26**, that is, $\boldsymbol{b}_t(i) = \sum_{i'=0}^{N_Z} \boldsymbol{b}_{t'}(i') \odot \boldsymbol{b}_t(i)$. The eighth equality isolates the common term $\boldsymbol{b}_t(i)$.

□

Lemma **L-28** shows the identification formula for $E(Y(t) \cdot \mathbf{1}[T = t]|\boldsymbol{S} \in \Sigma_t(i))$ that will be useful in our final proof.

Lemma L-28. Consider the IV model (1)–(3) in which unordered monotonicity **A-3** holds. Let $t \in \text{supp}(T)$ and $i \in \{1, \ldots, N_Z\}$ then:

$$E(Y(t) \cdot \mathbf{1}[T = t]|\boldsymbol{S} \in \Sigma_t(i)) = E(Y(t)|\boldsymbol{S} \in \Sigma_t(i)) \cdot P(T = t|\boldsymbol{S} \in \Sigma_t(i)), \tag{A.78}$$

and $E(Y(t) \cdot \mathbf{1}[T = t]|\boldsymbol{S} \in \Sigma_t(i))P(\boldsymbol{S} \in \Sigma_t(i))$ is identified by:

$$E(Y(t) \cdot \mathbf{1}[T = t]|\boldsymbol{S} \in \Sigma_t(i))P(\boldsymbol{S} \in \Sigma_t(i)) = \boldsymbol{b}_t(i)\big(\boldsymbol{B}_{t'}^{+}\boldsymbol{Q}_Z(t')\big) \odot \big(\boldsymbol{B}_{t'}'\boldsymbol{Pr}_{\boldsymbol{Z}}\big). \tag{A.79}$$

Proof. Equation (A.79) is demonstrated below:

$$\begin{aligned}
E(Y(t) \cdot \mathbf{1}[T = t]|\boldsymbol{S} \in \Sigma_t(i)) &= \sum_{\boldsymbol{s} \in \Sigma_t(i)} E(Y(t) \cdot \mathbf{1}[T = t]|\boldsymbol{S} = \boldsymbol{s}) \frac{P(\boldsymbol{S} = \boldsymbol{s})}{P(\boldsymbol{S} \in \Sigma_t(i))} \\
&= \sum_{\boldsymbol{s} \in \Sigma_t(i)} E(Y(t)|\boldsymbol{S} = \boldsymbol{s}) \cdot E(\mathbf{1}[T = t]|\boldsymbol{S} = \boldsymbol{s}) \frac{P(\boldsymbol{S} = \boldsymbol{s})}{P(\boldsymbol{S} \in \Sigma_t(i))} \\
&= \frac{\sum_{\boldsymbol{s} \in \Sigma_t(i)} E(Y(t)|\boldsymbol{S} = \boldsymbol{s}) \cdot P(T = t|\boldsymbol{S} = \boldsymbol{s}) P(\boldsymbol{S} = \boldsymbol{s})}{P(\boldsymbol{S} \in \Sigma_t(i))} \\
\Rightarrow E(Y(t) \cdot \mathbf{1}[T = t]|\boldsymbol{S} \in \Sigma_t(i))P(\boldsymbol{S} \in \Sigma_t(i) &= \sum_{\boldsymbol{s} \in \Sigma_t(i)} \Big(E(Y(t)|\boldsymbol{S} = \boldsymbol{s})P(\boldsymbol{S} = \boldsymbol{s})\Big) P(T = t|\boldsymbol{S} = \boldsymbol{s}) \\
&= P(T = t|\boldsymbol{S} \in \Sigma_t(i)) \Big(\sum_{\boldsymbol{s} \in \Sigma_t(i)} E(Y(t)|\boldsymbol{S} = \boldsymbol{s})P(\boldsymbol{S} = \boldsymbol{s})\Big), \\
&= P(T = t|\boldsymbol{S} \in \Sigma_t(i)) \Big(E(Y(t)|\boldsymbol{S} \in \Sigma_t(i)) \cdot P(\boldsymbol{S} \in \Sigma_t(i))\Big), \\
\therefore E(Y(t) \cdot \mathbf{1}[T = t]|\boldsymbol{S} \in \Sigma_t(i)) &= P(T = t|\boldsymbol{S} \in \Sigma_t(i)) E(Y(t)|\boldsymbol{S} \in \Sigma_t(i)).
\end{aligned} \tag{A.80}$$

The fifth equality comes from Equation (A.69) of Lemma **L-25** which state that $P(T = t|\boldsymbol{S} = \boldsymbol{s}) = P(T = t|\boldsymbol{S} \in \Sigma_t(i))$ for all $\boldsymbol{s} \in \Sigma_t(i)$. The last equation eliminates the term $P(\boldsymbol{S} \in \Sigma_t(i))$ in both sides of the equation.

According to Theorem **T-6**, $E(Y(t)|\boldsymbol{S} \in \Sigma_t(i)) \cdot P(\boldsymbol{S} \in \Sigma_t(i))$ is identified by $E(Y(t)|\boldsymbol{S} \in \Sigma_t(i)) \cdot P(\boldsymbol{S} \in \Sigma_t(i)) = \boldsymbol{b}_t(i)\boldsymbol{B}_t^{+}\boldsymbol{Q}_Z(t)$. Note that $\boldsymbol{b}_t(i)$ is a row-vector of dimension $1 \times N_{\boldsymbol{S}}$ and $\boldsymbol{B}_t^{+}\boldsymbol{Q}_Z(t)$ is a vector of dimension $N_{\boldsymbol{S}} \times 1$. Let ξ represent the vector $\xi = \boldsymbol{B}_t^{+}\boldsymbol{Q}_Z(t)$. The remaining of the proof of this lemma follows the rationale of Lemma **L-25**:

$$
\begin{aligned}
&E(Y(t)\cdot \mathbf{1}[T=t]|\boldsymbol{S}\in\Sigma_t(i))P(\boldsymbol{S}\in\Sigma_t(i))\\
&\quad= P(T=t|\boldsymbol{S}\in\Sigma_t(i))\Big(E(Y(t)|\boldsymbol{S}\in\Sigma_t(i))P(\boldsymbol{S}\in\Sigma_t(i)\Big)\\
&\quad= P(T=t|\boldsymbol{S}\in\Sigma_t(i))\Big(\boldsymbol{b}_t(i)\boldsymbol{B}_t^{+}\boldsymbol{Q}_Z(t)\Big)\\
&\quad= P(T=t|\boldsymbol{S}\in\Sigma_t(i))\Big(\sum_{\boldsymbol{s}\in \mathrm{supp}(\boldsymbol{S})}\boldsymbol{b}_t(i)[1,\boldsymbol{s}]\cdot\xi[\boldsymbol{s},1]\Big)\\
&\quad= \sum_{\boldsymbol{s}\in \mathrm{supp}(\boldsymbol{S})}\boldsymbol{b}_t(i)[1,\boldsymbol{s}]\cdot\xi[\boldsymbol{s},1]\cdot P(T=t|\boldsymbol{S}\in\Sigma_t(i))\\
&\quad= \sum_{\boldsymbol{s}\in\Sigma_t(i)}\boldsymbol{b}_t(i)[1,\boldsymbol{s}]\cdot\xi[\boldsymbol{s},1]\cdot P(T=t|\boldsymbol{S}\in\Sigma_t(i))+\sum_{\boldsymbol{s}\notin\Sigma_t(i)}\boldsymbol{b}_t(i)[1,\boldsymbol{s}]\cdot\xi[\boldsymbol{s},1]\cdot P(T=t|\boldsymbol{S}\in\Sigma_t(i))\\
&\quad= \sum_{\boldsymbol{s}\in\Sigma_t(i)}\boldsymbol{b}_t(i)[1,\boldsymbol{s}]\cdot\xi[\boldsymbol{s},1]\cdot P(T=t|\boldsymbol{S}=\boldsymbol{s})+\sum_{\boldsymbol{s}\notin\Sigma_t(i)}\boldsymbol{b}_t(i)[1,\boldsymbol{s}]\cdot\xi[\boldsymbol{s},1]\cdot P(T=t|\boldsymbol{S}=\boldsymbol{s})\\
&\quad= \sum_{\boldsymbol{s}\in \mathrm{supp}(\boldsymbol{S})}\boldsymbol{b}_t(i)[1,\boldsymbol{s}]\cdot\xi[\boldsymbol{s},1]\cdot P(T=t|\boldsymbol{S}=\boldsymbol{s})\\
&\quad= \sum_{\boldsymbol{s}\in \mathrm{supp}(\boldsymbol{S})}\boldsymbol{b}_t(i)[1,\boldsymbol{s}]\cdot\xi[\boldsymbol{s},1]\cdot\big(\boldsymbol{B}_t[\cdot,\boldsymbol{s}]'\boldsymbol{Pr_Z}\big)\\
&\quad= \boldsymbol{b}_t(i)\Big(\xi\odot\big(\boldsymbol{B}_t'\boldsymbol{Pr_Z}\big)\Big)\\
&\quad= \boldsymbol{b}_t(i)\Big(\boldsymbol{B}_t^{+}\boldsymbol{Q}_Z(t)\Big)\odot\Big(\boldsymbol{B}_t'\boldsymbol{Pr_Z}\Big).
\end{aligned}
$$

The first equality applies the result in Equation (A.80). The second equality uses $E(Y(t)|\boldsymbol{S}\in\Sigma_t(i))\cdot P(\boldsymbol{S}\in\Sigma_t(i))=\boldsymbol{b}_t(i)\boldsymbol{B}_t^{+}\boldsymbol{Q}_Z(t)$ from Theorem **T-6**. The third equality transforms $\boldsymbol{b}_t(i)\boldsymbol{B}_t^{+}\boldsymbol{Q}_Z(t)$ into a summation where $\xi=\boldsymbol{B}_t^{+}\boldsymbol{Q}_Z(t)$. The fourth equality includes the value $P(T=t|\boldsymbol{S}\in\Sigma_t(i))$ inside the summation. The fifth equality splits the summation into two terms according to the following partition of the support of $\boldsymbol{S}:\Sigma_t(i)$ and $\mathrm{supp}(\boldsymbol{S})\setminus\Sigma_t(i)$. The sixth equality replaces $P(T=t|\boldsymbol{S}\in\Sigma_t(i))$ by $P(T=t|\boldsymbol{S}=\boldsymbol{s})$ in each term of the sum. The reasons for the replacement differ in each term of the summation. The replacement in the first term of the fourth equality is due to Equation (A.69) which states that $P(T=t|\boldsymbol{S}=\boldsymbol{s})=P(T=t|\boldsymbol{S}\in\Sigma_t(i))$ for all $\boldsymbol{s}\in\Sigma_t(i)$. The replacement in the second term of the fourth equality is due to the fact that $\boldsymbol{b}_t(i)[1,\boldsymbol{s}]=0$ for all $\boldsymbol{s}\notin\Sigma_t(i)$. The seventh equality regroups the summation terms. The sixth equality uses the Equation (A.71) ($P(T=t|\boldsymbol{S}=\boldsymbol{s})=\boldsymbol{B}_t[\cdot,\boldsymbol{s}]'\boldsymbol{Pr_Z}$.) The eighth equality expresses the summation into matrix form. The ninth equality uses the definition of the vector $\xi=\boldsymbol{B}_t^{+}\boldsymbol{Q}_Z(t)$. Note that if $\Sigma_t(i)=\emptyset$ then $\boldsymbol{b}_t(i)$ is a row-vector of elements zero and $P(T=t|\boldsymbol{S}\in\Sigma_t(i))=0$.

□

Proof. We now return to Equation (A.67):

$$\sum_{\boldsymbol{s}\in\Sigma_t(i)} \frac{\sum_{t'\in\bar{t}} E(Y(t')|\boldsymbol{S}=\boldsymbol{s})P(T=t'|\boldsymbol{S}=\boldsymbol{s})P(\boldsymbol{S}=\boldsymbol{s})}{P(\boldsymbol{S}\in\Sigma_t(i))P(T\in\bar{t}|\boldsymbol{S}=\boldsymbol{s})}$$
$$= \frac{\sum_{t'\in\bar{t}}\sum_{\boldsymbol{s}\in\Sigma_t(i)}\Big(E(Y(t')|\boldsymbol{S}=\boldsymbol{s})P(T=t'|\boldsymbol{S}=\boldsymbol{s})\Big)P(\boldsymbol{S}=\boldsymbol{s})}{P(T\in\bar{t}|\boldsymbol{S}\in\Sigma_t(i))P(\boldsymbol{S}\in\Sigma_t(i))}$$
$$= \frac{\sum_{t'\in\bar{t}}\sum_{\boldsymbol{s}\in\Sigma_t(i)}\Big(E(Y(t')\cdot\mathbf{1}[T=t']|\boldsymbol{S}=\boldsymbol{s})\Big)P(\boldsymbol{S}=\boldsymbol{s})}{P(T\in\bar{t}|\boldsymbol{S}\in\Sigma_t(i))P(\boldsymbol{S}\in\Sigma_t(i))}$$
$$= \frac{\sum_{t'\in\bar{t}} E(Y(t')\cdot\mathbf{1}[T=t']|\boldsymbol{S}\in\Sigma_t(i))P(\boldsymbol{S}\in\Sigma_t(i))}{P(T\in\bar{t}|\boldsymbol{S}\in\Sigma_t(i))P(\boldsymbol{S}\in\Sigma_t(i))}. \tag{A.81}$$

The first Equality uses Equation (A.76) of Lemma **L-27** which states that $P(T\in\bar{t}|\boldsymbol{S}=\boldsymbol{s}) = P(T\in\bar{t}|\boldsymbol{S}\in\Sigma_t(i))$ for all $\boldsymbol{s}\in\Sigma_t(i)$. The second equality uses the conditional independence $Y(t) \perp\!\!\!\perp T|\boldsymbol{S}$ from Lemma **L-1**. The denominator of (A.81) is given by Equation (A.77) of Lemma **L-27** which states that:

$$P(T\in\bar{t}|\boldsymbol{S}\in\Sigma_t(i))P(\boldsymbol{S}\in\Sigma_t(i)) = \boldsymbol{b}_t(i)\sum_{t'\in\bar{t}}\big(\boldsymbol{B}_{t'}^{+}\boldsymbol{P}_Z(t')\big)\odot\big(\boldsymbol{B}_{t'}'\boldsymbol{Pr_Z}\big).$$

This is the expected expression. To prove the theorem, it remains to derive the expression for the numerator of (A.81):

$$\left(\sum_{t'\in\bar{t}} E(Y(t')\cdot\mathbf{1}[T=t']|\boldsymbol{S}\in\Sigma_t(i))\right)P(\boldsymbol{S}\in\Sigma_t(i))$$
$$= \left(\sum_{t'\in\bar{t}}\sum_{i'=0}^{N_Z}\mathbf{1}[\Sigma_t(i)\cap\Sigma_{t'}(i')\neq\emptyset]E(Y(t')\cdot\mathbf{1}[T=t']|\boldsymbol{S}\in\Sigma_t(i)\cap\Sigma_{t'}(i'))\frac{P(\boldsymbol{S}\in\Sigma_t(i)\cap\Sigma_{t'}(i'))}{P(\boldsymbol{S}\in\Sigma_t(i))}\right)P(\boldsymbol{S}\in\Sigma_t(i))$$
$$= \sum_{t'\in\bar{t}}\sum_{i'=0}^{N_Z}\mathbf{1}[\Sigma_t(i)\cap\Sigma_{t'}(i')\neq\emptyset]E(Y(t')\cdot\mathbf{1}[T=t']|\boldsymbol{S}\in\Sigma_{t'}(i'))P(\boldsymbol{S}\in\Sigma_{t'}(i'))$$
$$= \sum_{t'\in\bar{t}}\left(\sum_{i'=0}^{N_Z}\mathbf{1}[\Sigma_t(i)\cap\Sigma_{t'}(i')\neq\emptyset]\boldsymbol{b}_{t'}(i')\right)\big(\boldsymbol{B}_t^{+}\boldsymbol{Q}_Z(t)\big)\odot\big(\boldsymbol{B}_t'\boldsymbol{Pr_Z}\big).$$
$$= \sum_{t'\in\bar{t}}\left(\sum_{i'=0}^{N_Z}\boldsymbol{b}_{t'}(i')\odot\boldsymbol{b}_t(i)\right)(\boldsymbol{B}_{t'}^{+}\boldsymbol{P}_Z(t'))\odot(\boldsymbol{B}_t'\boldsymbol{Pr_Z})$$
$$= \sum_{t'\in\bar{t}}\boldsymbol{b}_t(i)(\boldsymbol{B}_{t'}^{+}\boldsymbol{Q}_Z(t'))\odot(\boldsymbol{B}_t'\boldsymbol{Pr_Z})$$
$$\therefore \sum_{t'\in\bar{t}} E(Y(t')\cdot\mathbf{1}[T=t']|\boldsymbol{S}\in\Sigma_t(i))P(\boldsymbol{S}\in\Sigma_t(i)) = \boldsymbol{b}_t(i)\sum_{t'\in\bar{t}}(\boldsymbol{B}_{t'}^{+}\boldsymbol{Q}_Z(t'))\odot(\boldsymbol{B}_t'\boldsymbol{Pr_Z}).$$

The proof follows the rationale of Lemma **L-27**. The first equality applies Partition (A.72) of **L-26**. The second equality eliminates $P(\boldsymbol{S}\in\Sigma_t(i))$ and uses Property (A.73) of **L-26**. The third equality applies Equation A.79 of **L-28** which states that $E(Y(t)\cdot\mathbf{1}[T=t]|\boldsymbol{S}\in\Sigma_t(i))P(\boldsymbol{S}\in\Sigma_t(i)) = \boldsymbol{b}_t(i)\big(\boldsymbol{B}_{t'}^{+}\boldsymbol{Q}_Z(t')\big)\odot\big(\boldsymbol{B}_{t'}'\boldsymbol{Pr_Z}\big)$. The fourth equality applies Equation (A.74) of Lemma **L-26** while the fifth equality applies Equation (A.72) of **L-26**. The last equation isolates the common term $\boldsymbol{b}_t(i)$. ☐

A.19 Derivation of the Equations for the Example of Theorem T-9

Theorem T-9 generates the following formulas for $E(Y(t_a)|\boldsymbol{S} \in \Sigma_{t_a}(2))$ and $E(Y(\bar{t}_a)|\boldsymbol{S} \in \Sigma_{t_a}(2))$ for the response matrix $\boldsymbol{R}$ of Table 3:

$$E(Y(t_a)|\boldsymbol{S} \in \Sigma_{t_a}(2)) = E(Y(t_a)|\boldsymbol{S} \in \{s_2, s_3\}) = \frac{\boldsymbol{b}_{t_a}(2)\boldsymbol{B}_{t_a}^{+}\boldsymbol{Q}_Z(t_a)}{\boldsymbol{b}_{t_a}(2)\boldsymbol{B}_{t}^{+}\boldsymbol{P}_Z(t_a)} \tag{A.82}$$

$$E(Y(\bar{t}_a)|\boldsymbol{S} \in \Sigma_{t_a}(2)) = E(Y(\{t_b, t_c\})|\boldsymbol{S} \in \{s_2, s_3\}) =$$

$$= \frac{\boldsymbol{b}_{t_a}(2)\big(\boldsymbol{B}_{t_b}^{+}\boldsymbol{Q}_Z(t_b)\big) \odot \big(\boldsymbol{B}_{t_b}'\boldsymbol{Pr_Z}\big) + \boldsymbol{b}_{t_a}(2)\big(\boldsymbol{B}_{t_c}^{+}\boldsymbol{Q}_Z(t_c)\big) \odot \big(\boldsymbol{B}_{t_c}'\boldsymbol{Pr_Z}\big)}{\boldsymbol{b}_{t_a}(2)\big(\boldsymbol{B}_{t_b}^{+}\boldsymbol{P}_Z(t_b)\big) \odot \big(\boldsymbol{B}_{t_b}'\boldsymbol{Pr_Z}\big) + \boldsymbol{b}_{t_a}(2)\big(\boldsymbol{B}_{t_c}^{+}\boldsymbol{P}_Z(t_c)\big) \odot \big(\boldsymbol{B}_{t_c}'\boldsymbol{Pr_Z}\big)}. \tag{A.83}$$

$$\text{where } \boldsymbol{Pr_Z} = [P(Z = z_{no}), P(Z = z_a), P(Z = z_{bc})]'.$$

The components of Equation (A.82) 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 (A.82) that depend on the response matrix are:

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

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

Equation (A.82) renders the following expressions:

$$\boldsymbol{b}_{t_a}(2)\boldsymbol{B}_{t_a}^{+}\boldsymbol{Q}_Z(t_a) = E(Y\cdot\mathbf{1}[T=t_a]|Z=z_{no}) - E(Y\cdot\mathbf{1}[T=t_a]|Z=z_{bc}),$$

$$\boldsymbol{b}_{t_a}(2)\boldsymbol{B}_{t}^{+}\boldsymbol{P}_Z(t_a) = P(T=t_a|Z=z_{no}) - P(T=t_a|Z=z_{bc}), \tag{A.84}$$

$$\therefore E(Y(t_a)|t_a\in\Sigma_{t_a}(2)) = \frac{E(Y\cdot\mathbf{1}[T=t_a]|Z=z_{no}) - E(Y\cdot\mathbf{1}[T=t_a]|Z=z_{bc})}{P(T=t_a|Z=z_{no}) - P(T=t_a|Z=z_{bc})}. \tag{A.85}$$

We now examine Equation (A.83). We first target the terms $\boldsymbol{b}_{t_a}(2)\big(\boldsymbol{B}_{t_b}^{+}\boldsymbol{Q}_Z(t_b)\big)\odot\big(\boldsymbol{B}_{t_b}'\boldsymbol{Pr_Z}\big)$ and $\boldsymbol{b}_{t_a}(2)\big(\boldsymbol{B}_{t_b}^{+}\boldsymbol{P}_Z(t_b)\big)\odot\big(\boldsymbol{B}_{t_b}'\boldsymbol{Pr_Z}\big)$. The components of these terms that can be estimated from observed data are:

$$\boldsymbol{P}_Z(t_b) = [P(T=t_b|Z=z_{no}), P(T=t_b|Z=z_a), P(T=t_b|Z=z_{bc})]';$$

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

$$\boldsymbol{Pr_Z} = [P(Z=z_{no}), P(Z=z_a), P(Z=z_{bc})].$$

The components of these terms that depend on the response matrix are:

$$\boldsymbol{B}_{t_b} = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 \end{bmatrix} \Rightarrow \boldsymbol{B}_{t_b}^{+} = \begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \boldsymbol{B}_{t_b}'\boldsymbol{Pr_Z} = \begin{bmatrix} 0 \\ P(Z=z_{bc}) \\ 0 \\ P(Z=z_{no})+P(Z=z_{bc}) \\ 1 \\ 0 \\ 0 \end{bmatrix}.$$

The examined terms render the following expressions:

$$\boldsymbol{b}_{t_a}(2)\big(\boldsymbol{B}_{t_b}^{+}\boldsymbol{Q}_Z(t_b)\big)\odot\big(\boldsymbol{B}_{t_b}'\boldsymbol{Pr_Z}\big) = \big(E(Y\cdot\mathbf{1}[T=t_b]|Z=z_{bc}) - E(Y\cdot\mathbf{1}[T=t_b]|Z=z_{no})\big)P(Z=z_{bc}),$$

$$\boldsymbol{b}_{t_a}(2)\big(\boldsymbol{B}_{t_b}^{+}\boldsymbol{P}_Z(t_b)\big)\odot\big(\boldsymbol{B}_{t_b}'\boldsymbol{Pr_Z}\big) = \big(P(T=t_b|Z=z_{bc}) - P(T=t_b|Z=z_{no})\big)P(Z=z_{bc}).$$

Next, we target the terms $\boldsymbol{b}_{t_a}(2)\big(\boldsymbol{B}^{+}_{t_b}\boldsymbol{Q}_Z(t_b)\big)\odot\big(\boldsymbol{B}'_{t_b}\boldsymbol{Pr_Z}\big)$ and $\boldsymbol{b}_{t_a}(2)\big(\boldsymbol{B}^{+}_{t_b}\boldsymbol{P}_Z(t_b)\big)\odot\big(\boldsymbol{B}'_{t_b}\boldsymbol{Pr_Z}\big)$. The components of the terms that can be estimated from observed data are:

$$\boldsymbol{P}_Z(t_b) = [P(T = t_b|Z = z_{no}), P(T = t_b|Z = z_a), P(T = t_b|Z = z_{bc})]';$$

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

$$\boldsymbol{Pr_Z} = [P(Z = z_{no}), P(Z = z_a), P(Z = z_{bc})].$$

The components of these terms that depend on the response matrix are:

$$\boldsymbol{B}_{t_c} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 \end{bmatrix} \Rightarrow \boldsymbol{B}^{+}_{t_c} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \ \boldsymbol{B}'_{t_c}\boldsymbol{Pr_Z} = \begin{bmatrix} 0 \\ 0 \\ P(Z = z_{bc}) \\ 0 \\ 0 \\ P(Z = z_{no}) \\ 1 \end{bmatrix}.$$

The examined terms generate the following expressions:

$$\boldsymbol{b}_{t_a}(2)\big(\boldsymbol{B}^{+}_{t_c}\boldsymbol{Q}_Z(t_c)\big) \odot \big(\boldsymbol{B}'_{t_c}\boldsymbol{Pr_Z}\big) = \big(E(Y \cdot \mathbf{1}[T = t_c]|Z = z_{bc}) - E(Y \cdot \mathbf{1}[T = t_c]|Z = z_{no})\big)P(Z = z_{bc}),$$
$$\boldsymbol{b}_{t_a}(2)\big(\boldsymbol{B}^{+}_{t_c}\boldsymbol{P}_Z(t_c)\big) \odot \big(\boldsymbol{B}'_{t_c}\boldsymbol{Pr_Z}\big) = \big(P(T = t_c|Z = z_{bc}) - P(T = t_c|Z = z_{no})\big)P(Z = z_{bc}).$$

Combining the terms we have that:

$$E(Y(\bar{t}_a)|\boldsymbol{S} \in \Sigma_{t_a}(2)) = \frac{\big(E(Y \cdot \mathbf{1}[T = t_c]|Z = z_{bc}) - E(Y \cdot \mathbf{1}[T = t_c]|Z = z_{no})\big) + \big(E(Y \cdot \mathbf{1}[T = t_b]|Z = z_{bc}) - E(Y \cdot \mathbf{1}[T = t_b]|Z = z_{no})\big)}{\big(P(T = t_c|Z = z_{bc}) - P(T = t_c|Z = z_{no})\big) + \big(P(T = t_b|Z = z_{bc}) - P(T = t_b|Z = z_{no})\big)},$$

which can be also written as:

$$E(Y(\bar{t}_a)|\boldsymbol{S} \in \Sigma_{t_a}(2)) = \frac{\big(E(Y \cdot \mathbf{1}[T \neq t_a]|Z = z_{bc}) - E(Y \cdot \mathbf{1}[T \neq t_a]|Z = z_{no})\big)}{\big(P(T \neq t_a|Z = z_{bc}) - P(T \neq t_a|Z = z_{no})\big)},$$

A.20 Proof of Corollary C-5

Proof. Corollary **C-3** states that:

$$E(Y(t)|t\text{-Switchers}) = \frac{\big(\sum_{i=1}^{N_Z-1} \boldsymbol{b}_t(i)\big)\boldsymbol{B}_t^{+}\boldsymbol{Q}_Z(t)}{\big(\sum_{i=1}^{N_Z-1} \boldsymbol{b}_t(i)\big)\boldsymbol{B}_{t_a}^{+}\boldsymbol{P}_Z(t)}.$$

Thus it suffices to prove that

$$E(Y(\bar{t})|t\text{-Switchers}) = \frac{\boldsymbol{b}_t(i)\sum_{t'\in\bar{t}}\big(\boldsymbol{B}_{t'}^{+}\boldsymbol{Q}_Z(t')\big)\odot\big(\boldsymbol{B}_{t'}'\boldsymbol{Pr_Z}\big)}{\boldsymbol{b}_t(i)\sum_{t'\in\bar{t}}\big(\boldsymbol{B}_{t'}^{+}\boldsymbol{P}_Z(t')\big)\odot\big(\boldsymbol{B}_{t'}'\boldsymbol{Pr_Z}\big)}.$$

We use Theorem **T-9** to express $E(Y(\bar{t})|t\text{-Switchers})$ as:

$$\begin{aligned}
E(Y(\bar{t})|t\text{-Switchers}) &= E(Y(\bar{t})|\boldsymbol{S}\in\cup_{i=1}^{N_Z-1}\Sigma_t(i)) \\
&= \sum_{i=1}^{N_Z-1} E(Y(\bar{t})|\boldsymbol{S}\in\Sigma_t(i))\frac{P(\boldsymbol{S}\in\Sigma_t(i))}{P(\boldsymbol{S}\in\cup_{i=1}^{N_Z-1}\Sigma_t(i))} \\
&= \sum_{i=1}^{N_Z-1} \frac{\boldsymbol{b}_t(i)\sum_{t'\in\bar{t}}\big(\boldsymbol{B}_{t'}^{+}\boldsymbol{Q}_Z(t')\big)\odot\big(\boldsymbol{B}_{t'}'\boldsymbol{Pr_Z}\big)}{\boldsymbol{b}_t(i)\sum_{t'\in\bar{t}}\big(\boldsymbol{B}_{t'}^{+}\boldsymbol{P}_Z(t')\big)\odot\big(\boldsymbol{B}_{t'}'\boldsymbol{Pr_Z}\big)}\cdot\frac{\boldsymbol{b}_t(i)\boldsymbol{B}_t^{+}\boldsymbol{P}_Z(t)}{\Big(\sum_{i=1}^{N_Z-1}\boldsymbol{b}_t(i)\Big)\boldsymbol{B}_t^{+}\boldsymbol{P}_Z(t)}.
\end{aligned}$$

The third equality uses the equation for $P(\boldsymbol{S}\in t\text{-Switchers})$ of Corollary **C-2** and the expression for $P(\boldsymbol{S}\in\Sigma_t(i))$ of Theorem **T-6**. □

A.21 Derivation of Equations for Example of Corollary C-5

Corollary (**C-5**) states that $E(Y(t)|t\text{-Switchers})$ and $E(Y(\bar{t})|t\text{-Switchers})$ can be identified by:

$$E(Y(t)|t\text{-Switchers}) = \frac{\Big(\sum_{i=1}^{N_Z-1}\boldsymbol{b}_t(i)\Big)\boldsymbol{B}_t^{+}\boldsymbol{Q}_Z(t)}{\Big(\sum_{i=1}^{N_Z-1}\boldsymbol{b}_t(i)\Big)\boldsymbol{B}_t^{+}\boldsymbol{P}_Z(t)}, \tag{A.86}$$

$$E(Y(\bar{t})|t\text{-Switchers}) = \sum_{i=1}^{N_Z-1}\left(\frac{\boldsymbol{b}_t(i)\sum_{t'\in\bar{t}}\big(\boldsymbol{B}_{t'}^{+}\boldsymbol{Q}_Z(t')\big)\odot\big(\boldsymbol{B}_{t'}'\boldsymbol{Pr_Z}\big)}{\boldsymbol{b}_t(i)\sum_{t'\in\bar{t}}\big(\boldsymbol{B}_{t'}^{+}\boldsymbol{P}_Z(t')\big)\odot\big(\boldsymbol{B}_{t'}'\boldsymbol{Pr_Z}\big)}\right). \tag{A.87}$$

The parameter $E(Y(t_a)|t_a\text{-Switchers})$ is associated with the response matrix $\boldsymbol{R}$ of Table 3 is computed in Example 7.1. Thus we focus on Equation (A.87) for t_a-Switchers. Namely:

$$E(Y(\bar{t}_a)|t_a\text{-Switchers}) = \frac{\sum_{i=1}^{2}\left(\frac{\boldsymbol{b}_{t_a}(i)\left(\boldsymbol{B}_{t_b}^{+}\boldsymbol{Q}_Z(t_b)\right)\odot\left(\boldsymbol{B}_{t_b}'\boldsymbol{Pr_Z}\right)+\left(\boldsymbol{B}_{t_c}^{+}\boldsymbol{Q}_Z(t_c)\right)\odot\left(\boldsymbol{B}_{t_c}'\boldsymbol{Pr_Z}\right)}{\boldsymbol{b}_{t_a}(i)\left(\boldsymbol{B}_{t_b}^{+}\boldsymbol{P}_Z(t_b)\right)\odot\left(\boldsymbol{B}_{t_b}'\boldsymbol{Pr_Z}\right)+\left(\boldsymbol{B}_{t_c}^{+}\boldsymbol{P}_Z(t_c)\right)\odot\left(\boldsymbol{B}_{t_c}'\boldsymbol{Pr_Z}\right)}\right)}{\left(\sum_{i=1}^{2}\boldsymbol{b}_{t_a}(i)\right)\boldsymbol{B}_t^{+}\boldsymbol{P}_Z(t_a)}. \quad (A.88)$$

The denominator of Equation (A.88) is examined in Example 7.1 and is given by:

$$P(\boldsymbol{S}\in t_a\text{-Switchers}) = \left(\sum_{i=1}^{2}\boldsymbol{b}_{t_a}(i)\right)\boldsymbol{B}_t^{+}\boldsymbol{P}_Z(t_a) = P(T=t_a|Z=z_a) - P(T=t_a|Z=z_{bc}). \quad (A.89)$$

$$(\boldsymbol{B}_{t_b}^{+}\boldsymbol{Q}_Z(t_b))\odot(\boldsymbol{B}_{t_b}'\boldsymbol{Pr_Z}) = \begin{pmatrix} 0 \\ E(Y\cdot\mathbf{1}[T=t_b]|Z=z_{bc}) - E(Y\cdot\mathbf{1}[T=t_b]|Z=z_{no}) \\ 0 \\ E(Y\cdot\mathbf{1}[T=t_b]|Z=z_{no}) - E(Y\cdot\mathbf{1}[T=t_b]|Z=z_a) \\ E(Y\cdot\mathbf{1}[T=t_b]|Z=z_a) \\ 0 \\ 0 \end{pmatrix} \odot \begin{pmatrix} 0 \\ P(Z=z_{bc}) \\ 0 \\ P(Z=z_{no})+P(Z=z_{bc}) \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

$$(\boldsymbol{B}_{t_c}^{+}\boldsymbol{Q}_Z(t_c))\odot(\boldsymbol{B}_{t_c}'\boldsymbol{Pr_Z}) = \begin{pmatrix} 0 \\ 0 \\ E(Y\cdot\mathbf{1}[T=t_c]|Z=z_{bc}) - E(Y\cdot\mathbf{1}[T=t_c]|Z=z_{no}) \\ 0 \\ 0 \\ E(Y\cdot\mathbf{1}[T=t_c]|Z=z_{no}) - E(Y\cdot\mathbf{1}[T=t_b]|Z=z_b) \\ E(Y\cdot\mathbf{1}[T=t_c]|Z=z_{no}) \end{pmatrix} \odot \begin{pmatrix} 0 \\ 0 \\ P(Z=z_{bc}) \\ 0 \\ 0 \\ P(Z=z_{no}) \\ 1 \end{pmatrix}.$$

$$(\boldsymbol{B}_{t_b}^{+}\boldsymbol{P}_Z(t_b))\odot(\boldsymbol{B}_{t_b}'\boldsymbol{Pr_Z}) = \begin{pmatrix} 0 \\ P(T=t_b|Z=z_{bc}) - P(T=t_b|Z=z_{no}) \\ 0 \\ P(T=t_b|Z=z_{no}) - P(T=t_b|Z=z_a) \\ P(T=t_b|Z=z_a) \\ 0 \\ 0 \end{pmatrix} \odot \begin{pmatrix} 0 \\ P(Z=z_{bc}) \\ 0 \\ P(Z=z_{no})+P(Z=z_{bc}) \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

$$(\boldsymbol{B}_{t_c}^{+}\boldsymbol{P}_Z(t_c)) \odot (\boldsymbol{B}_{t_c}'\boldsymbol{Pr_Z}) = \begin{pmatrix} 0 \\ 0 \\ P(T=t_c|Z=z_{bc}) - P(T=t_c|Z=z_{no}) \\ 0 \\ 0 \\ P(T=t_c|Z=z_{no}) - P(T=t_b|Z=z_b) \\ P(T=t_c|Z=z_{no}) \end{pmatrix} \odot \begin{pmatrix} 0 \\ 0 \\ P(Z=z_{bc}) \\ 0 \\ 0 \\ P(Z=z_{no}) \\ 1 \end{pmatrix}.$$

Also, we have that $\boldsymbol{b}_{t_a}(1) = [0,0,0,1,0,1,0]'$ and $\boldsymbol{b}_{t_a}(2) = [0,1,1,0,0,0,0]'$, thus:

$$\frac{\boldsymbol{b}_{t_a}(1)\sum_{t\in\{t_b,t_c\}}\left(\boldsymbol{B}_{t'}^{+}\boldsymbol{Q}_Z(t')\right)\odot\left(\boldsymbol{B}_{t'}'\boldsymbol{Pr_Z}\right)}{\boldsymbol{b}_{t_a}(1)\sum_{t\in\{t_b,t_c\}}\left(\boldsymbol{B}_{t'}^{+}\boldsymbol{P}_Z(t')\right)\odot\left(\boldsymbol{B}_{t'}'\boldsymbol{Pr_Z}\right)} = \frac{\begin{pmatrix} \big(E(Y\cdot\mathbf{1}[T=t_b]|Z=z_{no}) - E(Y\cdot\mathbf{1}[T=t_b]|Z=z_a)\big)\big(P(Z=z_{no})+P(Z=z_{bc})\big) \\ +\big(E(Y\cdot\mathbf{1}[T=t_c]|Z=z_{no}) - E(Y\cdot\mathbf{1}[T=t_b]|Z=z_b)\big)P(Z=z_{no}) \end{pmatrix}}{\begin{pmatrix} \big(P(T=t_b|Z=z_{no}) - P(T=t_b|Z=z_a)\big)\big(P(Z=z_{no})+P(Z=z_{bc})\big) \\ +\big(P(T=t_c|Z=z_{no}) - P(T=t_b|Z=z_b)\big)P(Z=z_{no}) \end{pmatrix}} \quad \text{(A.90)}$$

$$\frac{\boldsymbol{b}_{t_a}(2)\sum_{t\in\{t_b,t_c\}}\left(\boldsymbol{B}_{t'}^{+}\boldsymbol{Q}_Z(t')\right)\odot\left(\boldsymbol{B}_{t'}'\boldsymbol{Pr_Z}\right)}{\boldsymbol{b}_{t_a}(2)\sum_{t\in\{t_b,t_c\}}\left(\boldsymbol{B}_{t'}^{+}\boldsymbol{P}_Z(t')\right)\odot\left(\boldsymbol{B}_{t'}'\boldsymbol{Pr_Z}\right)} = \frac{\begin{pmatrix} \big(E(Y\cdot\mathbf{1}[T=t_b]|Z=z_{bc}) - E(Y\cdot\mathbf{1}[T=t_b]|Z=z_{no})\big)P(Z=z_{bc}) \\ +\big(E(Y\cdot\mathbf{1}[T=t_c]|Z=z_{bc}) - E(Y\cdot\mathbf{1}[T=t_c]|Z=z_{no})\big)P(Z=z_{bc}) \end{pmatrix}}{\begin{pmatrix} \big(P(T=t_b|Z=z_{bc}) - P(T=t_b|Z=z_{no})\big)P(Z=z_{bc}) \\ +\big(P(T=t_c|Z=z_{bc}) - P(T=t_c|Z=z_{no})\big)P(Z=z_{bc}) \end{pmatrix}} \quad \text{(A.91)}$$

Parameter $E(Y(\bar{t}_a)|t_a\text{-Switchers})$ is given by the sum of the expressions in (A.90) and (A.91) divided by the probability in (A.89).

B Directed Graphs for IV and Strata

This section presents IV model (1)–(6) as a directed acyclic graph and introduces strata into this framework. The IV model defined in (1)–(3) can be equivalently restated as:

$$T = f_T(Z, \boldsymbol{V}), \quad Y = f_Y(T, \boldsymbol{V}, \epsilon_Y), \quad \boldsymbol{V} = f_{\boldsymbol{V}}(\epsilon_{\boldsymbol{V}}), \quad Z = f_Z(\epsilon_Z), \tag{B.1}$$

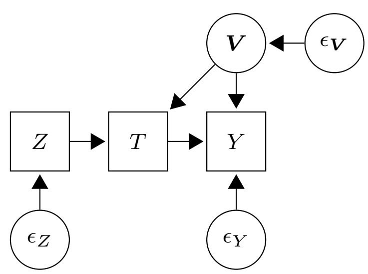
where

$$(\epsilon_Y, \epsilon_{\boldsymbol{V}}, \epsilon_Z) \text{ are mutually independent error terms.} \tag{B.2}$$

Equations (B.1) specify the causal directions of the IV model (1)–(3). Instrument Z affects T but does not directly affect Y. Z affects Y only through its effect on T.

Causal relationships are indicated by directed arrows. Unobserved variables are represented by circles. Squares represent observed variables. This leads to the Directed Acyclic Graph B.1.

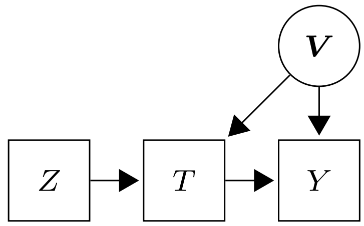
Figure B.1: **DAG for the Standard Choice Model with Instrumental Variables**



This figure represents the confounding model with instrumental variables as a DAG. Arrows represent direct causal relations. Circles represent unobserved variables. Squares represent observed variables.

It is standard (but sometimes confusing) not to depict the error terms $(\epsilon_Z, \epsilon_T, \epsilon_Y, \epsilon_{\boldsymbol{V}})$ which, in Figure B.1, are represented as circles with arrows pointing to their associated variables, and we follow this convention in Figure B.2:

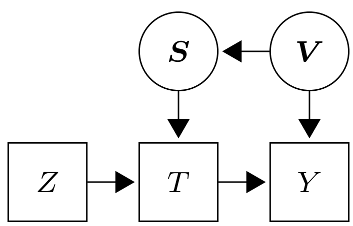
Figure B.2: **A Causal Model with Instrumental Variables**



This figure represents the confounding model with instrumental variables as a DAG. Arrows represent direct causal relations. Circles represent unobserved variables and the ϵ are kept implicit. Squares represent observed variables.

We note that strata add no new information not already present in the DAG of Figure B.2. See Figure B.3.

Figure B.3: **IV Model with Response Vector $\boldsymbol{S}$**



Notes: This figure represents the confounding model with instrumental variables and response vector $\boldsymbol{S}$ as a DAG. Arrows represent direct causal relations. Circles represent unobserved variables and the ϵ are kept implicit. Squares represent observed variables.

Strata are just representations of the model that give a coarser summary of the influence of $\boldsymbol{V}$ on T that is useful as a control function.

C Examples of the Benefits of Using Indicator Functions for the General Unordered Model

This section[6] provides a simple example that clarifies the ideas discussed in the main paper. Let $T_\omega \in \{1,2,3\}$ be the choice made by agent ω. Let $Z_\omega \in \{z_0, z_1\}$ be an instrumental variable. Y_ω is the observed outcome. $Y_\omega(t,z)$ denotes the counterfactual outcome when T_ω is fixed at $t \in \{1,2,3\}$ and Z_ω is fixed at $z \in \{z_0, z_1\}$. $T_\omega(z)$ is the counterfactual choice when the instrument is fixed at $z \in \{z_0, z_1\}$. Exclusion restrictions require that $Y_\omega(t, z_0) = Y_\omega(t, z_1) \equiv Y_\omega(t)$. Independence relation $\big(Y(t), T(z)\big) \perp\!\!\!\perp Z\ ;\ z \in \{z_0, z_1\}, t \in \{1,2,3\}$ is a version of *random assignment*.

If choices are ordered, one can invoke Ordered Monotonicity: $T_\omega(z_1) \geq T_\omega(z_0)$ for all ω. Under it, $E(Y|Z=z_1) - E(Y|Z=z_0)$ identifies:

$$\begin{aligned}
&E(Y|Z=z_1) - E(Y|Z=z_0) \\
=&\sum_{t=1}^{2}\sum_{t'=t+1}^{3} E(Y(t') - Y(t)|T(z_1)=t', T(z_0)=t)P(T(z_1)=t', T(z_0)=t). \qquad \text{(C.1)}
\end{aligned}$$

This is the gain (over all possible outcomes) arising from a change in the instrument from $Z = z_0$ to $Z = z_1$.

Proof. Ordered Monotonicity, that is, $T_\omega(z_1) \geq T_\omega(z_0)$ for all ω, generates the following response matrix:

$$\boldsymbol{R} = \begin{array}{c} \begin{matrix} \boldsymbol{s}_1 & \boldsymbol{s}_2 & \boldsymbol{s}_3 & \boldsymbol{s}_4 & \boldsymbol{s}_5 & \boldsymbol{s}_6 \end{matrix} \\ \begin{bmatrix} 1 & 2 & 3 & 1 & 1 & 2 \\ 1 & 2 & 3 & 2 & 3 & 3 \end{bmatrix} \end{array} \begin{matrix} \text{values for } T(z_0) \\ \text{values for } T(z_1) \end{matrix}. \qquad \text{(C.2)}$$

[6] This sub-appendix was motivated by the comments of Elie Tamer.

From Equation (14) of Theorem **T-1**, we have the following relationships:

$$
\begin{aligned}
E(Y \cdot \mathbf{1}[T=1]|Z=z_0) &= E(Y(1)|\boldsymbol{S}=\boldsymbol{s}_1)P(\boldsymbol{S}=\boldsymbol{s}_1) + E(Y(1)|\boldsymbol{S}=\boldsymbol{s}_4)P(\boldsymbol{S}=\boldsymbol{s}_4) \\
&\quad + E(Y(1)|\boldsymbol{S}=\boldsymbol{s}_5)P(\boldsymbol{S}=\boldsymbol{s}_5) \\
E(Y \cdot \mathbf{1}[T=2]|Z=z_0) &= E(Y(2)|\boldsymbol{S}=\boldsymbol{s}_2)P(\boldsymbol{S}=\boldsymbol{s}_2) + E(Y(2)|\boldsymbol{S}=\boldsymbol{s}_6)P(\boldsymbol{S}=\boldsymbol{s}_6) \\
E(Y \cdot \mathbf{1}[T=3]|Z=z_0) &= E(Y(3)|\boldsymbol{S}=\boldsymbol{s}_3)P(\boldsymbol{S}=\boldsymbol{s}_3)
\end{aligned}
$$

Also

$$
\begin{aligned}
E(Y \cdot \mathbf{1}[T=1]|Z=z_1) &= E(Y(1)|\boldsymbol{S}=\boldsymbol{s}_1)P(\boldsymbol{S}=\boldsymbol{s}_1) \\
E(Y \cdot \mathbf{1}[T=2]|Z=z_1) &= E(Y(2)|\boldsymbol{S}=\boldsymbol{s}_2)P(\boldsymbol{S}=\boldsymbol{s}_2) + E(Y(2)|\boldsymbol{S}=\boldsymbol{s}_4)P(\boldsymbol{S}=\boldsymbol{s}_4) \\
E(Y \cdot \mathbf{1}[T=3]|Z=z_1) &= E(Y(3)|\boldsymbol{S}=\boldsymbol{s}_3)P(\boldsymbol{S}=\boldsymbol{s}_3) + E(Y(3)|\boldsymbol{S}=\boldsymbol{s}_5)P(\boldsymbol{S}=\boldsymbol{s}_5) \\
&\quad + E(Y(3)|\boldsymbol{S}=\boldsymbol{s}_6)P(\boldsymbol{S}=\boldsymbol{s}_6)
\end{aligned}
$$

But we can express $E(Y|Z=z)$ as $E(Y|Z=z) = \sum_{t=1}^{3} E(Y \cdot \mathbf{1}[T=t]|Z=z)$, thus:

$$
E(Y|Z=z_1) - E(Y|Z=z_0) = \sum_{t=1}^{3} E(Y \cdot \mathbf{1}[T=t]|Z=z_1) - E(Y \cdot \mathbf{1}[T=t]|Z=z_0),
$$

where

$$
\begin{aligned}
E(Y \cdot \mathbf{1}[T=1]|Z=z_1) - E(Y \cdot \mathbf{1}[T=1]|Z=z_0) &= -\big(E(Y(1)|\boldsymbol{S}=\boldsymbol{s}_4)P(\boldsymbol{S}=\boldsymbol{s}_4) + E(Y(1)|\boldsymbol{S}=\boldsymbol{s}_5)P(\boldsymbol{S}=\boldsymbol{s}_5)\big) \\
E(Y \cdot \mathbf{1}[T=2]|Z=z_1) - E(Y \cdot \mathbf{1}[T=2]|Z=z_0) &= E(Y(2)|\boldsymbol{S}=\boldsymbol{s}_4)P(\boldsymbol{S}=\boldsymbol{s}_4) - E(Y(2)|\boldsymbol{S}=\boldsymbol{s}_6)P(\boldsymbol{S}=\boldsymbol{s}_6) \\
E(Y \cdot \mathbf{1}[T=3]|Z=z_1) - E(Y \cdot \mathbf{1}[T=3]|Z=z_0) &= E(Y(3)|\boldsymbol{S}=\boldsymbol{s}_5)P(\boldsymbol{S}=\boldsymbol{s}_5) + E(Y(3)|\boldsymbol{S}=\boldsymbol{s}_6)P(\boldsymbol{S}=\boldsymbol{s}_6)
\end{aligned}
$$

The summation of all the terms above generates Equation (C.1):

$$
E(Y|Z=z_1) - E(Y|Z=z_0) = \sum_{t=1}^{2} \sum_{t'=t+1}^{3} E(Y(t') - Y(t)|T(z_1)=t', T(z_0)=t)P(T(z_1)=t', T(z_0)=t).
$$

□

Our analysis requires no order on T. Our analysis is based on $\mathbf{1}[T_\omega(z)=t]$, which takes value 1 if $T_\omega(z)=t$ and zero otherwise. If choices are not ordered, economic theory can be used to justify monotonicity relationships based on indicator functions generated by $T_\omega(z)$ and not the $T_\omega(z)$. Using this notation, an example of monotonicity expressed in terms of indicator functions is the order $\mathbf{1}[T_\omega(z_1)=t] \geq \mathbf{1}[T_\omega(z_0)=t]; t \in \{1,3\}$, assumed to hold for all ω. For this case, $E(Y|Z=z_1) - E(Y|Z=z_0)$ identifies:

$$\begin{aligned} & E(Y|Z=z_1) - E(Y|Z=z_0) \\ &= \sum_{t\in\{1,3\}} E(Y(t)-Y(2)|T(z_1)=t, T(z_0)=2)P(T(z_1)=t, T(z_0)=2). \end{aligned} \tag{C.3}$$

Proof. Suppose instead that the following monotonicity relation holds:

$$\mathbf{1}[T_\omega(z_1)=t] \geq \mathbf{1}[T_\omega(z_0)=t]; t \in \{1,3\} \text{ for all } \omega.$$

In this case, the associated response matrix is given by:

$$\boldsymbol{R} = \begin{matrix} \boldsymbol{s}_1 & \boldsymbol{s}_2 & \boldsymbol{s}_3 & \boldsymbol{s}_4 & \boldsymbol{s}_5 & \\ \left[\begin{matrix} 1 & 2 & 3 & 2 & 2 \\ 1 & 2 & 3 & 1 & 3 \end{matrix}\right. & & & & \left.\right] & \begin{matrix} \text{values for } T(z_0) \\ \text{values for } T(z_1) \end{matrix} \end{matrix}. \tag{C.4}$$

According to Equation (14) of Theorem **T-1**, we have the following relations:

$$E(Y\cdot\mathbf{1}[T=1]|Z=z_0) = E(Y(1)|\boldsymbol{S}=\boldsymbol{s}_1)P(\boldsymbol{S}=\boldsymbol{s}_1)$$

$$E(Y\cdot\mathbf{1}[T=2]|Z=z_0) = E(Y(2)|\boldsymbol{S}=\boldsymbol{s}_2)P(\boldsymbol{S}=\boldsymbol{s}_2) + E(Y(2)|\boldsymbol{S}=\boldsymbol{s}_4)P(\boldsymbol{S}=\boldsymbol{s}_4) + E(Y(2)|\boldsymbol{S}=\boldsymbol{s}_5)P(\boldsymbol{S}=\boldsymbol{s}_5)$$

$$E(Y\cdot\mathbf{1}[T=3]|Z=z_0) = E(Y(3)|\boldsymbol{S}=\boldsymbol{s}_3)P(\boldsymbol{S}=\boldsymbol{s}_3)$$

Also $E(Y\cdot\mathbf{1}[T=1]|Z=z_1) = E(Y(1)|\boldsymbol{S}=\boldsymbol{s}_1)P(\boldsymbol{S}=\boldsymbol{s}_1) + E(Y(1)|\boldsymbol{S}=\boldsymbol{s}_4)P(\boldsymbol{S}=\boldsymbol{s}_4)$

$$E(Y\cdot\mathbf{1}[T=2]|Z=z_1) = E(Y(2)|\boldsymbol{S}=\boldsymbol{s}_2)P(\boldsymbol{S}=\boldsymbol{s}_2)$$

$$E(Y\cdot\mathbf{1}[T=3]|Z=z_1) = E(Y(3)|\boldsymbol{S}=\boldsymbol{s}_3)P(\boldsymbol{S}=\boldsymbol{s}_3) + E(Y(3)|\boldsymbol{S}=\boldsymbol{s}_5)P(\boldsymbol{S}=\boldsymbol{s}_5)$$

In the same fashion as previous analysis, we can express $E(Y|Z=z)$ as $E(Y|Z=z) = \sum_{t=1}^{3} E(Y\cdot\mathbf{1}[T=t]|Z=z)$, thus:

$$E(Y|Z=z_1) - E(Y|Z=z_0) = \sum_{t=1}^{3} E(Y\cdot\mathbf{1}[T=t]|Z=z_1) - E(Y\cdot\mathbf{1}[T=t]|Z=z_0),$$

where

$$E(Y\cdot\mathbf{1}[T=1]|Z=z_1) - E(Y\cdot\mathbf{1}[T=1]|Z=z_0) = E(Y(1)|\boldsymbol{S}=\boldsymbol{s}_4)P(\boldsymbol{S}=\boldsymbol{s}_4)$$

$$E(Y\cdot\mathbf{1}[T=2]|Z=z_1) - E(Y\cdot\mathbf{1}[T=2]|Z=z_0) = -\big(E(Y(2)|\boldsymbol{S}=\boldsymbol{s}_4)P(\boldsymbol{S}=\boldsymbol{s}_4) + E(Y(2)|\boldsymbol{S}=\boldsymbol{s}_5)P(\boldsymbol{S}=\boldsymbol{s}_5)\big)$$

$$E(Y\cdot\mathbf{1}[T=3]|Z=z_1) - E(Y\cdot\mathbf{1}[T=3]|Z=z_0) = E(Y(3)|\boldsymbol{S}=\boldsymbol{s}_5)P(\boldsymbol{S}=\boldsymbol{s}_5)$$

The summation of all the terms above generates Equation (C.3):

$$E(Y|Z=z_1) - E(Y|Z=z_0) = \sum_{t\in\{1,3\}} E(Y(t)-Y(2)|T(z_1)=t, T(z_0)=2)P(T(z_1)=t, T(z_0)=2).$$

□

D Response Type Elimination due to Monotonic Relationships and Choice Restrictions

Table D.1 presents the 27 response types of the multiple treatment model analyzed in Section 5. Table D.1 considers the case where Z takes values in $\text{supp}(Z) = \{z_{no}, z_a, z_{bc}\}$, and displays the restrictions on admissible strata imposed by the relationships presented at the base of the table.

Table D.1: Response Types and Elimination of Response Types for $\boldsymbol{S}$ when $\text{supp}(Z) = \{z_{no}, z_a, z_{bc}\}$ Under Restrictions (41)–(44)

Response Types	Values Instrumental Variable Z takes: No Voucher $T(z_{no})$	Voucher for a $T(z_a)$	Voucher for b or c $T(z_{bc})$	Restriction Analysis: Relation 1	Relation 2	Relation 3	Relation 4
1	t_a	t_a	t_a	✓	✓	✓	✓
2	t_a	t_a	t_b	✓	✓	✓	✓
3	t_a	t_a	t_c	✓	✓	✓	✓
4	t_a	t_b	t_a	✗	✗	✓	✗
5	t_a	t_b	t_b	✗	✓	✓	✓
6	t_a	t_b	t_c	✗	✓	✓	✓
7	t_a	t_c	t_a	✗	✗	✓	✗
8	t_a	t_c	t_b	✗	✓	✓	✓
9	t_a	t_c	t_c	✗	✓	✓	✓
10	t_b	t_a	t_a	✓	✓	✗	✓
11	t_b	t_a	t_b	✓	✓	✓	✓
12	t_b	t_a	t_c	✓	✓	✓	✓
13	t_b	t_b	t_a	✓	✗	✗	✗
14	t_b	t_b	t_b	✓	✓	✓	✓
15	t_b	t_b	t_c	✓	✓	✓	✓
16	t_b	t_c	t_a	✓	✗	✗	✗
17	t_b	t_c	t_b	✓	✓	✓	✓
18	t_b	t_c	t_c	✓	✓	✓	✓
19	t_c	t_a	t_a	✓	✓	✗	✓
20	t_c	t_a	t_b	✓	✓	✓	✓
21	t_c	t_a	t_c	✓	✓	✓	✓
22	t_c	t_b	t_a	✓	✗	✗	✗
23	t_c	t_b	t_b	✓	✓	✓	✓
24	t_c	t_b	t_c	✓	✓	✓	✓
25	t_c	t_c	t_a	✓	✗	✗	✗
26	t_c	t_c	t_b	✓	✓	✓	✓
27	t_c	t_c	t_c	✓	✓	✓	✓

This table presents all possible values that the response variable $\boldsymbol{S}$ can possibly take when instrumental variable Z ranges over $\text{supp}(Z) = \{z_{no}, z_a, z_{bc}\}$ and treatment status T ranges over $\{t_{no}, t_a, t_{bc}\}$. The first column enumerates the 27 possible response types. Columns 2 to 4 indicate the response types according to the vector of the values that $[T(z_{no}), T(z_a), T(z_{bc})]$ takes. Columns 5 to 8 indicate whether the response type violates any of the following monotonicity relations:

Relationship 1	$\mathbf{1}[T_\omega(z_{no}) = t_a]$	$\leq$	$\mathbf{1}[T_\omega(z_a) = t_a]$
Relationship 2	$\mathbf{1}[T_\omega(z_{bc}) = t_a]$	$\leq$	$\mathbf{1}[T_\omega(z_a) = t_a]$
Relationship 3	$\mathbf{1}[T_\omega(z_{no}) \in \{b,c\}]$	$\leq$	$\mathbf{1}[T_\omega(z_{bc}) \in \{b,c\}]$
Relationship 4	$\mathbf{1}[T_\omega(z_a) \in \{b,c\}]$	$\leq$	$\mathbf{1}[T_\omega(z_{bc}) \in \{b,c\}]$

A check mark sign indicates that the associated response type does not violate the relation. A cross sign indicates that the associated response type violates the relation.

The table shows which response types violate the monotonicity relationships in (41)–(44), which are restated below for sake of clarity:

$$
\begin{array}{lrcl}
\text{Relationship 1} & \mathbf{1}[T_\omega(z_{no}) = t_a] & \leq & \mathbf{1}[T_\omega(z_a) = t_a] \\
\text{Relationship 2} & \mathbf{1}[T_\omega(z_{bc}) = t_a] & \leq & \mathbf{1}[T_\omega(z_a) = t_a] \\
\text{Relationship 3} & \mathbf{1}[T_\omega(z_{no}) \in \{b, c\}] & \leq & \mathbf{1}[T_\omega(z_{bc}) \in \{b, c\}] \\
\text{Relationship 4} & \mathbf{1}[T_\omega(z_a) \in \{b, c\}] & \leq & \mathbf{1}[T_\omega(z_{bc}) \in \{b, c\}]
\end{array}
$$

The elimination process described in Table D.1 generates the response matrix $\boldsymbol{R}$ of Table 1 in the text reproduced as Table D.2.

Table D.2: Response Matrix Generated by Monotonicity relationship (41)–(44)

Instrumental Variables	Counterfactual Choices	Response Types of $\boldsymbol{S}$														
		$\boldsymbol{s}_1$	$\boldsymbol{s}_2$	$\boldsymbol{s}_3$	$\boldsymbol{s}_4$	$\boldsymbol{s}_5$	$\boldsymbol{s}_6$	$\boldsymbol{s}_7$	$\boldsymbol{s}_8$	$\boldsymbol{s}_9$	$\boldsymbol{s}_{10}$	$\boldsymbol{s}_{11}$	$\boldsymbol{s}_{12}$	$\boldsymbol{s}_{13}$	$\boldsymbol{s}_{14}$	$\boldsymbol{s}_{15}$
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

Table D.3 displays the response types that violate the following choice restrictions:

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	$Ch_\omega(z_a) = t_a$

The elimination process described in Table D.3 generates the response matrix $\boldsymbol{R}$ of Table D.4 that has 11 response types.

Table D.3: Restrictions on Response Vector $\boldsymbol{S}$ for $\text{supp}(Z) = \{z_{no}, z_a, z_{bc}\}$ Under 6 Choice Restrictions below

Response Types	Values Instrumental Variable Z takes			Restriction Analysis					
	No Voucher $T(z_{no})$	Voucher for a $T(z_a)$	Voucher for b or c $T(z_{bc})$	Res. 1	Res. 2	Res. 3	Res. 4	Res. 5	Res. 6
1	t_a	t_a	t_a	✓	✓	✓	✓	✓	✓
2	t_a	t_a	t_b	✓	✓	✓	✓	✓	✓
3	t_a	t_a	t_c	✓	✓	✓	✓	✓	✓
4	t_a	t_b	t_a	✗	✓	✓	✗	✓	✗
5	t_a	t_b	t_b	✗	✓	✓	✗	✓	✓
6	t_a	t_b	t_c	✗	✓	✓	✗	✓	✓
7	t_a	t_c	t_a	✗	✓	✓	✓	✗	✗
8	t_a	t_c	t_b	✗	✓	✓	✓	✗	✓
9	t_a	t_c	t_c	✗	✓	✓	✓	✗	✓
10	t_b	t_a	t_a	✓	✗	✓	✓	✓	✗
11	t_b	t_a	t_b	✓	✓	✓	✓	✓	✓
12	t_b	t_a	t_c	✓	✓	✓	✓	✓	✓
13	t_b	t_b	t_a	✓	✗	✓	✗	✓	✗
14	t_b	t_b	t_b	✓	✓	✓	✓	✓	✓
15	t_b	t_b	t_c	✓	✓	✓	✓	✓	✓
16	t_b	t_c	t_a	✓	✗	✓	✓	✗	✗
17	t_b	t_c	t_b	✓	✗	✓	✓	✗	✓
18	t_b	t_c	t_c	✓	✗	✓	✓	✗	✓
19	t_c	t_a	t_a	✓	✓	✗	✓	✓	✗
20	t_c	t_a	t_b	✓	✓	✓	✓	✓	✓
21	t_c	t_a	t_c	✓	✓	✓	✓	✓	✓
22	t_c	t_b	t_a	✓	✓	✗	✗	✓	✗
23	t_c	t_b	t_b	✓	✓	✗	✗	✓	✓
24	t_c	t_b	t_c	✓	✓	✗	✗	✓	✓
25	t_c	t_c	t_a	✓	✓	✗	✓	✗	✗
26	t_c	t_c	t_b	✓	✓	✓	✓	✓	✓
27	t_c	t_c	t_c	✓	✓	✓	✓	✓	✓

This table presents all possible values that the response variable $\boldsymbol{S}$ can possibly take when instrumental variable Z ranges over $\text{supp}(Z) = \{z_{no}, z_a, z_{bc}\}$ and treatment status T ranges over $\{t_a, t_b, t_c\}$. The first column enumerates the 27 possible response types. Columns 2 to 4 indicate the response types according to the vector of the values that $[T(z_{no}), T(z_a), T(z_{bc})]$ takes. The remaining six columns indicate whether the response type violates any of the following choice restrictions respectively:

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	$Ch_\omega(z_a) = t_a$

A check mark sign indicates that the associated response type does not violate the choice restriction. A cross sign indicates that the associated response type violates the choice restriction.

Table D.4 displays the response types arising from applying WARP (49) for the budget sets (46)–(48):

Table D.4: Admissible Response Types under WARP (49)

Instrumental Variables	Count. Choices	Response Types of $\boldsymbol{S}$										
		$\boldsymbol{s}_1$	$\boldsymbol{s}_2$	$\boldsymbol{s}_3$	$\boldsymbol{s}_4$	$\boldsymbol{s}_5$	$\boldsymbol{s}_6$	$\boldsymbol{s}_7$	$\boldsymbol{s}_8$	$\boldsymbol{s}_9$	$\boldsymbol{s}_{10}$	$\boldsymbol{s}_{11}$
No Voucher	$T(z_{no})$	t_a	t_a	t_a	t_b	t_b	t_b	t_b	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_a	t_a	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

Table D.5 illustrates the response types for the choice restrictions of Table 2 in the text:

Table D.5: Elimination of Response Types for $\text{supp}(Z) = \{z_{no}, z_a, z_{bc}\}$ and 7 Choice Restrictions

Response Types	Values Instrumental Variable Z takes			Restriction Analysis						
	No Voucher $T(z_{no})$	Voucher for a $T(z_a)$	Voucher for b or c $T(z_{bc})$	Res. 1	Res. 2	Res. 3	Res. 4	Res. 5	Res. 6	Res. 7
1	t_a	t_a	t_a	✓	✓	✓	✓	✓	✓	✓
2	t_a	t_a	t_b	✓	✓	✓	✓	✓	✓	✓
3	t_a	t_a	t_c	✓	✓	✓	✓	✓	✓	✓
4	t_a	t_b	t_a	✗	✓	✓	✗	✓	✗	✓
5	t_a	t_b	t_b	✗	✓	✓	✗	✓	✓	✓
6	t_a	t_b	t_c	✗	✓	✓	✗	✓	✓	✓
7	t_a	t_c	t_a	✗	✓	✓	✓	✗	✗	✓
8	t_a	t_c	t_b	✗	✓	✓	✓	✗	✓	✓
9	t_a	t_c	t_c	✗	✓	✓	✓	✗	✓	✓
10	t_b	t_a	t_a	✓	✗	✓	✓	✓	✗	✗
11	t_b	t_a	t_b	✓	✓	✓	✓	✓	✓	✓
12	t_b	t_a	t_c	✓	✓	✓	✓	✓	✓	✗
13	t_b	t_b	t_a	✓	✗	✓	✗	✓	✗	✗
14	t_b	t_b	t_b	✓	✓	✓	✓	✓	✓	✓
15	t_b	t_b	t_c	✓	✓	✓	✓	✓	✓	✗
16	t_b	t_c	t_a	✓	✗	✓	✓	✗	✗	✗
17	t_b	t_c	t_b	✓	✗	✓	✓	✗	✓	✓
18	t_b	t_c	t_c	✓	✗	✓	✓	✗	✓	✗
19	t_c	t_a	t_a	✓	✓	✗	✓	✓	✗	✗
20	t_c	t_a	t_b	✓	✓	✓	✓	✓	✓	✗
21	t_c	t_a	t_c	✓	✓	✓	✓	✓	✓	✓
22	t_c	t_b	t_a	✓	✓	✗	✗	✓	✗	✗
23	t_c	t_b	t_b	✓	✓	✗	✗	✓	✓	✗
24	t_c	t_b	t_c	✓	✓	✗	✗	✓	✓	✓
25	t_c	t_c	t_a	✓	✓	✗	✓	✗	✗	✗
26	t_c	t_c	t_b	✓	✓	✓	✓	✓	✓	✗
27	t_c	t_c	t_c	✓	✓	✓	✓	✓	✓	✓

This table presents all possible values that the response variable $\boldsymbol{S}$ can possibly take when instrumental variable Z ranges over $\text{supp}(Z) = \{z_{no}, z_a, z_{bc}\}$ and treatment status T ranges over $\{t_a, t_b, t_c\}$. The first column enumerates the 27 possible response types. Columns 2 to 4 indicate the response types according to the vector of the values that $[T(z_{no}), T(z_a), T(z_{bc})]$ takes. The remaining seven columns indicate whether the response type violates any of the following choice restrictions respectively:

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**	$Ch_\omega(z_{bc}) = t_a$	$\Rightarrow$	$Ch_\omega(z_{no}) = t_a = Ch_\omega(z_a) = t_a$		
Choice Restriction **7**	$Ch_\omega(z_{no}) \neq t_a$	$\Rightarrow$	$Ch_\omega(z_{bc}) = Ch_\omega(z_{no})$		

A check mark sign indicates that the associated response type does not violate the choice restriction. A cross sign indicates that the associate response type violates the choice restriction.

Table D.6 demonstrates the restrictions on admissible response types from the restrictions imposed in Table 4 in the text.

Table D.6: Elimination Response Types Under Unordered Monotonicity

Response Types	Values Z takes			Elimination of Response Types								
	$T(z_{no})$	$T(z_a)$	$T(z_{bc})$	Rel. 1	Rel. 2	Rel. 3	Rel. 4	Rel. 5	Rel. 6	Rel. 7	Rel. 8	Rel. 9
1	t_a	t_a	t_a	✓	✓	✓	✓	✓	✓	✓	✓	✓
2	t_a	t_a	t_b	✓	✓	✓	✓	✓	✓	✓	✓	✓
3	t_a	t_a	t_c	✓	✓	✓	✓	✓	✓	✓	✓	✓
4	t_a	t_b	t_a	✗	✓	✗	✗	✓	✗	✓	✓	✓
5	t_a	t_b	t_b	✗	✓	✓	✗	✓	✓	✓	✓	✓
6	t_a	t_b	t_c	✗	✓	✓	✗	✓	✗	✓	✓	✓
7	t_a	t_c	t_a	✗	✓	✗	✓	✓	✓	✗	✓	✗
8	t_a	t_c	t_b	✗	✓	✓	✓	✓	✓	✗	✓	✗
9	t_a	t_c	t_c	✗	✓	✓	✓	✓	✓	✗	✓	✓
10	t_b	t_a	t_a	✓	✗	✓	✓	✗	✓	✓	✓	✓
11	t_b	t_a	t_b	✓	✓	✓	✓	✓	✓	✓	✓	✓
12	t_b	t_a	t_c	✓	✓	✓	✓	✗	✓	✓	✓	✓
13	t_b	t_b	t_a	✓	✗	✗	✓	✗	✗	✓	✓	✓
14	t_b	t_b	t_b	✓	✓	✓	✓	✓	✓	✓	✓	✓
15	t_b	t_b	t_c	✓	✓	✓	✓	✗	✗	✓	✓	✓
16	t_b	t_c	t_a	✓	✗	✗	✓	✗	✓	✗	✓	✗
17	t_b	t_c	t_b	✓	✓	✓	✓	✓	✓	✗	✓	✗
18	t_b	t_c	t_c	✓	✓	✓	✓	✗	✓	✗	✓	✓
19	t_c	t_a	t_a	✓	✗	✓	✓	✓	✓	✓	✗	✓
20	t_c	t_a	t_b	✓	✓	✓	✓	✓	✓	✓	✗	✓
21	t_c	t_a	t_c	✓	✓	✓	✓	✓	✓	✓	✓	✓
22	t_c	t_b	t_a	✓	✗	✗	✗	✓	✗	✓	✗	✓
23	t_c	t_b	t_b	✓	✓	✓	✗	✓	✓	✓	✗	✓
24	t_c	t_b	t_c	✓	✓	✓	✗	✓	✗	✓	✓	✓
25	t_c	t_c	t_a	✓	✗	✗	✓	✓	✓	✓	✗	✗
26	t_c	t_c	t_b	✓	✓	✓	✓	✓	✓	✓	✗	✗
27	t_c	t_c	t_c	✓	✓	✓	✓	✓	✓	✓	✓	✓

This table presents all possible values that the response variable $\boldsymbol{S}$ can possibly take when instrumental variable Z ranges over, $\text{supp}(Z) = \{z_{no}, z_a, z_{bc}\}$ and treatment status T ranges over $\text{supp}(T) = \{t_a, t_b, t_c\}$. The first column enumerates the 27 possible response types. Columns 2 to 4 presents the response types according to the vector of the values that $[T_\omega(z_{no}) \, , \, T_\omega(z_a) \, , \, T_\omega(z_{bc})]$ takes. Columns 5 to 13 indicate whether the response type violates any of the following monotonicity relations:

Monotonicity Relation 1	$\mathbf{1}[T_\omega(z_{no}) = t_a]$	$\leq$	$\mathbf{1}[T_\omega(z_a) = t_a]$
Monotonicity Relation 2	$\mathbf{1}[T_\omega(z_{bc}) = t_a]$	$\leq$	$\mathbf{1}[T_\omega(z_{no}) = t_a]$
Monotonicity Relation 3	$\mathbf{1}[T_\omega(z_{bc}) = t_a]$	$\leq$	$\mathbf{1}[T_\omega(z_a) = t_a]$
Monotonicity Relation 4	$\mathbf{1}[T_\omega(z_a) = t_b]$	$\leq$	$\mathbf{1}[T_\omega(z_{no}) = t_b]$
Monotonicity Relation 5	$\mathbf{1}[T_\omega(z_{no}) = t_b]$	$\leq$	$\mathbf{1}[T_\omega(z_{bc}) = t_b]$
Monotonicity Relation 6	$\mathbf{1}[T_\omega(z_a) = t_b]$	$\leq$	$\mathbf{1}[T_\omega(z_{bc}) = t_b]$
Monotonicity Relation 7	$\mathbf{1}[T_\omega(z_a) = t_c]$	$\leq$	$\mathbf{1}[T_\omega(z_{no}) = t_c]$
Monotonicity Relation 8	$\mathbf{1}[T_\omega(z_{no}) = t_c]$	$\leq$	$\mathbf{1}[T_\omega(z_{bc}) = t_c]$
Monotonicity Relation 9	$\mathbf{1}[T_\omega(z_a) = t_c]$	$\leq$	$\mathbf{1}[T_\omega(z_{bc}) = t_c]$

A check mark sign indicates that the associated response type does not violates the relation. A cross sign indicates that the associated response type violates the relation.

We present additional examples next.

D.1 Another Example of Choice Restrictions That Generate an Unordered Monotonic Response

Suppose that Z takes values in $\text{supp}(Z) = \{z_{no}, z_b, z_{bc}\}$. Following the rationale of Section 5.2, assume the following budget relationships:

$$\text{Budget Relationships for } t_a \quad \Lambda_\omega(z_{no}, t_a) = \Lambda_\omega(z_b, t_a) = \Lambda_\omega(z_{bc}, t_a)$$

$$\text{Budget Relationships for } t_b \quad \Lambda_\omega(z_{no}, t_b) \subset \Lambda_\omega(z_b, t_b) = \Lambda_\omega(z_{bc}, t_b)$$

$$\text{Budget Relationships for } t_c \quad \Lambda_\omega(z_{no}, t_c) = \Lambda_\omega(z_b, t_c) \subset \Lambda_\omega(z_{bc}, t_c)$$

The budget set relationships above can be used as input to WARP (49), which generates the following choice restrictions:

Choice Restriction **1**	$Ch_\omega(z_{no}) = t_a$	$\Rightarrow$	$Ch_\omega(z_b) \neq t_c$		
Choice Restriction **2**	$Ch_\omega(z_{no}) = t_b$	$\Rightarrow$	$Ch_\omega(z_b) = t_b$	and	$Ch_\omega(z_c) \neq t_a$
Choice Restriction **3**	$Ch_\omega(z_{no}) = t_c$	$\Rightarrow$	$Ch_\omega(z_b) \neq t_a$	and	$Ch_\omega(z_c) \neq t_a$
Choice Restriction **4**	$Ch_\omega(z_b) = t_a$	$\Rightarrow$	$Ch_\omega(z_a) = t_a$	and	$Ch_\omega(z_c) \neq t_b$
Choice Restriction **5**	$Ch_\omega(z_b) = t_b$	$\Rightarrow$	$Ch_\omega(z_c) \neq t_a$		
Choice Restriction **6**	$Ch_\omega(z_b) = t_c$	$\Rightarrow$	$Ch_\omega(z_a) = t_c$	and	$Ch_\omega(z_c) = t_c$
Choice Restriction **7**	$Ch_\omega(z_{bc}) = t_a$	$\Rightarrow$	$Ch_\omega(z_a) = t_a$	and	$Ch_\omega(z_b) = t_a$
Choice Restriction **8**	$Ch_\omega(z_{bc}) = t_b$	$\Rightarrow$	$Ch_\omega(z_b) = t_b$		

The elimination process described below in Table D.7 generates the response matrix D.8:

Table D.7: Restrictions on Response Vector $\boldsymbol{S}$ for $\text{supp}(Z) = \{z_{no}, z_b, z_{bc}\}$ Under Choice Restrictions below

Response Types	Values Instrumental Variable Z takes			Choice Restriction Analysis							
	No Voucher $T(z_{no})$	Voucher for b $T(z_b)$	Voucher b and c $T(z_{bc})$	1	2	3	4	5	6	7	8
1	t_a	t_a	t_a	✓	✓	✓	✓	✓	✓	✓	✓
2	t_a	t_a	t_b	✓	✓	✓	✗	✓	✓	✓	✗
3	t_a	t_a	t_c	✓	✓	✓	✓	✓	✓	✓	✓
4	t_a	t_b	t_a	✓	✓	✓	✓	✗	✓	✗	✓
5	t_a	t_b	t_b	✓	✓	✓	✓	✓	✓	✓	✓
6	t_a	t_b	t_c	✓	✓	✓	✓	✓	✓	✓	✓
7	t_a	t_c	t_a	✗	✓	✓	✓	✓	✗	✗	✓
8	t_a	t_c	t_b	✗	✓	✓	✓	✓	✗	✓	✗
9	t_a	t_c	t_c	✗	✓	✓	✓	✓	✗	✓	✓
10	t_b	t_a	t_a	✓	✗	✓	✗	✓	✓	✗	✓
11	t_b	t_a	t_b	✓	✗	✓	✗	✓	✓	✓	✗
12	t_b	t_a	t_c	✓	✗	✓	✗	✓	✓	✓	✓
13	t_b	t_b	t_a	✓	✗	✓	✓	✗	✓	✗	✓
14	t_b	t_b	t_b	✓	✓	✓	✓	✓	✓	✓	✓
15	t_b	t_b	t_c	✓	✓	✓	✓	✓	✓	✓	✓
16	t_b	t_c	t_a	✓	✗	✓	✓	✓	✗	✗	✓
17	t_b	t_c	t_b	✓	✗	✓	✓	✓	✗	✓	✗
18	t_b	t_c	t_c	✓	✗	✓	✓	✓	✗	✓	✓
19	t_c	t_a	t_a	✓	✓	✗	✗	✓	✓	✗	✓
20	t_c	t_a	t_b	✓	✓	✗	✗	✓	✓	✓	✗
21	t_c	t_a	t_c	✓	✓	✗	✗	✓	✓	✓	✓
22	t_c	t_b	t_a	✓	✓	✗	✓	✗	✓	✗	✓
23	t_c	t_b	t_b	✓	✓	✓	✓	✓	✓	✓	✓
24	t_c	t_b	t_c	✓	✓	✓	✓	✓	✓	✓	✓
25	t_c	t_c	t_a	✓	✓	✗	✓	✓	✗	✗	✓
26	t_c	t_c	t_b	✓	✓	✓	✓	✓	✗	✓	✗
27	t_c	t_c	t_c	✓	✓	✓	✓	✓	✓	✓	✓

This table presents all possible values that the response variable $\boldsymbol{S}$ can possibly take when instrumental variable Z ranges over $\text{supp}(Z) = \{z_{no}, z_b, z_{bc}\}$ and treatment status T ranges over $\{t_a, t_b, t_c\}$. The first column enumerates the 27 possible response types. Columns 2 to 4 indicate the response types according to the vector of the values that $[T(z_{no}), T(z_b), T(z_{bc})]$ takes. The remaining eight columns indicate whether the response type violates any of the following choice restrictions:

Choice Restriction **1**	$Ch_\omega(z_{no}) = t_a$	$\Rightarrow$	$Ch_\omega(z_b) \neq t_c$		
Choice Restriction **2**	$Ch_\omega(z_{no}) = t_b$	$\Rightarrow$	$Ch_\omega(z_b) = t_b$	and	$Ch_\omega(z_c) \neq t_a$
Choice Restriction **3**	$Ch_\omega(z_{no}) = t_c$	$\Rightarrow$	$Ch_\omega(z_b) \neq t_a$	and	$Ch_\omega(z_c) \neq t_a$
Choice Restriction **4**	$Ch_\omega(z_b) = t_a$	$\Rightarrow$	$Ch_\omega(z_a) = t_a$	and	$Ch_\omega(z_c) \neq t_b$
Choice Restriction **5**	$Ch_\omega(z_b) = t_b$	$\Rightarrow$	$Ch_\omega(z_c) \neq t_a$		
Choice Restriction **6**	$Ch_\omega(z_b) = t_c$	$\Rightarrow$	$Ch_\omega(z_a) = t_c$	and	$Ch_\omega(z_c) = t_c$
Choice Restriction **7**	$Ch_\omega(z_{bc}) = t_a$	$\Rightarrow$	$Ch_\omega(z_a) = t_a$	and	$Ch_\omega(z_b) = t_a$
Choice Restriction **8**	$Ch_\omega(z_{bc}) = t_b$	$\Rightarrow$	$Ch_\omega(z_b) = t_b$		

A check mark sign indicates that the associated response type does not violate the choice restriction. A cross sign indicates that the associated response type violates the choice restriction.

Table D.8: Response Types Generated by WARP only for supp$(Z) = \{z_{no}, z_b, z_{bc}\}$.

Instrumental Variables	Count. Choices	Response Types of $\boldsymbol{S}$								
		$\boldsymbol{s}_1$	$\boldsymbol{s}_2$	$\boldsymbol{s}_3$	$\boldsymbol{s}_4$	$\boldsymbol{s}_5$	$\boldsymbol{s}_6$	$\boldsymbol{s}_7$	$\boldsymbol{s}_8$	$\boldsymbol{s}_9$
No Voucher	$T(z_{no})$	t_a	t_a	t_a	t_a	t_b	t_b	t_c	t_c	t_c
Voucher for b	$T(z_b)$	t_a	t_a	t_b	t_b	t_b	t_b	t_b	t_b	t_c
Voucher for b or c	$T(z_{bc})$	t_a	t_c	t_b	t_c	t_b	t_c	t_b	t_c	t_c

If we also assume neutral income effects, we can eliminate response types $\boldsymbol{s}_7$ and $\boldsymbol{s}_8$ above and obtain the response matrix of Table 5 of the main paper, presented here as D.9:

Table D.9: Response Types Generated by WARP and Neutral Income Effects for supp$(Z) = \{z_{no}, z_b, z_{bc}\}$.

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

D.2 An Example Where Choice Restrictions Fail to Generate Identification

Choice restrictions do not always generate response matrices that achieve identification. Table D.10 presents a response matrix generated by the reveled preference analysis when Z takes values in supp$(Z) = \{z_{no}, z_b, z_c\}$. The generated response matrix is not consistent with unordered monotonicity **A-3** and the rank of its associated binary matrix $\boldsymbol{B_T}$ is equal to 7, which is less that the number of response types, i.e., 8. Thus, response-type probabilities are not identified (Corollary **C-1**).

Table D.10: Response Types Generated by WARP and Normal Choices for supp(Z) = $\{z_{no}, z_b, z_c\}$.

Instrumental Variables	Count. Choices	Response Types of $\boldsymbol{S}$							
		s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8
No Voucher	$T(z_{no})$	t_a	t_a	t_a	t_a	t_b	t_b	t_c	t_c
Voucher for car b	$T(z_b)$	t_a	t_a	t_b	t_b	t_b	t_b	t_b	t_c
Voucher for car c	$T(z_c)$	t_a	t_c	t_a	t_c	t_b	t_c	t_c	t_c

As another example, suppose Z takes values in supp$(Z) = \{z_{no}, z_b, z_c\}$. Following the same rationale of Section 5.2, we assume the following budget relations:

$$\text{Budget Relationships for } t_a \quad \Lambda_\omega(z_{no}, t_a) = \Lambda_\omega(z_b, t_a) = \Lambda_\omega(z_c, t_a)$$

$$\text{Budget Relationships for } t_b \quad \Lambda_\omega(z_{no}, t_b) = \Lambda_\omega(z_c, t_b) \subset \Lambda_\omega(z_b, t_b)$$

$$\text{Budget Relationships for } t_c \quad \Lambda_\omega(z_{no}, t_c) = \Lambda_\omega(z_b, t_c) \subset \Lambda_\omega(z_c, t_c)$$

The budget set relations above can be used as input to WARP (49), which generates the following choice restrictions:

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

The elimination process described in Table D.11 generates the response matrix of Table D.10, also presented below as D.12:

Table D.11: Restrictions on Response Vector $\boldsymbol{S}$ for $\text{supp}(Z) = \{z_{no}, z_b, z_c\}$ Under Choice Restrictions below

Response Types	Values Instrumental Variable Z takes: No Voucher $T(z_{no})$	Voucher for b $T(z_b)$	Voucher c $T(z_c)$	Choice Restriction Analysis: 1	2	3	4	5	6	7
1	t_a	t_a	t_a	✓	✓	✓	✓	✓	✓	✓
2	t_a	t_a	t_b	✗	✓	✓	✗	✓	✓	✗
3	t_a	t_a	t_c	✓	✓	✓	✓	✓	✓	✓
4	t_a	t_b	t_a	✓	✓	✓	✓	✓	✓	✓
5	t_a	t_b	t_b	✗	✓	✓	✓	✓	✓	✗
6	t_a	t_b	t_c	✓	✓	✓	✓	✓	✓	✓
7	t_a	t_c	t_a	✗	✓	✓	✓	✗	✗	✓
8	t_a	t_c	t_b	✗	✓	✓	✓	✗	✓	✗
9	t_a	t_c	t_c	✗	✓	✓	✓	✗	✓	✓
10	t_b	t_a	t_a	✓	✗	✓	✗	✓	✗	✓
11	t_b	t_a	t_b	✓	✗	✓	✗	✓	✓	✗
12	t_b	t_a	t_c	✓	✗	✓	✗	✓	✓	✓
13	t_b	t_b	t_a	✓	✗	✓	✓	✓	✗	✓
14	t_b	t_b	t_b	✓	✓	✓	✓	✓	✓	✓
15	t_b	t_b	t_c	✓	✓	✓	✓	✓	✓	✓
16	t_b	t_c	t_a	✓	✗	✓	✓	✗	✗	✓
17	t_b	t_c	t_b	✓	✗	✓	✓	✗	✓	✗
18	t_b	t_c	t_c	✓	✗	✓	✓	✗	✓	✓
19	t_c	t_a	t_a	✓	✓	✗	✗	✓	✗	✓
20	t_c	t_a	t_b	✓	✓	✗	✗	✓	✓	✗
21	t_c	t_a	t_c	✓	✓	✗	✗	✓	✓	✓
22	t_c	t_b	t_a	✓	✓	✗	✓	✓	✗	✓
23	t_c	t_b	t_b	✓	✓	✗	✓	✓	✓	✗
24	t_c	t_b	t_c	✓	✓	✓	✓	✓	✓	✓
25	t_c	t_c	t_a	✓	✓	✗	✓	✗	✗	✓
26	t_c	t_c	t_b	✓	✓	✗	✓	✗	✓	✗
27	t_c	t_c	t_c	✓	✓	✓	✓	✓	✓	✓

This table presents all possible values that the response variable $\boldsymbol{S}$ can possibly take when instrumental variable Z ranges over $\text{supp}(Z) = \{z_{no}, z_b, z_c\}$ and treatment status T ranges over $\{t_a, t_b, t_c\}$. The first column enumerates the 27 possible response types. Columns 2 to 4 indicate the response types according to the vector of the values that $[T(z_{no})\ ,\ T(z_b)\ ,\ T(z_c)]$ takes. The remaining seven columns indicate whether the response type violates any of the following choice restrictions:

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

A check mark sign indicates that the associated response type does not violate the choice restriction. A cross sign indicates that the associated response type violates the choice restriction.

Table D.12: Response Types Generated by WARP for $\text{supp}(Z) = \{z_{no}, z_b, z_c\}$.

Instrumental Variables	Count. Choices	Response Types of $\boldsymbol{S}$ $\boldsymbol{s}_1$	$\boldsymbol{s}_2$	$\boldsymbol{s}_3$	$\boldsymbol{s}_4$	$\boldsymbol{s}_5$	$\boldsymbol{s}_6$	$\boldsymbol{s}_7$	$\boldsymbol{s}_8$
No Voucher	$T(z_{no})$	t_a	t_a	t_a	t_a	t_b	t_b	t_c	t_c
Voucher for car b	$T(z_b)$	t_a	t_a	t_b	t_b	t_b	t_b	t_b	t_c
Voucher for car c	$T(z_c)$	t_a	t_c	t_a	t_c	t_b	t_c	t_c	t_c

D.3 An Example Without Unordered Monotonicity that Generates Identification

Assuming that Z takes values in $\text{supp}(Z) = \{z_c, z_b, z_{bc}\}$. Following the same rationale of Section 5.2, we assume the following budget relations:

$$\text{Budget Relations for } t_a \quad \Lambda_\omega(z_c, t_a) = \Lambda_\omega(z_b, t_a) = \Lambda_\omega(z_{bc}, t_a)$$

$$\text{Budget Relations for } t_b \quad \Lambda_\omega(z_c, t_b) \subset \Lambda_\omega(z_b, t_b) = \Lambda_\omega(z_{bc}, t_b)$$

$$\text{Budget Relations for } t_c \quad \Lambda_\omega(z_b, t_c) \subset \Lambda_\omega(z_c, t_c) = \Lambda_\omega(z_{bc}, t_c)$$

The budget set relations above can be used as input to WARP (49), which generates the following choice restrictions:

Choice Restriction **1**	$Ch_\omega(z_c) = t_a$	$\Rightarrow$	$Ch_\omega(z_b) \neq t_c$	and	$Ch_\omega(z_c) \neq t_c$
Choice Restriction **2**	$Ch_\omega(z_c) = t_b$	$\Rightarrow$	$Ch_\omega(z_b) = t_b$	and	$Ch_\omega(z_c) = t_b$
Choice Restriction **3**	$Ch_\omega(z_c) = t_c$	$\Rightarrow$	$Ch_\omega(z_c) \neq t_a$		
Choice Restriction **4**	$Ch_\omega(z_b) = t_a$	$\Rightarrow$	$Ch_\omega(z_a) \neq t_b$	and	$Ch_\omega(z_c) \neq t_b$
Choice Restriction **5**	$Ch_\omega(z_b) = t_b$	$\Rightarrow$	$Ch_\omega(z_c) \neq t_a$		
Choice Restriction **6**	$Ch_\omega(z_b) = t_c$	$\Rightarrow$	$Ch_\omega(z_a) = t_c$	and	$Ch_\omega(z_c) = t_c$
Choice Restriction **7**	$Ch_\omega(z_{bc}) = t_a$	$\Rightarrow$	$Ch_\omega(z_a) = t_a$	and	$Ch_\omega(z_b) = t_a$
Choice Restriction **8**	$Ch_\omega(z_{bc}) = t_b$	$\Rightarrow$	$Ch_\omega(z_b) = t_b$		
Choice Restriction **9**	$Ch_\omega(z_{bc}) = t_c$	$\Rightarrow$	$Ch_\omega(z_a) = t_c$		

The elimination process described in Table D.13 generates the response matrix of Table 6, also presented below as D.14:

Table D.13: Restrictions on Response Vector $\boldsymbol{S}$ for $\text{supp}(Z) = \{z_c, z_b, z_{bc}\}$ Under Choice Restrictions below

Response Types	Values Instrumental Variable Z takes: Voucher for c $T(z_c)$	Voucher for b $T(z_b)$	Voucher for b and c $T(z_{bc})$	Choice Restriction Analysis: 1	2	3	4	5	6	7	8	9
1	t_a	t_a	t_a	✓	✓	✓	✓	✓	✓	✓	✓	✓
2	t_a	t_a	t_b	✓	✓	✓	✗	✓	✓	✓	✗	✓
3	t_a	t_a	t_c	✗	✓	✓	✓	✓	✓	✓	✓	✗
4	t_a	t_b	t_a	✓	✓	✓	✓	✗	✓	✗	✓	✓
5	t_a	t_b	t_b	✓	✓	✓	✓	✓	✓	✓	✓	✓
6	t_a	t_b	t_c	✗	✓	✓	✓	✓	✓	✓	✓	✗
7	t_a	t_c	t_a	✗	✓	✓	✓	✓	✗	✗	✓	✓
8	t_a	t_c	t_b	✗	✓	✓	✓	✓	✗	✓	✗	✓
9	t_a	t_c	t_c	✗	✓	✓	✓	✓	✗	✓	✓	✗
10	t_b	t_a	t_a	✓	✗	✓	✗	✓	✓	✗	✓	✓
11	t_b	t_a	t_b	✓	✗	✓	✗	✓	✓	✓	✗	✓
12	t_b	t_a	t_c	✓	✗	✓	✗	✓	✓	✓	✓	✗
13	t_b	t_b	t_a	✓	✗	✓	✓	✗	✓	✗	✓	✓
14	t_b	t_b	t_b	✓	✓	✓	✓	✓	✓	✓	✓	✓
15	t_b	t_b	t_c	✓	✗	✓	✓	✓	✓	✓	✓	✗
16	t_b	t_c	t_a	✓	✗	✓	✓	✓	✗	✗	✓	✓
17	t_b	t_c	t_b	✓	✗	✓	✓	✓	✗	✓	✗	✓
18	t_b	t_c	t_c	✓	✗	✓	✓	✓	✗	✓	✓	✗
19	t_c	t_a	t_a	✓	✓	✗	✓	✓	✓	✗	✓	✓
20	t_c	t_a	t_b	✓	✓	✓	✗	✓	✓	✓	✗	✓
21	t_c	t_a	t_c	✓	✓	✓	✓	✓	✓	✓	✓	✓
22	t_c	t_b	t_a	✓	✓	✗	✓	✗	✓	✗	✓	✓
23	t_c	t_b	t_b	✓	✓	✓	✓	✓	✓	✓	✓	✓
24	t_c	t_b	t_c	✓	✓	✓	✓	✓	✓	✓	✓	✓
25	t_c	t_c	t_a	✓	✓	✗	✓	✓	✗	✗	✓	✓
26	t_c	t_c	t_b	✓	✓	✓	✓	✓	✗	✓	✗	✓
27	t_c	t_c	t_c	✓	✓	✓	✓	✓	✓	✓	✓	✓

This table presents all possible values that the response variable $\boldsymbol{S}$ can possibly take when instrumental variable Z ranges over $\text{supp}(Z) = \{z_c, z_b, z_{bc}\}$ and treatment status T ranges over $\{t_a, t_b, t_c\}$. The first column enumerates the 27 possible response types. Columns 2 to 4 indicate the response types according to the vector of the values that $[T(z_c)\,,\, T(z_b)\,,\, T(z_{bc})]$ takes. The remaining nine columns indicate whether the response type violates any of the following choice restrictions:

Choice Restriction **1**	$Ch_\omega(z_c) = t_a$	$\Rightarrow$	$Ch_\omega(z_b) \neq t_c$	and	$Ch_\omega(z_c) \neq t_c$
Choice Restriction **2**	$Ch_\omega(z_c) = t_b$	$\Rightarrow$	$Ch_\omega(z_b) = t_b$	and	$Ch_\omega(z_c) = t_b$
Choice Restriction **3**	$Ch_\omega(z_c) = t_c$	$\Rightarrow$	$Ch_\omega(z_c) \neq t_a$		
Choice Restriction **4**	$Ch_\omega(z_b) = t_a$	$\Rightarrow$	$Ch_\omega(z_a) \neq t_b$	and	$Ch_\omega(z_c) \neq t_b$
Choice Restriction **5**	$Ch_\omega(z_b) = t_b$	$\Rightarrow$	$Ch_\omega(z_c) \neq t_a$		
Choice Restriction **6**	$Ch_\omega(z_b) = t_c$	$\Rightarrow$	$Ch_\omega(z_a) = t_c$	and	$Ch_\omega(z_c) = t_c$
Choice Restriction **7**	$Ch_\omega(z_{bc}) = t_a$	$\Rightarrow$	$Ch_\omega(z_a) = t_a$	and	$Ch_\omega(z_b) = t_a$
Choice Restriction **8**	$Ch_\omega(z_{bc}) = t_b$	$\Rightarrow$	$Ch_\omega(z_b) = t_b$		
Choice Restriction **9**	$Ch_\omega(z_{bc}) = t_c$	$\Rightarrow$	$Ch_\omega(z_a) = t_c$		

A check mark sign indicates that the associated response type does not violate the choice restriction. A cross sign indicates that the associated response type violates the choice restriction.

Table D.14: Response Types Generated by WARP for $\text{supp}(Z) = \{z_c, z_b, z_{bc}\}$.

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 $\boldsymbol{S}$						
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

E T-3 Implies Vytlacil's Theorem (2002)

In the binary case where $T \in \{0, 1\}$, T and $\mathbf{1}[T = 1]$ are equivalent and condition (iii) reduces to:

$$\big(T|Z = z, \boldsymbol{V} = \boldsymbol{v}\big) \geq \big(T|Z = z', \boldsymbol{V} = \boldsymbol{v}\big) \text{ or } \big(T|Z = z, \boldsymbol{V} = \boldsymbol{v}\big) \leq \big(T|Z = z', \boldsymbol{V} = \boldsymbol{v}\big), \forall \boldsymbol{v} \in \text{supp}(\boldsymbol{V}). \tag{E.1}$$

Equation (E.1) can be written in terms of an agent ω for whom $\boldsymbol{V}_\omega = \boldsymbol{v} \in \text{supp}(\boldsymbol{V})$ as:

$$T_\omega(z) \geq T_\omega(z') \text{ or } T_\omega(z) \leq T_\omega(z'), \tag{E.2}$$

which is the monotonicity condition of Imbens and Angrist (1994). Condition (iv) reduces to:

$$P\Big(T = \mathbf{1}[\varphi(V) + g(Z) \geq 0]\Big) = 1,$$

which is the separable representation of Vytlacil (2002). Under (E.2), the response matrix of the binary treatment model is lower triangular. This implies that matrices $\boldsymbol{B}_1, \boldsymbol{B}_0$ are maximal matrices and thereby lonesum, which corroborates conditions (i) and (ii) of **T-3**.

F Examples of Unordered Monotonic Response Matrices

Section 5.2 examines the case of multiple treatments, in which the treatment indicator takes three values and the instrumental variable also takes three values. This setup generates 27 possible response types. The number of response matrices generated by the combination of 7 response types taken from these 27 possible ones totals 888,030. Among those, there are 66 response matrices that are unordered monotonic responses. Namely, response matrices whose binary indicator matrices associated with each treatment choice are lonesum. Those are listed in Table F.1.

Table F.1: All Response Matrices that are Unordered Monotonic Responses

t_1	t_2	t_3	t_1	t_1	t_1	t_1	t_1	t_2	t_3	t_1	t_3	t_3	t_3	t_1	t_2	t_3	t_1	t_2	t_2	t_2	t_1	t_2	t_3	t_1	t_2	t_3	t_3	t_1	t_2	t_3	t_1	t_2	t_2	t_2
t_1	t_2	t_3	t_1	t_1	t_2	t_3	t_1	t_2	t_3	t_1	t_1	t_2	t_3	t_1	t_2	t_3	t_1	t_1	t_2	t_3	t_1	t_2	t_3	t_2	t_2	t_2	t_2	t_1	t_2	t_3	t_3	t_2	t_3	t_3
t_1	t_2	t_3	t_2	t_3	t_2	t_3	t_1	t_2	t_3	t_2	t_2	t_2	t_2	t_1	t_2	t_3	t_3	t_3	t_3	t_3	t_1	t_2	t_3	t_1	t_1	t_1	t_3	t_1	t_2	t_3	t_1	t_1	t_1	t_3
t_1	t_2	t_3	t_1	t_1	t_2	t_3	t_1	t_2	t_3	t_1	t_3	t_3	t_3	t_1	t_2	t_3	t_1	t_2	t_3	t_3	t_1	t_2	t_3	t_1	t_2	t_2	t_3	t_1	t_2	t_3	t_1	t_2	t_3	t_3
t_1	t_2	t_3	t_1	t_1	t_1	t_1	t_1	t_2	t_3	t_1	t_1	t_1	t_2	t_1	t_2	t_3	t_1	t_2	t_1	t_2	t_1	t_2	t_3	t_2	t_2	t_2	t_2	t_1	t_2	t_3	t_3	t_3	t_3	t_3
t_1	t_2	t_3	t_2	t_3	t_2	t_3	t_1	t_2	t_3	t_2	t_2	t_3	t_2	t_1	t_2	t_3	t_3	t_3	t_3	t_3	t_1	t_2	t_3	t_1	t_1	t_3	t_3	t_1	t_2	t_3	t_1	t_2	t_1	t_2
t_1	t_2	t_3	t_1	t_1	t_1	t_1	t_1	t_2	t_3	t_1	t_1	t_1	t_1	t_1	t_2	t_3	t_1	t_1	t_1	t_1	t_1	t_2	t_3	t_1	t_3	t_3	t_3	t_1	t_2	t_3	t_1	t_2	t_2	t_3
t_1	t_2	t_3	t_1	t_2	t_3	t_3	t_1	t_2	t_3	t_1	t_2	t_2	t_3	t_1	t_2	t_3	t_2	t_2	t_2	t_3	t_1	t_2	t_3	t_2	t_2	t_2	t_3	t_1	t_2	t_3	t_3	t_3	t_3	t_3
t_1	t_2	t_3	t_2	t_2	t_2	t_3	t_1	t_2	t_3	t_3	t_2	t_3	t_3	t_1	t_2	t_3	t_1	t_2	t_3	t_3	t_1	t_2	t_3	t_1	t_1	t_2	t_1	t_1	t_2	t_3	t_1	t_1	t_2	t_1
t_1	t_2	t_3	t_1	t_1	t_1	t_3	t_1	t_2	t_3	t_1	t_1	t_1	t_3	t_1	t_2	t_3	t_1	t_1	t_1	t_3	t_1	t_2	t_3	t_1	t_3	t_3	t_3	t_1	t_2	t_3	t_1	t_2	t_2	t_2
t_1	t_2	t_3	t_1	t_2	t_3	t_3	t_1	t_2	t_3	t_1	t_2	t_2	t_2	t_1	t_2	t_3	t_2	t_2	t_2	t_2	t_1	t_2	t_3	t_2	t_2	t_2	t_2	t_1	t_2	t_3	t_3	t_3	t_3	t_3
t_1	t_2	t_3	t_2	t_2	t_2	t_2	t_1	t_2	t_3	t_3	t_2	t_3	t_3	t_1	t_2	t_3	t_1	t_2	t_3	t_3	t_1	t_2	t_3	t_1	t_1	t_2	t_3	t_1	t_2	t_3	t_1	t_1	t_2	t_3
t_1	t_2	t_3	t_1	t_1	t_3	t_3	t_1	t_2	t_3	t_1	t_1	t_1	t_2	t_1	t_2	t_3	t_1	t_1	t_1	t_1	t_1	t_2	t_3	t_1	t_1	t_1	t_1	t_1	t_2	t_3	t_2	t_2	t_3	t_3
t_1	t_2	t_3	t_1	t_2	t_2	t_3	t_1	t_2	t_3	t_1	t_2	t_3	t_2	t_1	t_2	t_3	t_2	t_2	t_3	t_3	t_1	t_2	t_3	t_2	t_3	t_3	t_3	t_1	t_2	t_3	t_1	t_1	t_1	t_1
t_1	t_2	t_3	t_2	t_2	t_2	t_2	t_1	t_2	t_3	t_3	t_3	t_3	t_3	t_1	t_2	t_3	t_1	t_2	t_1	t_3	t_1	t_2	t_3	t_2	t_1	t_2	t_3	t_1	t_2	t_3	t_1	t_2	t_1	t_3
t_1	t_2	t_3	t_1	t_1	t_1	t_2	t_1	t_2	t_3	t_1	t_1	t_2	t_3	t_1	t_2	t_3	t_1	t_1	t_3	t_3	t_1	t_2	t_3	t_1	t_1	t_1	t_3	t_1	t_2	t_3	t_2	t_2	t_2	t_2
t_1	t_2	t_3	t_1	t_3	t_3	t_3	t_1	t_2	t_3	t_1	t_2	t_2	t_2	t_1	t_2	t_3	t_2	t_2	t_2	t_2	t_1	t_2	t_3	t_2	t_3	t_3	t_3	t_1	t_2	t_3	t_1	t_1	t_3	t_3
t_1	t_2	t_3	t_2	t_2	t_3	t_2	t_1	t_2	t_3	t_3	t_3	t_3	t_3	t_1	t_2	t_3	t_1	t_2	t_2	t_3	t_1	t_2	t_3	t_2	t_1	t_2	t_2	t_1	t_2	t_3	t_1	t_2	t_2	t_3
t_1	t_2	t_3	t_1	t_1	t_2	t_3	t_1	t_2	t_3	t_1	t_1	t_2	t_2	t_1	t_2	t_3	t_1	t_1	t_1	t_2	t_1	t_2	t_3	t_1	t_1	t_1	t_2	t_1	t_2	t_3	t_2	t_2	t_2	t_3
t_1	t_2	t_3	t_1	t_3	t_3	t_3	t_1	t_2	t_3	t_1	t_3	t_2	t_3	t_1	t_2	t_3	t_2	t_2	t_3	t_2	t_1	t_2	t_3	t_3	t_3	t_3	t_3	t_1	t_2	t_3	t_1	t_1	t_1	t_1
t_1	t_2	t_3	t_2	t_2	t_2	t_2	t_1	t_2	t_3	t_3	t_3	t_3	t_3	t_1	t_2	t_3	t_1	t_3	t_3	t_3	t_1	t_2	t_3	t_1	t_2	t_3	t_2	t_1	t_2	t_3	t_1	t_2	t_3	t_3
t_1	t_2	t_3	t_1	t_2	t_3	t_3	t_1	t_2	t_3	t_1	t_2	t_2	t_3	t_1	t_2	t_3	t_1	t_1	t_2	t_3	t_1	t_2	t_3	t_1	t_1	t_2	t_3	t_1	t_2	t_3	t_2	t_2	t_2	t_2
t_1	t_2	t_3	t_1	t_1	t_1	t_3	t_1	t_2	t_3	t_1	t_1	t_1	t_1	t_1	t_2	t_3	t_2	t_2	t_2	t_2	t_1	t_2	t_3	t_3	t_3	t_3	t_3	t_1	t_2	t_3	t_1	t_1	t_1	t_3
t_1	t_2	t_3	t_2	t_2	t_2	t_2	t_1	t_2	t_3	t_3	t_2	t_3	t_3	t_1	t_2	t_3	t_1	t_3	t_3	t_3	t_1	t_2	t_3	t_1	t_2	t_2	t_2	t_1	t_2	t_3	t_1	t_2	t_3	t_3
t_1	t_2	t_3	t_1	t_2	t_3	t_3	t_1	t_2	t_3	t_1	t_2	t_2	t_2	t_1	t_2	t_3	t_1	t_1	t_2	t_3	t_1	t_2	t_3	t_1	t_1	t_2	t_2	t_1	t_2	t_3	t_2	t_2	t_2	t_3
t_1	t_2	t_3	t_1	t_1	t_1	t_1	t_1	t_2	t_3	t_1	t_1	t_1	t_3	t_1	t_2	t_3	t_2	t_3	t_2	t_3	t_1	t_2	t_3	t_3	t_3	t_3	t_3	t_1	t_2	t_3	t_1	t_1	t_2	t_1
t_1	t_2	t_3	t_2	t_2	t_2	t_3	t_1	t_2	t_3	t_3	t_2	t_3	t_3	t_1	t_2	t_3	t_1	t_1	t_1	t_1	t_1	t_2	t_3	t_1	t_3	t_2	t_3	t_1	t_2	t_3	t_1	t_3	t_3	t_3
t_1	t_2	t_3	t_1	t_2	t_2	t_3	t_1	t_2	t_3	t_1	t_2	t_2	t_3	t_1	t_2	t_3	t_1	t_2	t_3	t_3	t_1	t_2	t_3	t_1	t_2	t_2	t_3	t_1	t_2	t_3	t_2	t_2	t_2	t_2
t_1	t_2	t_3	t_1	t_1	t_3	t_3	t_1	t_2	t_3	t_1	t_1	t_2	t_1	t_1	t_2	t_3	t_2	t_2	t_2	t_3	t_1	t_2	t_3	t_3	t_2	t_3	t_3	t_1	t_2	t_3	t_1	t_1	t_2	t_3
t_1	t_2	t_3	t_2	t_2	t_2	t_2	t_1	t_2	t_3	t_3	t_3	t_3	t_3	t_1	t_2	t_3	t_1	t_1	t_1	t_1	t_1	t_2	t_3	t_1	t_1	t_1	t_1	t_1	t_2	t_3	t_1	t_3	t_3	t_3
t_1	t_2	t_3	t_3	t_3	t_3	t_3	t_1	t_2	t_3	t_3	t_3	t_3	t_3	t_1	t_2	t_3	t_3	t_3	t_3	t_3	t_1	t_2	t_3	t_3	t_3	t_3	t_3	t_1	t_2	t_3	t_3	t_3	t_3	t_3
t_1	t_2	t_3	t_1	t_2	t_3	t_3	t_1	t_2	t_3	t_1	t_2	t_2	t_3	t_1	t_2	t_3	t_1	t_2	t_2	t_2	t_1	t_2	t_3	t_1	t_1	t_2	t_3	t_1	t_2	t_3	t_1	t_1	t_2	t_3
t_1	t_2	t_3	t_1	t_2	t_1	t_2	t_1	t_2	t_3	t_1	t_1	t_2	t_1	t_1	t_2	t_3	t_1	t_1	t_2	t_3	t_1	t_2	t_3	t_1	t_2	t_2	t_2	t_1	t_2	t_3	t_1	t_2	t_2	t_2

This table presents all response matrices consisting of seven response types that comply with the separability condition of Theorem **T-3**.

G Why Do We Get Separability?

This appendix motivates why separability condition (iv) of **T-3** holds. We now present a detailed discussion of how the lonesum property of matrix $\boldsymbol{B}_t$ generates separability of the choice equation in item (iv) of **T-3**. Our proof is in three steps. We first show that if all $\boldsymbol{B}_t; t \in \text{supp}(T)$ are lonesum, then $\boldsymbol{B}_t[i,j]$ can be expressed in terms of its column and row sums by:[7]

$$\boldsymbol{B}_t[i,j] = \mathbf{1}\Bigg[\underbrace{\left(\sum_{j'=1}^{N_{\boldsymbol{S}}} \mathbf{1}\Bigg[\sum_{i'=1}^{N_Z} \boldsymbol{B}_t[i',j] \leq \sum_{i'=1}^{N_Z} \boldsymbol{B}_t[i',j']\Bigg]\right)}_{\text{Number of columns whose sum isbigger than or equal to the column sum of } \boldsymbol{B}_t[\cdot,j]} \leq \underbrace{\left(\sum_{j'=1}^{N_{\boldsymbol{S}}} \boldsymbol{B}_t[i,j']\right)}_{\text{row sum of } \boldsymbol{B}_t[i,\cdot]}\Bigg], \tag{G.1}$$

Next, we use (G.1) to show that $\boldsymbol{B}_t[i,j]$ can be also expressed in terms of propensity scores and response-type probabilities as described in Equation (G.2).[8]

$$\boldsymbol{B}_t[i,j] = \mathbf{1}\Bigg[\Bigg(\sum_{j'=1}^{N_{\boldsymbol{S}}} P(\boldsymbol{S}=\boldsymbol{s}_{j'}) \cdot \mathbf{1}\Bigg[\sum_{i'=1}^{N_Z} \mathrm{P}(Z=z_{i'})\boldsymbol{B}_t[i',j] \leq \sum_{i'=1}^{N_Z} \mathrm{P}(Z=z_{i'})\boldsymbol{B}_t[i',j']\Bigg]\Bigg) \leq \Bigg(\sum_{j'=1}^{N_{\boldsymbol{S}}} P(\boldsymbol{S}=\boldsymbol{s}_{j'}) \cdot \boldsymbol{B}_t[i,j']\Bigg)\Bigg]. \tag{G.2}$$

We then use Equations (15) and (17) in the text to replace the terms $\boldsymbol{B}_t$ in Equation (G.2) with propensity scores and the probability of $T=t$ conditional on $\boldsymbol{S}$, which is a function of the values that T and $\boldsymbol{V}$ take.[9]

A consequence of (i) in **T-3** is that each binary matrix $\boldsymbol{B}_t$ is equivalent to its maximal. This property generates a key ingredient of the proof. $\boldsymbol{B}_t[i,j]$ can be expressed in terms of its column and row sums using the following relationship:

$$\boldsymbol{B}_t[i,j] = \mathbf{1}\Bigg[\underbrace{\left(\sum_{j'=1}^{N_{\boldsymbol{S}}} \mathbf{1}\Bigg[\sum_{i'=1}^{N_Z} \boldsymbol{B}_t[i',j] \leq \sum_{i'=1}^{N_Z} \boldsymbol{B}_t[i',j']\Bigg]\right)}_{\text{Number of columns whose sum is bigger than or equal to the column sum of } \boldsymbol{B}_t[\cdot,j]} \leq \underbrace{\left(\sum_{j'=1}^{N_{\boldsymbol{S}}} \boldsymbol{B}_t[i,j']\right)}_{\text{row sum of } \boldsymbol{B}_t[i,\cdot]}\Bigg], \tag{G.3}$$

[7] See Lemma **L-12** of Appendix A for a formal proof.

[8] See Lemmas **L-5**–**L-8** and Lemma **L-14** of Appendix A for a formal proof.

[9] For instance, we show that $\mathbf{1}[T=t|Z=z_i,\boldsymbol{S}=\boldsymbol{s}_j]$ can be expressed as $\mathbf{1}[T=t|Z=z_i,\boldsymbol{S}=\boldsymbol{s}_j] = \mathbf{1}\big[\varphi(\boldsymbol{s}_j,t)+g(z_i,t)\geq 0\big]$ where $g(z_i,t)=\mathrm{P}(T=t|Z=z_i)$ and $\varphi(\boldsymbol{s}_j,t) = -\sum_{j'=1}^{N_{\boldsymbol{S}}} \mathrm{P}(\boldsymbol{S}=\boldsymbol{s}_{j'})\cdot\mathbf{1}\big[\mathrm{P}(T=t|\boldsymbol{S}=\boldsymbol{s}_j)\leq \mathrm{P}(T=t|\boldsymbol{S}=\boldsymbol{s}_{j'})\big]$.

where $i \in \{1, \ldots, N_Z\}$ and $j \in \{1, \ldots, N_{\boldsymbol{S}}\}$. Equation (G.3) states that we can determine $\boldsymbol{B}_t[i,j]$ by comparing the i^{th} *row-sum* with the *number of columns* (including the j^{th} column) whose column sum is greater than or equal to the j^{th} column sum. If the row sum is equal to or greater than this number of columns then $\boldsymbol{B}_t[i,j] = 1$, otherwise, $\boldsymbol{B}_t[i,j] = 0$.

An example clarifies Equation (G.3). Consider the binary matrix $\boldsymbol{B}$ in Table G.1. The table also shows the matrix generated by reordering the columns of matrix $\boldsymbol{B}$ in decreasing column sums and its rows in increasing row sums. The rows of the reordered matrix consist of a sequence of elements 1 followed by a sequence of elements 0. Thus matrix $\boldsymbol{B}$ is equivalent to its maximal.

Consider the last column of the reordered matrix. It consists of elements $[0,0,0]'$ whose column sum is 0. All 5 columns have column sums greater than or equal to 0. The row sums (1,3 and 4) are all less than 5, which generates the elements 0. The second column consists of elements $[0,1,1]$, whose column sum is 2. There are 3 columns whose column sum is equal or greater than 2 (first, second and third columns). The sum of the first row is 1, which is less than 3, and this generates the element 0. The second and third row sums are 3 and 4, both greater than or equal to 3, thereby generating the elements 1.

Equation (G.3) also holds for the original matrix. The first column of the original matrix consists of elements $[1,0,0]'$, whose column sum is 1. There are 4 columns whose column sums are greater than or equal to 1 (first, second, fourth and fifth columns). The sum of the first row is 4, which is greater or equal than 4, generating the element 1. The sums of the second and third rows are 1 and 3, both less than 4, generating the elements 0.

Table G.1: Example of a Binary Matrix $\boldsymbol{B}$ that is Equivalent to its Maximal

Row	Row Sum	Binary Matrix $\boldsymbol{B}$					Row	Row Sum	Reordered Matrix $\boldsymbol{B}$				
$\boldsymbol{B}[1,\cdot]$	4	1	1	0	1	1	$\boldsymbol{B}[2,\cdot]$	1	1	0	0	0	0
$\boldsymbol{B}[2,\cdot]$	1	0	0	0	1	0	$\boldsymbol{B}[3,\cdot]$	3	1	1	1	0	0
$\boldsymbol{B}[3,\cdot]$	3	0	1	0	1	1	$\boldsymbol{B}[1,\cdot]$	4	1	1	1	1	0
		$\boldsymbol{B}[\cdot,1]$	$\boldsymbol{B}[\cdot,2]$	$\boldsymbol{B}[\cdot,3]$	$\boldsymbol{B}[\cdot,4]$	$\boldsymbol{B}[\cdot,5]$			$\boldsymbol{B}[\cdot,4]$	$\boldsymbol{B}[\cdot,2]$	$\boldsymbol{B}[\cdot,5]$	$\boldsymbol{B}[\cdot,1]$	$\boldsymbol{B}[\cdot,3]$
Column Sum		1	2	0	3	2			3	2	2	1	0

The lonesum property of $\boldsymbol{B}_t$ also generates the following equality:[10]

$$\mathbf{1}\left[\left(\sum_{j'=1}^{N_{\boldsymbol{S}}}\mathbf{1}\left[\sum_{i'=1}^{N_Z}\boldsymbol{B}_t[i',j]\leq\sum_{i'=1}^{N_Z}\boldsymbol{B}_t[i',j']\right]\right)\leq\left(\sum_{j'=1}^{N_{\boldsymbol{S}}}\right)\boldsymbol{B}_t[i,j']\right]=$$
$$=\mathbf{1}\left[\left(\sum_{j'=1}^{N_{\boldsymbol{S}}}P(\boldsymbol{S}=\boldsymbol{s}_{j'})\cdot\mathbf{1}\left[\sum_{i'=1}^{N_Z}\boldsymbol{B}_t[i',j]\leq\sum_{i'=1}^{N_Z}\boldsymbol{B}_t[i',j']\right]\right)\leq\left(\sum_{j'=1}^{N_{\boldsymbol{S}}}P(\boldsymbol{S}=\boldsymbol{s}_{j'})\cdot\boldsymbol{B}_t[i,j']\right)\mathbf{1}\right]. \qquad \text{(G.4)}$$

According to Remark 6.3, no 2×2 sub-matrix of each $\boldsymbol{B}_t$ takes the prohibited patterns (52).[11] As a consequence, we have that:[12]

$$\sum_{i'=1}^{N_Z}\boldsymbol{B}_t[i',j]\leq\sum_{i'=1}^{N_Z}\boldsymbol{B}_t[i',j']\Leftrightarrow\boldsymbol{B}_t[i',j]\leq\boldsymbol{B}_t[i',j']\quad\forall\ i'\in\{1,\dots,N_Z\},$$

thereby $\mathbf{1}\left[\sum_{i'=1}^{N_Z}\boldsymbol{B}_t[i',j]\leq\sum_{i'=1}^{N_Z}\boldsymbol{B}_t[i',j']\right]=\mathbf{1}\left[\sum_{i'=1}^{N_Z}\mathrm{P}(Z=z_{i'})\boldsymbol{B}_t[i',j]\leq\sum_{i'=1}^{N_Z}\mathrm{P}(Z=z_{i'})\boldsymbol{B}_t[i',j']\right]. \qquad \text{(G.5)}$

If we substitute Equations (G.4)–(G.5) into (G.3) we obtain:

$$\boldsymbol{B}_t[i,j]=\mathbf{1}\left[\left(\sum_{j'=1}^{N_{\boldsymbol{S}}}P(\boldsymbol{S}=\boldsymbol{s}_{j'})\cdot\mathbf{1}\left[\sum_{i'=1}^{N_Z}\mathrm{P}(Z=z_{i'})\boldsymbol{B}_t[i',j]\leq\sum_{i'=1}^{N_Z}\mathrm{P}(Z=z_{i'})\boldsymbol{B}_t[i',j']\right]\right)\leq\left(\sum_{j'=1}^{N_{\boldsymbol{S}}}P(\boldsymbol{S}=\boldsymbol{s}_{j'})\cdot\boldsymbol{B}_t[i,j']\right)\right]. \qquad \text{(G.6)}$$

[10] See Lemma **L-13** of Appendix A for a formal proof.
[11] See Lemmas **L-5**–**L-8** of Appendix A for a formal proof.
[12] See Lemma **L-14** of Appendix A for a formal proof.

Next we can represent Equations (15)–(17) in terms of $\boldsymbol{B}_t$ as:

$$\mathrm{P}(T=t|Z=z_i) = \sum_{j=1}^{N_{\boldsymbol{S}}} \boldsymbol{B}_t[i,j]\,\mathrm{P}(\boldsymbol{S}=\boldsymbol{s}_j), \tag{G.7}$$

$$\mathrm{P}(T=t|\boldsymbol{S}=\boldsymbol{s}_j) = \sum_{i=1}^{N_Z} \boldsymbol{B}_t[i,j]\,\mathrm{P}(Z=z_i), \tag{G.8}$$

where $\boldsymbol{S}$ is a balancing score for V.

Equations (G.7)–(G.8) are useful for translating the summations over $\boldsymbol{B}_t$ into propensity scores of t, conditional on Z and $\boldsymbol{V}$. We substitute Equations (G.7)–(G.8) into (G.6). As we show here and in the proof of Theorem **T-3**, we can construct $g(z_i,t)$ and $\varphi(\boldsymbol{s}_j,t)$ from the following relationships:

$$\mathbf{1}[T=t|Z=z_i,\boldsymbol{S}=\boldsymbol{s}_j] = \mathbf{1}\big[\varphi(\boldsymbol{s}_j,t)+g(z_i,t)\geq 0\big],\ \text{where } g(z_i,t)=\mathrm{P}(T=t|Z=z_i) \text{ and}$$

$$\varphi(\boldsymbol{s}_j,t) = -\sum_{j'=1}^{N_{\boldsymbol{S}}} \mathrm{P}(\boldsymbol{S}=\boldsymbol{s}_{j'})\cdot \mathbf{1}\big[\,\mathrm{P}(T=t|\boldsymbol{S}=\boldsymbol{s}_j)\leq \mathrm{P}(T=t|\boldsymbol{S}=\boldsymbol{s}_{j'})\big].$$

H Examining the Threshold Property of Condition (iv) of Theorem T-3

Condition (iv) of **T-3** states that the treatment choice T can be expressed as:

$$\mathbf{1}[T=t|\boldsymbol{V}=\boldsymbol{v},Z=z] = \mathbf{1}[\varphi(\boldsymbol{V},t)+g(Z,t)\geq 0]. \tag{H.1}$$

This representation can be understood as the combination of a *separability condition* and a *threshold property*:

1. *Separability Condition* refers to the separable equation $\varphi(\boldsymbol{v},t)+g(z,t)$ of the Equality (H.1).

2. *Threshold Property* refers to the fact that T takes value 1 if $\varphi(\boldsymbol{v},t)+g(z,t)$ in (H.1) is

greater or equal than the threshold value 0.

To clarify the role of the Threshold Property, consider the case in which the separability condition holds but the threshold property does not. In this case, Equality (H.1) would be replaced by Equality (H.2) below.

$$\mathbf{1}[T = t|\boldsymbol{V} = \boldsymbol{v}, Z = z] = \mathbf{1}[\varphi(\boldsymbol{V}, t) \geq -g(Z, t)]. \tag{H.2}$$

This section clarifies why the threshold property is necessary in Equality (H.1) of condition (iv). Namely, we show that separability alone is not enough to produce unordered monotonicity.

Let T take value $t^* \in \text{supp}(T)$ when $\boldsymbol{V} = \boldsymbol{v}$ and $Z = z$. Then the separability condition **and** the threshold property discussed above imply that:

$$\text{if } t^* = \text{argmax}_{t \in \text{supp}(T)} \varphi(\boldsymbol{v}, t) + g(z, t),$$
$$\text{then } \varphi(\boldsymbol{v}, t^*) + g(z, t^*) \geq 0 \text{ and } \varphi(\boldsymbol{v}, t') + g(z, t') < 0 \text{ for all } t' \neq t^*.$$

We use the following strategy to prove that separability alone is not sufficient to produce unordered monotonicity.

1. We present an example in which the treatment choice T is expressed by an equation where separability holds but the threshold property does not.

2. We show that this example can generate a response matrix that contains the prohibited pattern (i.e., Equation (52)).

3. We then evoke Condition (iii) of **T-3** which states that the prohibited pattern implies that unordered monotonicity does not hold.

Let $\Psi(t, z, \boldsymbol{v})$ represent the utility of choice $T = t$ when $\boldsymbol{V} = \boldsymbol{v}$ and $Z = z$. We assume that $\Psi(t, z, \boldsymbol{v})$ is separable, that is, $\Psi(t, z, \boldsymbol{v}) = u(\boldsymbol{v}, t) + h(z, t)$ (but we do not invoke the threshold property). It suffices to show that $\Psi(t, z, \boldsymbol{v})$ can generate the prohibited pattern defined by Equation (52). Let $\text{supp}(T) = \{t_1, t_2, t_3\}$, $\text{supp}(Z) = \{z, z'\}$, $\text{supp}(\boldsymbol{S}) = \{\boldsymbol{s}, \boldsymbol{s}'\}$.

The associated response matrix is given by:

$$\boldsymbol{R} = \begin{pmatrix} (T|\boldsymbol{S}=\boldsymbol{s}, Z=z) & (T|\boldsymbol{S}=\boldsymbol{s}', Z=z) \\ (T|\boldsymbol{S}=\boldsymbol{s}, Z=z') & (T|\boldsymbol{S}=\boldsymbol{s}', Z=z') \end{pmatrix}$$

$$= \begin{pmatrix} \underset{t\in\{t_1,t_2,t_3\}}{\operatorname{argmax}} \big(u(\boldsymbol{s},t)+h(z,t)\big) & \underset{t\in\{t_1,t_2,t_3\}}{\operatorname{argmax}} \big(u(\boldsymbol{s}',t)+h(z,t)\big) \\ \underset{t\in\{t_1,t_2,t_3\}}{\operatorname{argmax}} \big(u(\boldsymbol{s},t)+h(z',t)\big) & \underset{t\in\{t_1,t_2,t_3\}}{\operatorname{argmax}} \big((u(\boldsymbol{s}',t)+h(z',t)\big) \end{pmatrix} \tag{H.3}$$

The prohibited pattern arises in the response matrix $\boldsymbol{R}$ above has t_1 on the diagonal but has choice values other than t_1 on the off-diagonal. For this to happen we need the following inequalities to hold:

$$u(\boldsymbol{s},t_1)+h(z,t_1) > \max(u(\boldsymbol{s},t_2)+h(z,t_2), u(\boldsymbol{s},t_3)+h(z,t_3)),$$

$$u(\boldsymbol{s}',t_1)+h(z',t_1) > \max(u(\boldsymbol{s}',t_2)+h(z',t_2), u(\boldsymbol{s}',t_3)+h(z',t_3)),$$

$$u(\boldsymbol{s},t_1)+h(z',t_1) < \max(u(\boldsymbol{s},t_2)+h(z',t_2), u(\boldsymbol{s},t_3)+h(z',t_3)),$$

$$u(\boldsymbol{s}',t_1)+h(z,t_1) < \max(u(\boldsymbol{s}',t_2)+h(z,t_2), u(\boldsymbol{s}',t_3)+h(z,t_3)).$$

Now consider the following values for $u(\boldsymbol{s},t), h(z,t)$:

$$u(\boldsymbol{s},t_1) = h(z,t_1) = u(\boldsymbol{s}',t_1) = h(z',t_1) = 0,$$

$$u(\boldsymbol{s},t_2) = h(z,t_3) = u(\boldsymbol{s}',t_3) = h(z',t_2) = 1,$$

$$u(\boldsymbol{s},t_3) = h(z,t_2) = u(\boldsymbol{s}',t_2) = h(z',t_3) = -2.$$

The threshold property is violated as

$$u(\boldsymbol{s},t_1)+h(z,t_1) = 0 = \max_{t\in\{t_1,t_2,t_3\}} u(\boldsymbol{s},t)+h(z,t)$$

but

$$u(\boldsymbol{s},t_2)+h(z',t_2) = 2 = \max_{t\in\{t_1,t_2,t_3\}} u(\boldsymbol{s},t)+h(z',t)$$

while

$$u(\boldsymbol{s},t_1)+h(z',t_1) = 0.$$

Thus, the separability condition of $\Psi(t, z, \boldsymbol{v})$ is not sufficient to guarantee that unordered monotonicity **A-3** holds.

References

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.

DAWID, A. P. (1979): "Conditional independence in statistical theory (with Discussion)," *Journal of the Royal Statistical Society. Series B (Statistical Methodological)*, 41, 1–31.

HECKMAN, J. J. AND R. PINTO (2015): "Alternative Ways to Identify P_S," Unpublished manuscript, University of Chicago, August 2015.

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

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

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.

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