NBER WORKING PAPER SERIES

TREATMENT EFFECTS WITH MULTIPLE OUTCOMES

John Mullahy

Working Paper 25307
http://www.nber.org/papers/w25307

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2018

I am grateful to Domenico Depalo, Paddy Gillespie, Hyunseung Kang, Chuck Manski, and participants in presentations at Queen's University Belfast, the University of Bergamo, the University of Chicago, and the University of Wisconsin-Madison for helpful comments, suggestions, and guidance on the literature. Partial support was provided by the Robert Wood Johnson Foundation (Evidence for Action Grant 73336), and by the UW-Madison Center for Demography and Ecology (NICHD grant P2CHD047873). The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

Treatment Effects with Multiple Outcomes
John Mullahy
NBER Working Paper No. 25307
November 2018
JEL No. C18,D04,I1

ABSTRACT

This paper proposes strategies for defining, identifying, and estimating features of treatment-effect distributions in contexts where multiple outcomes are of interest. After describing existing empirical approaches used in such settings, the paper develops a notion of treatment preference that is shown to be a feature of standard treatment-effect analysis in the single-outcome case. Focusing largely on binary outcomes, treatment-preference probability treatment effects (PTEs) are defined and are seen to correspond to familiar average treatment effects in the single-outcome case. The paper suggests seven possible characterizations of treatment preference appropriate to multiple-outcome contexts. Under standard assumptions about unconfoundedness of treatment assignment, the PTEs are shown to be point identified for three of the seven characterizations and set identified for the other four. Probability bounds are derived and empirical approaches to estimating the bounds—or the PTEs themselves in the point-identified cases—are suggested. These empirical approaches are straightforward, involving in most instances little more than estimation of binary-outcome probability models of what are commonly known as composite outcomes. The results are illustrated with simulated data and in analyses of two microdata samples. Finally, the main results are extended to situations where the component outcomes are ordered or categorical.

John Mullahy
University of Wisconsin-Madison
Dept. of Population Health Sciences
787 WARF, 610 N. Walnut Street
Madison, WI 53726
and NBER
jmullahy@facstaff.wisc.edu

1. Introduction

Obtaining a clear picture of the effect of a treatment or intervention on a single outcome of interest can be daunting. Familiar challenges include treatment-effect heterogeneity, confounded or endogenous treatment assignment, and generalization from experimental to population circumstances. Many now-familiar strategies to mitigate the estimation biases that can arise from such challenges have been developed, and still more are evolving.[1]

The challenges proliferate when multiple outcomes are of interest. Even if the obstacles noted above are absent, the simple notion of what is meant by "a treatment effect" is no longer obvious when two or more outcomes are of concern. This paper's main goals are to develop an integrated framework for understanding treatment effects with multiple outcomes, to determine how features of the population distribution of such treatment effects might be identified, and to suggest empirical approaches to learning from data about these features.

Multiple Outcomes

The existing literature on treatment effects (TEs) offers only limited guidance for understanding multiple outcomes. Abadie and Cattaneo, 2018, note that "in practice, researchers may be interested in a multiplicity of treatments and outcomes," but then conduct their analysis treating both treatment (their W) and outcome (their Y) as scalar random variables. Athey and Imbens, 2017, recognize explicitly multiple-outcome contexts, but do so largely with regard to multiple-testing problems.[2] Manski and Tetenov, 2018, consider optimizing clinical trial sizes when multiple outcomes are of interest. Athey et al., 2016, assess how multiple surrogate outcomes can inform understanding of treatment effects.

Such limited scope is unfortunate as considerations of multiple outcomes arise broadly in empirical work. In health and clinical research multiple outcomes are often commonplace when measuring some population's health status or health behaviors, or when attempting to understand their determinants. Buttorff et al., 2017, consider how chronic conditions vary across the U.S. adult population. Hoynes et al., 2015, explore how public policies affect multiple infant-health outcomes. Ludwig et al., 2011 and 2013, use data from the Moving to Opportunity experiment to assess how residential changes affect physical and mental health outcomes. Pesko et al., 2016, consider how regulations may affect use of cigarettes, e-cigarettes, and nicotine replacement therapy.

---

[1] See Imbens and Wooldridge, 2009, for a valuable review.

[2] While some research has considered TEs in multiple-*treatment* contexts (e.g. Angrist and Imbens, 1995), this is unrelated to the multiple-outcome contexts considered here. That said, Appendix B considers briefly extensions of this paper's results to contexts involving more than two treatments.

Multiple outcomes are also frequently involved in studies of clinical populations.[3] Prominent in cardiovascular and diabetes research, for instance, are clinical trials that focus on multiple outcomes such as mortality, stroke, myocardial infarction, and hospital readmission, often aggregated into composite outcomes (e.g. Look AHEAD Research Group, 2013; Parving et al., 2012; Rosenfield et al., 2016).[4] Beyond serving as primary or secondary outcomes in such clinical research, multiple outcomes are often studied in the form of treatment-specific adverse events. For example, Parving et al., 2012, report the rates of 21 adverse events potentially affecting subjects in the two arms of their study of diabetes treatments. In some instances a study may examine a single primary outcome and a single adverse event or safety outcome, as such examining data having a multiple-outcome character. In their study of transcatheter mitral-valve repair, for instance, Stone et al., 2018, focus on hospitalization for heart failure within 24 months of follow-up as their primary outcome and device-related complications at 12 months as their main safety outcome.

Healthcare quality measurement is another research and policy area where consideration of multiple outcomes is of central interest. For example Cebul et al., 2011, consider how clinical practices' electronic health record use is related to four measures of care-process quality and five measures of patient outcomes (see also IOM, 2006, and Shwartz et al., 2015).

Beyond health and health care, considerations of multiple outcomes arise in contexts as diverse as welfare and poverty (e.g. multiple deprivations, Nolan and Whelan, 2010), education (e.g. school accountability, Loeb and Figlio, 2011; teacher quality, Jackson, 2018), nutrition (e.g. food security, Coleman-Jensen et al., 2018), finance (e.g. financial-institution soundness, FDIC, 1997), and many others.

For the most part the wide array of data structures mentioned in the preceding paragraphs falls within this paper's scope. While the strategies proposed here will not necessarily be suited for understanding treatment effects related to all multiple-outcome contexts they will likely be applicable to many, even if the study of such outcomes has not traditionally been approached in the manner suggested below. While in many instances the focus of a multiple-outcomes analysis will be on subjects whose relevant outcomes at a point in time are summarized by an M-component vector, the analytics described here accommodate other multiple-outcome contexts. For example, the M outcomes may also be a univariate outcome characterizing each subject over M time periods,[5] or an M'-dimension outcome for each subject over T time periods with $\mathrm{T} \times \mathrm{M}' = \mathrm{M}$, i.e.

[3] See Manski, 2018, for an assessment of medical decisionmaking under uncertainty.

[4] Composites are considered in section 5 and in an in-progress companion paper (Mullahy, 2018b).

[5] For example, in their study of treatments for episodic migraine Stauffer et al., 2018, use a binary "migraine headache day" outcome, a univariate outcome measured daily over the follow-up period.

multiple-outcome panel data (see also footnote 26). The key in any case is that outcomes can be imagined arising under alternative treatments, interventions, or policies.

Understanding Treatment Effects with Multiple Outcomes—Existing Approaches

Given such outcome data some questions naturally arise, perhaps most prominent being: How does one conceptualize a TE, or perhaps a set of TEs, when studying multiple outcomes?

To provide context for the paper's main analysis, suppose one observes M binary outcomes $\mathbf{y}=\left[y_1,\ldots,y_M\right]$ and a vector of covariates $\mathbf{x}=\left[x_T,\mathbf{x}_{oth}\right]$, where x_T measures exogenous scalar treatment and $\mathbf{x}_{oth}$ are other exogenous covariates.[6] With such data at least three analytical strategies for estimation of average treatment effects (ATEs) have been prominent in practice.

In the first,[7] separate ATEs are estimated for each of the M components y_m of $\mathbf{y}$ based on estimates of some parametric or nonparametric probability model

$$\Pr\left(y_m=1\middle|\mathbf{x}\right)=p_m\left(\mathbf{x}\right),\quad m=1,\ldots,M. \tag{1}$$

Strategies like this result in estimates of M separate ATEs.[8]

A second approach[9] treats the sum $s_{\mathbf{y}}=\sum_{m=1}^{M}w_m y_m$ as a continuous, ordered, or count outcome, and considers parametric or nonparametric models for its conditional mean

$$E\left[s_{\mathbf{y}}\middle|\mathbf{x}\right]=\mu_s\left(\mathbf{x}\right) \tag{2}$$

and/or conditional probability structure

---

[6] The notation will be formalized in sections 2 and 3.

[7] See, e.g., Sarma et al., 2015. Only sometimes are multiple testing issues addressed; see Romano et al., 2010, and Dmitrienko and D'Agostino, 2018.

[8] That the effects may be heterogeneous due to $\mathbf{x}_{oth}$ and to unobservables is in general an important consideration but one that will be ignored in this paper.

[9] See, e.g., Dodge et al., 2014, Hanssen et al., 2014, Jackson, 2018, Khan et al., 2008, Siebert et al., 2016, and Walitt et al., 2016, for a sampling of approaches that have been used. These include linear and ordered-outcome regression, count-data models (e.g. Poisson, negative binomial), and others. Note that the weighted-sum construct includes approaches like principal components.

$$\Pr\left(s_{\mathbf{y}} = n \middle| \mathbf{x}\right) = p_s\left(n, \mathbf{x}\right). \tag{3}$$

This approach might yield a single ATE estimate or a set of ATE estimates across the values of n.

A third approach[10]—one that will be seen in section 5 to be of particular interest in this paper—is in essence a coarsening of $s_{\mathbf{y}}$ into a binary outcome, i.e. $1\left(s_{\mathbf{y}} \geq t\right)$, where t is some relevant threshold or cut point. Binary measures like this are encountered often in health and clinical research as one form of so-called composite outcomes. Two important cases are t=1 ("any") and t=M ("all").[11] To understand such outcomes analysts typically specify and estimate some parametric or nonparametric conditional probability model

$$\Pr\left(s_{\mathbf{y}} \geq t \middle| \mathbf{x}\right) = p_c\left(\mathbf{x}\right) \tag{4}$$

on which estimates of ATEs are based.[12,13]

To anticipate some of what follows, it is useful to consider specifically how one would use ATE estimates like those described in the preceding paragraphs to arrive at a conclusion that one treatment (x_T) is "better than" or "preferable to" a comparator treatment ($x_{T'}$). When an approach like (1) is used, the analyst must consider how to draw a conclusion or inference from M separate estimates that have been obtained. When an approach like (2) is used, the analyst may determine that any particular conclusion about treatment superiority could depend on the manner in which the outcomes are weighted (or on the fact that they are implicitly weighted the same when all the w_m are the same). When an approach like (4) is used, the analyst might recognize that conclusions depend on the particular threshold that is selected. The strategies pursued in this paper will be seen, for the most part, to circumvent these concerns.

---

[10] See, e.g., Geronimus et al., 2006, and Fleishman et al., 2014.

[11] Many of the studies summarized in the previous subsection use some form of composite outcome (see U.S. FDA, 2017, Mullahy, 2018b, and the discussion in this paper's section 5).

[12] A fourth approach is to consider simultaneously the entire portfolio of outcomes $\mathbf{y}$, model its joint probability structure $\Pr\left(\mathbf{y} \middle| \mathbf{x}\right)$, and estimate ATEs corresponding to some or all of its 2^M particular values $\Pr\left(\mathbf{y} = \mathbf{q} \middle| \mathbf{x}\right)$. Investigation of this approach is underway.

[13] In each of the cases described here the ATE would often be estimated by something akin to the difference between the estimated conditional-on-$\mathbf{x}$ probability or moment model evaluated at two different values of x_T and then averaged over $\mathbf{x}_{oth}$.

This Paper's Strategy

The paper's main strategy is to adapt for and adopt in multiple-outcome settings an underappreciated interpretation of familiar ATEs from the single-outcome context. While not typically invoked, this interpretation turns out to be fundamental to standard ATE definitions in the single-binary-outcome case. It is generalized here to provide a unified framework for considering population treatment effects when multiple outcomes are considered.

For any $M \geq 1$, the approach suggested here results in population-level statements about TEs akin to how ATEs in the single-outcome context summarize individual TEs across a population. It will be shown that the TE parameter proposed here, termed a *treatment-preference probability treatment effect*, or PTE, is essentially no different in the multiple-outcome and single-outcome contexts except that—under unconfounded treatment assignment—the PTE is point identified in the single-outcome case whereas in the multiple-outcome case some characterizations permit point identification while for others only partial or set identification is possible.

The approach proposed here for the multiple-outcome case has several attractive features. Except for a benchmark case that is developed as a link to some existing empirical practice, this paper's strategies require neither any across-outcome measurement comparability[14] nor any relative weighting of the outcomes; indeed, no equal or differential weighting of the outcomes is implied. More generally, the approach proposed here requires no aggregation of outcomes *per se*. If inference is of concern (the paper's focus is largely on definition and identification, although see footnote 37), no issues of multiple testing arise since the relevant parameters are ultimately scalar. Finally, once the relevant parameters are point or set identified computation is straightforward in most instances, with empirical computation generally requiring nothing more than specification and estimation of binary-outcome probability models (defined and discussed in section 5).

A cost of this approach is that decisionmakers must adopt one or more characterizations of treatment preference, an exercise akin to specifying what loss or value function is germane for the decision at hand. Seven such characterizations are suggested here, although others are imaginable. It is shown in section 2 that a particular notion of treatment preference is implied in considerations of ATEs in the single-outcome case. As such the demand that the decisionmaker adopt such a standard in the multiple-outcome case does not seem unreasonable even though, because of its simplicity, the decisionmaker is not explicitly confronted with such a decision when $M=1$.

While the focus here is largely on empirical issues the paper's emphases on treatment

[14] Comparability of measures in the binary-outcome case might seem, and indeed might be, trivially satisfied (although see footnote 20). Such considerations are more salient in the ordered-outcome cases considered in section 7.

preference necessitate consideration of the nature of preferences in multiple-outcome settings. At a practical level the considerations involve determining what outcomes matter to decisionmakers (e.g. patients or patient-provider teams, a fundamental concern in efforts to deliver high-value health care; see Lynn et al., 2015; Manski, 2018). At a theoretical level, the framework proposed by Manski, 2004, provides essential guidance for considering treatment preference for outcomes of arbitrary dimension. The particular notions of treatment preference described in section 3 are offered to be both potentially reasonable characterizations of what outcomes might actually concern decisionmakers as well as simple enough to serve as the basis of treatment-effect definitions that can be described and implemented empirically in a straightforward manner.

Plan for This Paper

Section 2 reviews treatment-effect analysis in the single-outcome context and develops a notion of *treatment preference* that is seen to be a feature, albeit one not typically recognized, of standard treatment-effect analysis in the one-outcome case. Using this idea, *treatment-preference probability treatment effects* (PTEs) are defined and are seen to correspond to familiar average treatment effects in the single-outcome case. In section 3 the focus turns to multiple-outcome settings. Focusing on binary outcomes, seven characterizations of treatment preference appropriate to multiple-outcome contexts are proposed, their corresponding probabilities and PTEs are derived, and the seven characterizations are compared and contrasted. Under standard assumptions about unconfoundedness of treatment assignment, section 4 shows that the corresponding PTEs are point identified for three of the characterizations and set identified for the other four, for which cases bounds are derived. Section 5 considers empirical approaches to bounds estimation, or the PTEs themselves in the point-identified cases. The results are illustrated with simulated data and, in section 6, in analyses of two microdata samples. Section 7 generalizes the binary-outcome case to consider treatment effects with multiple ordered outcomes, and suggests how that framework might be used to address some questions involving multiple continuous outcomes. Section 8 summarizes. While a fair amount of elementary probability algebra is used to derive results, the paper's main ideas as well as their empirical implementation turn out to be quite straightforward.

Ultimately if this paper accomplishes nothing more than stimulating readers to reassess their approaches to understanding multiple outcomes and how treatments and interventions effect them, it will have served some valuable purpose.

2. Treatment Preference and Treatment Effects with One Outcome

This section reviews standard TE analytics for the one-outcome case when the potential outcomes are binary, and then considers an interpretation of this setting's ATE that provides the foundation for the analysis of multiple-outcome treatment effects considered subsequently.

Potential Outcomes, Treatment Preference, and Treatment Effects When M=1

The setup is standard for M=1. There are two possible treatments, T_j and T_k, one of which will be assigned or administered. The treatments' features are summarized in observable vectors $\mathbf{x}_j$ and $\mathbf{x}_k$, which often have just one element. A single (M=1) potential outcome, y_j or y_k, is observed given treatment T_j or T_k. $y_\bullet$ is the generic y_j or y_k. $\mathbf{Y} = \left[y_j, y_k\right]$ denotes the 1×2 vector of potential outcomes, only one of which will be observed. The TE is $y_k - y_j$.

Until section 7 the focus is on binary[15] outcomes, with 1 and 0 indicating "bad" (e.g. unhealthy) and "good" (e.g. healthy) outcomes, respectively.[16] As such $TE \in \left\{-1, 0, 1\right\}$. Let treatment preference be characterized simply as preferring a "good" outcome to a "bad" outcome. Using "C" to denote this characterization of treatment preference, T_j is preferred to T_k ($T_j \succ_C T_k$) if $TE = -1$, T_k is preferred to T_j ($T_k \succ_C T_j$) if $TE = 1$, and the preference is neutral ($T_j \sim_C T_k$) if $TE = 0$.[17] This obvious point about treatment preference plays a central role in what follows.

The probability structure of the binary potential-outcome data is summarized in exhibit 1's contingency table, in which the $\pi_{\bullet\bullet}$ and $\pi_\bullet$ denote joint and marginal probabilities:

Exhibit 1

		y_j		
		0	1	
y_k	0	π_{00}	π_{10}	$1 - \pi_k$
	1	π_{01}	π_{11}	π_k
		$1 - \pi_j$	π_j	1

From this information the ATE can be expressed in three equivalent ways:

$$\text{ATE} \;=\; \mathrm{E}\left[y_k - y_j\right] \;=\; \pi_k - \pi_j \;=\; \pi_{01} - \pi_{10}\,. \tag{5}$$

[15] Section 7 considers the case of ordered outcomes. Brief considerations of continuous outcomes appear in several places in what follows.

[16] For concreteness, in health contexts one might think of "bad" binary outcomes along the lines of mortality, chronic illness diagnosis or onset, 30-day readmission, substance abuse, etc.

[17] It will become evident in section 3 why the "C" subscript—indicating a particular characterization of "bad" and "good" outcomes—is used alongside the preference indicator ("$\succ_C$").

Treatment-Preference-Probability Treatment Effects

When conceiving an ATE with binary outcomes it is natural to focus on the difference in marginal means or marginal probabilities, $\pi_k - \pi_j$ in (5). For this paper's purposes, however, the rightmost expression in (5), $\pi_{01} - \pi_{10}$, plays the central role. In light of the treatment-preference characterization noted above, $\pi_{01} = \Pr\left(y_j = 0 \wedge y_k = 1\right)$ and $\pi_{10} = \Pr\left(y_j = 1 \wedge y_k = 0\right)$, correspond respectively to the probabilities that T_j is preferred to T_k ($\Pr\left(T_j \succ_C T_k\right)$) and that T_k is preferred to T_j ($\Pr\left(T_k \succ_C T_j\right)$).[18] These are henceforth called "treatment-preference" probabilities. Thus the ATE (5) can be expressed as the difference in treatment-preference probabilities,

$$\mathrm{ATE}_{j \succ k, C} = \Pr\left(y_j \succ_C y_k\right) - \Pr\left(y_k \succ_C y_j\right). \tag{6}$$

From ATEs to PTEs

While the expression in (6) is algebraically little more than one summary of the information in exhibit 1, it is of fundamental importance for this paper's main goals. Specifically, the difference in treatment-preference probabilities expressed in (6) is the basis of the paper's strategy for characterizing population-level treatment effects when M>1, an idea developed in section 3. It is proposed that, for any M, such treatment effects can be meaningfully defined by the difference in treatment-preference probabilities as in (6), given some suitable characterization $C\bullet$ of what it means for one treatment to be preferred to another, i.e. $T_j \succ_{C\bullet} T_k$ and $T_k \succ_{C\bullet} T_j$. When M=1 and the outcomes are binary there is only one logical characterization of treatment preference, as suggested above. When M>1, however, there is no single unambiguous characterization of what it means for one treatment to be preferred to another.

For any M≥1 and any characterization of treatment preference, the events $T_j \succ_{C\bullet} T_k$ and $T_k \succ_{C\bullet} T_j$ have probabilities $\Pr\left(T_j \succ_{C\bullet} T_k\right)$ and $\Pr\left(T_k \succ_{C\bullet} T_j\right)$, whose difference in turn defines a *treatment-preference-probability treatment effect*, or "PTE":

$$\mathrm{PTE}_{j \succ k, C\bullet} = \Pr\left(T_j \succ_{C\bullet} T_k\right) - \Pr\left(T_k \succ_{C\bullet} T_k\right), \tag{7}$$

where the subscript "$j \succ k$" signifies the ordering of the minuend and subtrahend in (7). That is, it

[18] Note that the preference "events" $T_j \succ_C T_k$ and $T_k \succ_C T_j$ are stochastic as they depend on **Y**. Standard notation is used: "$\wedge$" denotes "and" (intersection), "$\vee$" denotes "or" (union).

is suggested that quantity used to summarize individual-level TEs across the population is the PTE, regardless of M. As such the sign and magnitude of $\text{PTE}_{\text{j}\succ\text{k},\text{C}\bullet}$ are proposed as standards for assessing treatment success, clinical significance, etc., for any value of M. Its interpretation is natural since it is exactly the familiar ATE in cases where M=1 and the outcomes are binary.

Rather than writing out repeatedly $\Pr\left(\text{T}_\text{j} \succ_{\text{C}\bullet} \text{T}_\text{k}\right)$, the shorthand $\mathcal{P}_{\text{j}\succ\text{k},\text{C}\bullet}$ will be used henceforth.[19] As such,

$$\text{PTE}_{\text{j}\succ\text{k},\text{C}\bullet} = \mathcal{P}_{\text{j}\succ\text{k},\text{C}\bullet} - \mathcal{P}_{\text{k}\succ\text{j},\text{C}\bullet}\,. \tag{8}$$

When M=1 the PTE (8) is identical to the ATE (6).[20] When M>1 the task is to determine decision-relevant characterizations $\text{C}\bullet$ of what is meant by "treatment preference" since comparisons of vector outcomes is less straightforward than comparison of scalar outcomes. Yet once $\text{C}\bullet$ is selected to characterize treatment-preference and define a corresponding PTE, the problem of defining a treatment effect in the M-dimension context is reduced to decisionmaking in a one-dimension context. The paper turns now to formal development of this idea.

---

[19] For M=1 essentially the same idea can be applied to continuously distributed univariate outcomes—e.g. measures where "larger" corresponds to "better" like survival time—for which $\mathcal{P}_{\text{j}\succ\text{k},\text{C}\bullet} = \Pr\left(\text{y}_\text{j} > \text{y}_\text{k}\right)$ (Mullahy, 2018a). Since there are no ties, $\Pr\left(\text{y}_\text{j} > \text{y}_\text{k}\right) = 1 - \Pr\left(\text{y}_\text{k} > \text{y}_\text{j}\right)$ so that $\text{PTE}_{\text{j}\succ\text{k},\text{C}\bullet} = 2\mathcal{P}_{\text{j}\succ\text{k},\text{C}\bullet} - 1$. Mapping continuous outcomes into binary representations when M>1 (e.g., Alkire and Foster, 2010) permits analyses to be conducted within this paper's framework albeit at the cost of potential information waste. See section 7 for additional discussion.

[20] Endowing a nominal- or ordinal-scale binary-outcome measure with ratio- or interval-scale properties may be questionable when computing quantities like an ATE. The restriction to the particular $\{0,1\}$ measure is easily loosened: Let the two possible outcome values be arbitrary values $\{\text{a},\text{b}\}$ and rewrite the vector of joint probabilities as $\left[\pi_{\text{aa}}, \pi_{\text{ab}}, \pi_{\text{ba}}, \pi_{\text{bb}}\right]$. Then

$$\text{ATE} \;=\; \text{E}\left[\text{y}_\text{k} - \text{y}_\text{j}\right] \;=\; \left(\text{b}-\text{a}\right)\times\left(\pi_\text{k} - \pi_\text{j}\right) \;=\; \left(\text{b}-\text{a}\right)\times\left(\pi_{\text{ab}} - \pi_{\text{ba}}\right),$$

a rescaled version of the ATE in (5). Whether the notion of an ATE is meaningful in such a binary-outcome context may depend on an application's structure and decisionmaker's objectives. Note that the ATE's expression in terms of $\pi_\text{k} - \pi_\text{j}$ arises from the ATE's averaging of the outcomes, not as a direct assertion of what constitutes a population-level TE. Yet without appealing at all to an ATE one might simply assert that $\pi_\text{k} - \pi_\text{j}$ or, equivalently, $\text{PTE}_{\text{j}\succ\text{k},\text{C}\bullet}$ is an appropriate TE; this circumvents concerns about measurement properties of the binary outcomes and whether these may or shouldn't lend themselves to numerical averaging as in (5).

3. Treatment Preference and Treatment Effects with Multiple Outcomes

The previous section showed that when M=1 it is straightforward to characterize what it means for one treatment to be preferred to another and to define—at least conceptually—the probability of such preference. Observation of the single outcome arising under each of the treatments reveals—at least conceptually—the relevant information. Considerations of identification are taken up in sections 4 and 5, but in anticipation it is useful to note that when M=1 $\mathcal{P}_{j\succ k,C}$ and $\mathcal{P}_{k\succ j,C}$ will not generally be point identified even under the best circumstances—e.g. unconfounded treatment assignment—although the $\text{PTE}_{j\succ k,C}$ they define would be.

With multiple outcomes a first-order concern not arising when M=1 is how to determine whether one treatment is preferred to another from a comparison of the M>1 potential outcomes arising from competing treatments. This issue has attracted surprisingly little attention in the health evaluation literature, particularly since samples containing and studies using data on multiple health outcomes are commonplace. While the empirical literature has handled *ad hoc* such data structures in a variety of ways (as noted in section 1), consideration of how to conceive of treatment preference and treatment effects in the multiple-outcome case is largely absent.

This section explores such issues and offers a set of criteria or characterizations by which one might assess the extent to which one treatment is preferred to another when treatments result in M>1 outcomes of interest. Not surprisingly matters are more complicated when M>1, but managing such complications should be a small price to pay for an integrated structure within which questions about multiple-outcome treatment effects can be explored.

Definitions

For $M\geq 1$ let $\mathbf{y}_j=\left[y_{j,1}\ \ldots\ y_{j,M}\right]$ and $\mathbf{y}_k=\left[y_{k,1}\ \ldots\ y_{k,M}\right]$ be M-vectors of binary potential outcomes, and let $\mathbf{Y}=\left[\mathbf{y}_j,\mathbf{y}_k\right]$ $(1\times 2M)$. $\mathbf{y}_\bullet$ denotes the generic version of either $\mathbf{y}_j$ or $\mathbf{y}_k$.[21] Let $\Pr\left(\mathbf{Y}\right)$ and $\Pr\left(\mathbf{y}_\bullet\right)$ denote the joint and joint-marginal probabilities of the potential outcomes. Let $\mathcal{Q}=\left\{\mathbf{q}\middle| q_m\in\{0,1\},\ m=1,\ldots,M\right\}$ be the set of all 2^M possible values of the potential outcomes $\mathbf{y}_\bullet$. For arbitrary vectors $\mathbf{a}$ and $\mathbf{b}$ let $\mathbf{a}>\mathbf{b}$ denote element-by-element strict inequality ($a_m>b_m$ for all m, or weak monotonicity) and let $\mathbf{a}\geq\mathbf{b}$ denote element-by-element weak inequality with at least one strict inequality (strong monotonicity). In what follows the $\mathbf{y}_\bullet$ are assumed to have a

[21] The components of the $\mathbf{y}_\bullet$ are considered fixed, but their particular specification is a key consideration in practice. For example, much effort is devoted to defining core outcome measures and standardized outcome sets (Porter et al., 2016; Williamson et al., 2017). See Mullahy, 2018b.

"multivariate" but not "multinomial" structure; that is, for all m $\Pr\left(y_{\bullet,m}=1\middle|\mathbf{y}_{\bullet,-m}=\mathbf{0}\right)\neq 1$, where $\mathbf{y}_{\bullet,-m}$ denotes $\mathbf{y}_{\bullet}$ without its m-th element. Boldface fonts denote vectors.

Using various characterizations $C\bullet$ of treatment preference, the main objective here is to define the treatment preferences, $T_j \succ_{C\bullet} T_k$ and $T_k \succ_{C\bullet} T_j$, the corresponding treatment-preference probabilities, $\mathcal{P}_{j\succ k,C\bullet}$ and $\mathcal{P}_{k\succ j,C\bullet}$, and the implied $PTE_{j\succ k,C\bullet}$ as in (8) that, in the multiple-outcome context, correspond to those quantities described in section 2 for M=1. After these definitions are provided, the discussion compares the properties of the various characterizations.[22] While the characterizations offered are hopefully both intuitive and reasonable, two considerations might be noted: first, other reasonable characterizations can be advanced; second, any standard for what it means for T_j to be preferred to T_k should be linked ideally to decisionmakers' values.

Characterization 1 (C1)

An intuitively obvious way to compare $\mathbf{y}_j$ and $\mathbf{y}_k$ when M>1 is to consider the events $\mathbf{y}_k \geq \mathbf{y}_j$ and $\mathbf{y}_j \geq \mathbf{y}_k$. In this instance there is no single standard—i.e. no particular value(s) of the $\mathbf{y}_{\bullet}$—defining a "good" or "bad" outcome. Instead the focus is on the set of inequality relationships that may obtain between $\mathbf{y}_j$ and $\mathbf{y}_k$ across the 2^{2M} possible values of $\mathbf{Y}$. In light of how "good" and "bad" are defined for each of the M component outcomes, $\mathbf{y}_k \geq \mathbf{y}_j$ may be a reasonable and natural characterization of T_j being preferred to T_k (and symmetrically, $\mathbf{y}_j \geq \mathbf{y}_k$ a reasonable characterization of T_k being preferred to T_j).

Formally T_j is preferred to T_k by characterization C1, denoted $T_j \succ_{C1} T_k$, if and only if $\mathbf{y}_k \geq \mathbf{y}_j$. In essence this corresponds to standard formal notions of strongly monotonic (decreasing) preferences. It follows that the probability that T_j is preferred to T_k under C1 is

$$\mathcal{P}_{j\succ k,C1} = \Pr\left(\mathbf{y}_k \geq \mathbf{y}_j\right), \tag{9}$$

[22] An obvious way to conceive of treatment preference is via treatment-specific utility. That is, one might consider a utility function $V(\ldots)$ and the expected utilities associated with treatments T_j and T_k, $EU_{\bullet} = \sum_{\mathbf{y}_{\bullet}\in Q} V\left(\mathbf{y}_{\bullet}\right)\times \Pr\left(\mathbf{y}_{\bullet}\right)$, where Q is the set of all 2^M possible outcomes. Under expected utility, $T_j \succ_{C,EU} T_k$ when $EU_j > EU_k$. Specifying $V(\ldots)$ determines how the elements of the $\mathbf{y}_{\bullet}$ are weighted; see Manski and Tetenov, 2018, for an example with M=2.

where this probability is necessarily defined from the full joint probability $\Pr\left(\mathbf{Y}\right)$. Define $\mathbb{Y}_{\mathrm{j}\succ\mathrm{k},\mathrm{C1}} = \left\{\mathbf{Y}\middle|\mathbf{y}_{\mathrm{k}} \geq \mathbf{y}_{\mathrm{j}}\right\}$ so that $\mathcal{P}_{\mathrm{j}\succ\mathrm{k},\mathrm{C1}} = \Pr\left(\mathbf{Y} \in \mathbb{Y}_{\mathrm{j}\succ\mathrm{k},\mathrm{C1}}\right)$. As such $\#\,\mathbb{Y}_{\mathrm{j}\succ\mathrm{k},\mathrm{C1}} = 3^{\mathrm{M}} - 2^{\mathrm{M}}$ (oeis.org, A001047), so that there are $2^{2\mathrm{M}} - 2\left(3^{\mathrm{M}} - 2^{\mathrm{M}}\right)$ values among the $2^{2\mathrm{M}}$ possible values of $\mathbf{Y}$ for which $\mathrm{T}_{\mathrm{j}} \sim_{\mathrm{C1}} \mathrm{T}_{\mathrm{k}}$. The PTE corresponding to (9) is

$$\mathrm{PTE}_{\mathrm{j}\succ\mathrm{k},\mathrm{C1}} \;=\; \mathcal{P}_{\mathrm{j}\succ\mathrm{k},\mathrm{C1}} - \mathcal{P}_{\mathrm{k}\succ\mathrm{j},\mathrm{C1}} \;=\; \Pr\left(\mathbf{y}_{\mathrm{k}} \geq \mathbf{y}_{\mathrm{j}}\right) \;-\; \Pr\left(\mathbf{y}_{\mathrm{j}} \geq \mathbf{y}_{\mathrm{k}}\right). \tag{10}$$

Characterization 2 (C2)

Characterization C2, which turns out to be a special case of C1, specifies that a "good" outcome is one where $\mathbf{y}_{\bullet} = \mathbf{0}$ whereas a "bad" outcome is any outcome where $\mathbf{y}_{\bullet} \neq \mathbf{0}$. That is, a treatment failure or bad outcome is one where *at least one* bad component outcome occurs. As such, $\mathrm{T}_{\mathrm{j}} \succ_{\mathrm{C2}} \mathrm{T}_{\mathrm{k}}$ if and only if $\mathbf{y}_{\mathrm{j}} = \mathbf{0} \wedge \mathbf{y}_{\mathrm{k}} \neq \mathbf{0}$, that is, when T_{j} results in no bad component outcome while T_{k} results in at least one bad component outcome. It follows that

$$\mathcal{P}_{\mathrm{j}\succ\mathrm{k},\mathrm{C2}} = \Pr\left(\mathbf{y}_{\mathrm{j}} = \mathbf{0} \wedge \mathbf{y}_{\mathrm{k}} \neq \mathbf{0}\right) \tag{11}$$

For C2 a neutral treatment preference occurs when $\mathbf{y}_{\mathrm{j}} = \mathbf{0} \wedge \mathbf{y}_{\mathrm{k}} = \mathbf{0}$ or when $\mathbf{y}_{\mathrm{j}} \neq \mathbf{0} \wedge \mathbf{y}_{\mathrm{k}} \neq \mathbf{0}$. Define $\mathbb{Y}_{\mathrm{j}\succ\mathrm{k},\mathrm{C2}} = \left\{\mathbf{Y}\middle|\mathbf{y}_{\mathrm{j}} = \mathbf{0}, \mathbf{y}_{\mathrm{k}} \neq \mathbf{0}\right\}$. Then $\#\,\mathbb{Y}_{\mathrm{j}\succ\mathrm{k},\mathrm{C2}} = 2^{\mathrm{M}} - 1$ so that $2^{2\mathrm{M}} - 2\left(2^{\mathrm{M}} - 1\right)$ values of $\mathbf{Y}$ result in $\mathrm{T}_{\mathrm{j}} \sim_{\mathrm{C2}} \mathrm{T}_{\mathrm{k}}$. The PTE corresponding to (11) is

$$\mathrm{PTE}_{\mathrm{j}\succ\mathrm{k},\mathrm{C2}} \;=\; \mathcal{P}_{\mathrm{j}\succ\mathrm{k},\mathrm{C2}} - \mathcal{P}_{\mathrm{k}\succ\mathrm{j},\mathrm{C2}} \;=\; \Pr\left(\mathbf{y}_{\mathrm{j}} = \mathbf{0} \wedge \mathbf{y}_{\mathrm{k}} \neq \mathbf{0}\right) \;-\; \Pr\left(\mathbf{y}_{\mathrm{j}} \neq \mathbf{0} \wedge \mathbf{y}_{\mathrm{k}} = \mathbf{0}\right). \tag{12}$$

Characterization 3 (C3)

The third characterization considered here, C3, is the mirror image of C2 and is also a special case of C1. Here a "good" outcome is one where $\mathbf{y}_{\bullet} \neq \mathbf{1}$ while a "bad" outcome occurs when $\mathbf{y}_{\bullet} = \mathbf{1}$. That is, a "bad" outcome under C3 is one where *all* component outcomes are "bad"; if any component is not "bad" then the overall outcome is "good." Thus $\mathrm{T}_{\mathrm{j}} \succ_{\mathrm{C3}} \mathrm{T}_{\mathrm{k}}$ if and only if $\mathbf{y}_{\mathrm{j}} \neq \mathbf{1} \wedge \mathbf{y}_{\mathrm{k}} = \mathbf{1}$, that is, when T_{j} results in at least one "good" component outcome while T_{k} results in no "good" component outcome. It follows that

$$\mathcal{P}_{\mathrm{j}\succ\mathrm{k},\mathrm{C3}} = \Pr\left(\mathbf{y}_{\mathrm{j}} \neq \mathbf{1} \wedge \mathbf{y}_{\mathrm{k}} = \mathbf{1}\right) \tag{13}$$

A neutral treatment preference with C3 occurs when $\mathbf{y}_j = \mathbf{1} \wedge \mathbf{y}_k = \mathbf{1}$ or when $\mathbf{y}_j \neq \mathbf{1} \wedge \mathbf{y}_k \neq \mathbf{1}$. Define $\mathbb{Y}_{j\succ k,C3} = \left\{\mathbf{Y} \middle| \mathbf{y}_j \neq \mathbf{1}, \mathbf{y}_k = \mathbf{1}\right\}$. Then as with C2 $\#\mathbb{Y}_{j\succ k,C3} = 2^M - 1$ so that $2^{2M} - 2\left(2^M - 1\right)$ values of $\mathbf{Y}$ result in $T_j \sim_{C3} T_k$. The PTE corresponding to (13) is

$$PTE_{j\succ k,C3} = \mathcal{P}_{j\succ k,C3} - \mathcal{P}_{k\succ j,C3} = \Pr\left(\mathbf{y}_j \neq \mathbf{1} \wedge \mathbf{y}_k = \mathbf{1}\right) - \Pr\left(\mathbf{y}_j = \mathbf{1} \wedge \mathbf{y}_k \neq \mathbf{1}\right). \quad (14)$$

Characterization 4 (C4)

C4 is the union of C2 and C3. There is no single "good" or "bad" outcome. Rather treatment preference is determined as $T_j \succ_{C4} T_k$ if and only if either $\left(\mathbf{y}_j = \mathbf{0} \wedge \mathbf{y}_k \neq \mathbf{0}\right)$ or $\left(\mathbf{y}_j \neq \mathbf{1} \wedge \mathbf{y}_k = \mathbf{1}\right)$. $T_j \succ_{C4} T_k$ if T_j results in no "bad" component outcome while T_k results in at least one "bad" component outcome, or if T_k results in all "bad" component outcomes while T_j results in at least one "good" component outcome. For health outcomes, C4 means "perfect" health is better than "imperfect" health and partially "imperfect" health is better than completely "imperfect" health. Thus

$$\mathcal{P}_{j\succ k,C4} = \Pr\left(\left(\mathbf{y}_j = \mathbf{0} \wedge \mathbf{y}_k \neq \mathbf{0}\right) \vee \left(\mathbf{y}_j \neq \mathbf{1} \wedge \mathbf{y}_k = \mathbf{1}\right)\right). \quad (15)$$

Neutral treatment preference occurs if $\mathbf{y}_j = \mathbf{y}_k = \mathbf{0}$ or $\mathbf{y}_j = \mathbf{y}_k = \mathbf{1}$ or if neither $\mathbf{y}_j$ nor $\mathbf{y}_k$ is in $\left\{\mathbf{0},\mathbf{1}\right\}$. Let

$$\mathbb{Y}_{j\succ k,C4} = \left\{\mathbf{Y} \middle| \left(\mathbf{y}_j = \mathbf{0} \wedge \mathbf{y}_k \neq \mathbf{0}\right) \vee \left(\mathbf{y}_j \neq \mathbf{1} \wedge \mathbf{y}_k = \mathbf{1}\right)\right\} = \mathbb{Y}_{j\succ k,C2} \bigcup \mathbb{Y}_{j\succ k,C3}. \quad (16)$$

$\#\mathbb{Y}_{j\succ k,C4} = 2^{M+1} - 3$ so that $2^{2M} - 2^{M+2} + 6$ among the 2^{2M} possible values of $\mathbf{Y}$ result in $T_j \sim_{C4} T_k$. The PTE corresponding to (15) is

$$\begin{aligned} PTE_{j\succ k,C4} = \mathcal{P}_{j\succ k,C4} - \mathcal{P}_{k\succ j,C4} = {} & \Pr\left(\left(\mathbf{y}_j = \mathbf{0} \wedge \mathbf{y}_k \neq \mathbf{0}\right) \vee \left(\mathbf{y}_j \neq \mathbf{1} \wedge \mathbf{y}_k = \mathbf{1}\right)\right) - \\ & \Pr\left(\left(\mathbf{y}_j \neq \mathbf{0} \wedge \mathbf{y}_k = \mathbf{0}\right) \vee \left(\mathbf{y}_j = \mathbf{1} \wedge \mathbf{y}_k \neq \mathbf{1}\right)\right) \end{aligned} \quad (17)$$

Characterization 5 (C5)

C5 is the intersection of C2 and C3. A "good" outcome occurs when $\mathbf{y}_\bullet = \mathbf{0}$ while a "bad" outcome occurs when $\mathbf{y}_\bullet = \mathbf{1}$. Treatment preference is determined as $T_j \succ_{C5} T_k$ if and only if

$\mathbf{y}_j = \mathbf{0} \wedge \mathbf{y}_k = \mathbf{1}$. That is, $T_j \succ_{C5} T_k$ if and only if T_j results in no "bad" component outcome while T_k results in all "bad" component outcomes. With binary outcomes this is equivalent to $\mathbf{y}_k > \mathbf{y}_j$, corresponding to weakly monotonic preferences. For health outcomes, C5 suggests that "perfect" health is better than "worst possible" health, and otherwise does not adjudicate. Here

$$\mathcal{P}_{j\succ k,C5} = \Pr\left(\mathbf{y}_j = \mathbf{0} \wedge \mathbf{y}_k = \mathbf{1}\right). \tag{18}$$

A neutral treatment preference under C5 occurs in all cases except $\mathbf{y}_j = \mathbf{0} \wedge \mathbf{y}_k = \mathbf{1}$. Define

$$\mathbb{Y}_{j\succ k,C5} \;=\; \left\{\mathbf{Y}\middle|\mathbf{y}_j = \mathbf{0} \wedge \mathbf{y}_k = \mathbf{1}\right\} \;=\; \mathbb{Y}_{j\succ k,C2} \bigcap \mathbb{Y}_{j\succ k,C3}. \tag{19}$$

$\#\,\mathbb{Y}_{j\succ k,C5} = 1$ so that $2^{2M} - 2$ values of $\mathbf{Y}$ result in $T_j \sim_{C5} T_k$. The PTE corresponding to (18) is

$$\mathrm{PTE}_{j\succ k,C5} \;=\; \mathcal{P}_{j\succ k,C5} - \mathcal{P}_{k\succ j,C5} \;=\; \Pr\left(\mathbf{y}_j = \mathbf{0} \wedge \mathbf{y}_k = \mathbf{1}\right) \;-\; \Pr\left(\mathbf{y}_j = \mathbf{1} \wedge \mathbf{y}_k = \mathbf{0}\right). \tag{20}$$

Characterization 6 (C6)

C6 is a generic characterization that will be seen to have important commonalities with characterizations C2 and C3 when considerations of identification are raised in sections 4 and 5. Define the set $Z \subset Q$ and its complement in Q as Z^c. For C6 "good" outcomes are those where $\mathbf{y}_\bullet \in Z$ while "bad" outcomes occur when $\mathbf{y}_\bullet \in Z^c$.[23] As such, treatment preference is determined as

[23] C6 is offered to encompass various multiple-outcome settings encountered in practice. In applied health research, criteria beyond ones based simply on $Z = \{\mathbf{0}\}$ or $Z^c = \{\mathbf{1}\}$ are used to define composite outcomes (U.S. FDA, 2017). Consider two such cases. In the first a "good" outcome is one where no more than z<M component outcomes are "bad", i.e. a "bad" outcome requires at least z+1 "bad" component outcomes (e.g., metabolic syndrome (M=5, z=2); U.S. NHLBI, 2018); which particular component outcomes are "bad" doesn't matter, only that at least z+1 of them are. In the second a "bad" outcome is one where w particular component(s) and at least z other components be "bad", otherwise the outcome is "good" (e.g., DSM-V narcolepsy (M=4, w=1, z=1); Ruoff and Rye, 2016). In the extreme, w components may represent outcomes particularly important to a decisionmaker so that a "good" outcome is any outcome where these w outcomes are "good" (akin to "essential factors," Färe and Svensson, 1980). Treatment preference depends here only on the essential component(s) (i.e. z=0); in effect M=w. This encompasses so-called primary and secondary outcomes (U.S. FDA, 2017). In technology evaluations (e.g. RCTs) outcomes are sometimes prioritized as primary or secondary and the technology is deemed successful if the (cont.)

$T_j \succ_{C6(Z)} T_k$ if and only if $\mathbf{y}_j \in Z \wedge \mathbf{y}_k \in Z^c$. Thus,

$$\mathcal{P}_{j\succ k,C6(Z)} = \Pr\left(\mathbf{y}_j \in Z \wedge \mathbf{y}_k \in Z^c\right). \tag{21}$$

A neutral treatment preference under C6 occurs when $\mathbf{y}_j \in Z \wedge \mathbf{y}_k \in Z$ or when $\mathbf{y}_j \in Z^c \wedge \mathbf{y}_k \in Z^c$. Define $\mathbb{Y}_{j\succ k,C6(Z)} = \left\{\mathbf{Y}\middle|\mathbf{y}_j \in Z \wedge \mathbf{y}_k \in Z^c\right\}$. The PTE corresponding to (21) is

$$PTE_{j\succ k,C6(Z)} \;=\; \mathcal{P}_{j\succ k,C6(Z)} - \mathcal{P}_{k\succ j,C6(Z)} \;=\; \Pr\left(\mathbf{y}_j \in Z \wedge \mathbf{y}_k \in Z^c\right) \;-\; \Pr\left(\mathbf{y}_j \in Z^c \wedge \mathbf{y}_k \in Z\right). \tag{22}$$

Outcome Counts: A Benchmark Characterization (C0)

Each of the characterizations C1-C6 provides an unambiguous basis for comparing the outcome vectors $\mathbf{y}_j$ and $\mathbf{y}_k$, and does so in a way that relates treatment preference to a particular interpretation of what it means for one binary vector to be element-by-element "better" than another. Some potentially relevant characterizations of treatment preference, though, will not be based on such vector relationships. One such characterization is both obvious—since its empirical counterparts are encountered frequently in applications—and provides a benchmark against which properties and implications of other characterizations might be assessed.

Specifically, with M>1 binary outcomes a prominent data-aggregation or dimension-reduction approach used in practice is based on the count of "bad" outcomes (see Khan et al., 2008, for one example). Let $s_j = \sum_{m=1}^{M} y_{j,m}$ and likewise define s_k. Like C1, characterization C0 does not involve a particular standard for $\mathbf{y}_\bullet$ of what constitutes a "good" or "bad" outcome, but relies on a comparison of s_j and s_k wherein $T_j \succ_{C0} T_k$ if and only if $s_k > s_j$. It follows that

$$\mathcal{P}_{j\succ k,C0} = \Pr\left(s_k > s_j\right). \tag{23}$$

(cont.)
primary outcomes are all "good" regardless of the secondary components' outcomes. For illustration, let M=3, w=1, and z=1. With $y_{\bullet,1}$ being the essential component, the two scenarios correspond to

$Z = \left\{[0,0,0],[1,0,0],[0,1,0],[0,0,1]\right\}$, $Z^c = \left\{[1,1,0],[1,0,1],[0,1,1],[1,1,1]\right\}$, and
$Z = \left\{[0,0,0],[1,0,0],[0,1,0],[0,0,1],[0,1,1]\right\}$, $Z^c = \left\{[1,1,0],[1,0,1],[1,1,1]\right\}$.

Define $\mathbb{Y}_{j\succ k,C0} = \left\{\mathbf{Y}\middle| s_k > s_j\right\}$. $\#\mathbb{Y}_{j\succ k,C0} = 2^{2M-1} - \frac{(2M-1)!}{M!(M-1)!}$; $(2M)!/(M!)^2$ values of $\mathbf{Y}$ result in $T_j \sim_{C0} T_k$.[24] The PTE corresponding to (23) is

$$PTE_{j\succ k,C0} \;=\; \mathcal{P}_{j\succ k,C0} - \mathcal{P}_{k\succ j,C0} \;=\; \Pr\left(s_k > s_j\right) - \Pr\left(s_j > s_k\right). \tag{24}$$

Comparing the Treatment-Preference Characterizations

The treatment-preference characterizations are summarized in exhibit 2.

Exhibit 2: Summary of $T_j \succ_{C\bullet} T_k$ Characterizations

	$T_j \succ_{C\bullet} T_k$
C0	$s_k > s_j$
C1	$\mathbf{y}_k \geq \mathbf{y}_j$
C2	$\mathbf{y}_j = \mathbf{0} \wedge \mathbf{y}_k \neq \mathbf{0}$
C3	$\mathbf{y}_j \neq \mathbf{1} \wedge \mathbf{y}_k = \mathbf{1}$
C4	$\left(\mathbf{y}_j = \mathbf{0} \wedge \mathbf{y}_k \neq \mathbf{0}\right) \vee \left(\mathbf{y}_j \neq \mathbf{1} \wedge \mathbf{y}_k = \mathbf{1}\right)$
C5	$\mathbf{y}_j = \mathbf{0} \wedge \mathbf{y}_k = \mathbf{1}$
C6	$\mathbf{y}_j \in Z \wedge \mathbf{y}_k \in Z^c$

Using the definitions of the $\mathbb{Y}_{j\succ k,C\bullet}$ above, it is straightforward to show that for M>2

$$\mathbb{Y}_{j\succ k,C5} \;\subset\; \left\{\begin{array}{c} \mathbb{Y}_{j\succ k,C2} \\ \mathbb{Y}_{j\succ k,C3} \end{array}\right\} \;\subset\; \mathbb{Y}_{j\succ k,C4} \;\subset\; \mathbb{Y}_{j\succ k,C1} \;\subset\; \mathbb{Y}_{j\succ k,C0} \tag{25}$$

(for M=2 $\mathbb{Y}_{j\succ k,C0} = \mathbb{Y}_{j\succ k,C1} = \mathbb{Y}_{j\succ k,C4}$). C1 is more restrictive than C0 (for M>2) since the set of events $\mathbf{y}_k \geq \mathbf{y}_j$ is a subset of all events for which $s_k > s_j$. In turn C2 and C3 are more restrictive than C1 since inequalities $\mathbf{y}_k \geq \mathbf{y}_j$ with $\mathbf{y}_j = \mathbf{0}$ or with $\mathbf{y}_k = \mathbf{1}$ are a subset of all possible $\mathbf{y}_k \geq \mathbf{y}_j$ inequalities (as for C1); C2 and C3 are more restrictive than C4 since $\mathbb{Y}_{j\succ k,C4} = \mathbb{Y}_{j\succ k,C2} \bigcup \mathbb{Y}_{j\succ k,C3}$. Similarly C4 is more restrictive than C1 since all $\mathbf{Y}$ in $\mathbb{Y}_{j\succ k,C4}$ satisfy $\mathbf{y}_k \geq \mathbf{y}_j$ but some $\mathbf{Y}$ with

[24] See oeis.org, entries A000346 and A000984.

$\mathbf{y}_k \geq \mathbf{y}_j$ are not in $\mathbb{Y}_{j\succ k,C4}$. C5 is more restrictive than either C2 or C3 since it is the intersection of those characterizations. C2-C5 are all special cases of the $\mathbf{y}_k \geq \mathbf{y}_j$ inequalities that define C1. C2 and C3 are special cases of C6.

Exhibit 3: $\mathbf{Y}$ Values Relevant in C2-C5 Characterizations

		$\mathbf{y}_j$		
		$\mathbf{y}_j = \mathbf{0}$	$\mathbf{y}_j \neq \mathbf{0} \wedge \mathbf{y}_j \neq \mathbf{1}$	$\mathbf{y}_j = \mathbf{1}$
$\mathbf{y}_k$	$\mathbf{y}_k = \mathbf{0}$	$\mathbb{Y}_{0,0}$	$\mathbb{Y}_{\sim,0}$	$\mathbb{Y}_{1,0}$
	$\mathbf{y}_k \neq \mathbf{0} \wedge \mathbf{y}_k \neq \mathbf{1}$	$\mathbb{Y}_{0,\sim}$	$\mathbb{Y}_{\sim,\sim}$	$\mathbb{Y}_{1,\sim}$
	$\mathbf{y}_k = \mathbf{1}$	$\mathbb{Y}_{0,1}$	$\mathbb{Y}_{\sim,1}$	$\mathbb{Y}_{1,1}$

In exhibit 3's cells the $\mathbb{Y}_{\bullet,\bullet}$ represent sets of $\mathbf{Y}$ values corresponding to the particular margin definitions (e.g. $\mathbb{Y}_{0,0} = \left\{\mathbf{Y}\middle|\mathbf{y}_j = \mathbf{0} \wedge \mathbf{y}_k = \mathbf{0}\right\}$). Thus $T_j \succ_{C\bullet} T_k$ is determined by:

C2: $\mathbf{y}_j = \mathbf{0} \wedge \mathbf{y}_k \neq \mathbf{0}$, i.e. $\mathbf{Y} \in \mathbb{Y}_{0,\sim} \bigcup \mathbb{Y}_{0,1}$

C3: $\mathbf{y}_j \neq \mathbf{1} \wedge \mathbf{y}_k = \mathbf{1}$, i.e. $\mathbf{Y} \in \mathbb{Y}_{0,1} \bigcup \mathbb{Y}_{\sim,1}$

C4: $\left(\mathbf{y}_j = \mathbf{0} \wedge \mathbf{y}_k \neq \mathbf{0}\right) \vee \left(\mathbf{y}_j \neq \mathbf{1} \wedge \mathbf{y}_k = \mathbf{1}\right)$, i.e. $\mathbf{Y} \in \mathbb{Y}_{0,\sim} \bigcup \mathbb{Y}_{0,1} \bigcup \mathbb{Y}_{\sim,1}$

C5: $\left(\mathbf{y}_j = \mathbf{0} \wedge \mathbf{y}_k = \mathbf{1}\right)$, i.e. $\mathbf{Y} \in \mathbb{Y}_{0,1}$

If one imagines the entries in the exhibit as a 3×3 array, then C2-C5 are different ways of asserting that outcomes represented in its lower off-diagonal elements are ones where $T_j \succ_{C\bullet} T_k$.

Ignoring for the moment C0 and C6, another perspective recognizes that C1 and C5 bracket C2, C3, and C4 in the sense that C1 is defined by weakest possible vector inequality $\mathbf{y}_k \geq \mathbf{y}_j$ while C5 is defined by the the strictest vector inequality $\mathbf{y}_k > \mathbf{y}_j$. C2, C3, and C4 fall in-between, more structured than C1 but less stringent than C5. Since C1 and C5 correspond to strongly and weakly monotone preference structures, C2-C4 are thus special cases of strongly monotone preferences. For illustrative purposes table 1 shows relationships among the $C\bullet$ for M=3.

[table 1 about here]

From (25) it follows that[25]

$$\mathcal{P}_{\text{j}\succ\text{k,C0}} \;\geq\; \mathcal{P}_{\text{j}\succ\text{k,C1}} \;\geq\; \mathcal{P}_{\text{j}\succ\text{k,C4}} \;\geq\; \left\{ \begin{array}{c} \mathcal{P}_{\text{j}\succ\text{k,C2}} \\ \mathcal{P}_{\text{j}\succ\text{k,C3}} \end{array} \right\} \;\geq\; \mathcal{P}_{\text{j}\succ\text{k,C5}}\,, \tag{26}$$

with weak inequalities accommodating the possibility that the probabilities of the particular joint events that define the differences among the $\mathbb{Y}_{\text{j}\succ\text{k,C}\bullet}$ sets may be zero.

Finally since $\mathcal{P}_{\text{j}\succ\text{k,C}\bullet}$, $\text{PTE}_{\text{j}\succ\text{k,C}\bullet}$, or both might be of interest in particular contexts, the following section considers identification of each.

4. Identifying $\mathcal{P}_{\text{j}\succ\text{k,C}\bullet}$ and $\text{PTE}_{\text{j}\succ\text{k,C}\bullet}$

The treatment-preference probabilities discussed in section 3 are defined by features of the full joint distribution $\Pr\left(\mathbf{Y}\right)$. As such, while point identifying one or more of the $\mathcal{P}_{\text{j}\succ\text{k,C}\bullet}$ and the corresponding PTEs may be desirable, this will not always be possible. Nonetheless, partial or set identification is generally possible so long as the joint marginals $\Pr\left(\mathbf{y}_{\bullet}\right)$ are identified, and for three of the characterizations described in the previous section point identification of the $\text{PTE}_{\text{j}\succ\text{k,C}\bullet}$ is possible. This section explores these identification questions.

The standard sense of identification is used here, i.e. that the parameters of interest are uniquely determined by the (joint) distribution of observable/observed data (Hansen, 2018, section 2.33). This conveys notions of both conceptual (population) as well as empirical (sample) identification. Treatment assignment is assumed exogenous, i.e. unconfoundedness is assumed, so point or set identification of this paper's key parameters is broadly straightforward: using analog principles, features of conditional-on-treatment distributions of observed outcomes will generally suffice to estimate consistently the corresponding features of treatment-specific potential-outcome distributions. It is assumed that the joint marginals $\Pr\left(\mathbf{y}_{\bullet}\right)$ are knowable and accessible, which will typically be the case.[26] If covariates beyond treatment-status indicators are relevant, they are assumed to be benign and can innocuously condition probability statements.

With unconfounded treatment assignment, $\Pr\left(\mathbf{y}_{\bullet}\right)$ will generally be nonparametrically

---

[25] It can be shown that $\mathcal{P}_{\text{j}\succ\text{k,C2}} \gtreqless \mathcal{P}_{\text{j}\succ\text{k,C3}}$ as $\Pr\left(\mathbf{y}_{\text{k}} \notin \left\{\mathbf{0},\mathbf{1}\right\}\right) \gtreqless \Pr\left(\mathbf{y}_{\text{j}} \notin \left\{\mathbf{0},\mathbf{1}\right\}\right)$.

[26] Note that the empirical joint marginals may be obtained from two arms of a randomized trial or even from two separate samples (e.g. repeated cross sections, synthetic cohorts). The key in any case is that the samples represent the same population, however that population be defined.

identified and estimable from observable data as $\widehat{\Pr}\left(\mathbf{y}\middle|\mathbf{x}=\mathbf{x}_{\bullet}\right)$ (section 5). As seen below this will typically suffice to set identify $\mathcal{P}_{j\succ k,C\bullet}$ and $\mathcal{P}_{k\succ j,C\bullet}$, and then estimate bounds thereon. It is also shown that point identification of the $\text{PTE}_{j\succ k,C\bullet}$ is possible for C2, C3, and C6.

Probability Bounds: General Results and Results for M=1

The approach described by Boole, 1854 (chapter XIX), and others[27] is used to determine the probability and PTE bounds. In general Boole's upper bounds (UB) and lower bounds (LB) are straightforward to derive given knowledge of the joint marginal probabilities. For the most part the required bounds will be seen to be those on conjunction probabilities ("intersection," "and").[28]

To illustrate, consider first the M=1 case discussed in section 2. With reference to exhibit 1, consider $\mathcal{P}_{j\succ k,C}=\Pr\left(y_j=0\wedge y_k=1\right)=\pi_{01}$. (The results are shown here for $\mathcal{P}_{j\succ k,C\bullet}$; switching subscripts gives the results for $\mathcal{P}_{k\succ j,C\bullet}$.) While π_{01} cannot be point identified it can be bounded by using the identified marginal distribution probabilities π_j and π_k as

$$\text{UB}\left(\mathcal{P}_{j\succ k,C}\right) \;=\; \min\left\{\Pr\left(y_j=0\right),\Pr\left(y_k=1\right)\right\} \;=\; \min\left\{1-\pi_j,\pi_k\right\} \tag{27}$$

and

$$\text{LB}\left(\mathcal{P}_{j\succ k,C}\right) \;=\; \max\left\{\Pr\left(y_j=0\right)+\Pr\left(y_k=1\right)-1,0\right\} \;=\; \max\left\{\pi_k-\pi_j,0\right\}. \tag{28}$$

For nondegenerate cases the UB is always informative (i.e. less than one) while the LB may or may not be informative (i.e. exceed zero).[29]

Even though $\mathcal{P}_{j\succ k,C}$ and $\mathcal{P}_{k\succ j,C}$ are not themselves point identified, the corresponding

---

[27] Although such bounds are often referred to as Fréchet or Fréchet-Hoeffding bounds, Boole's 1854 treatise was published before either Fréchet or Hoeffding was born.

[28] For N events e_n jointly distributed in the population as $\Pr\left(e_1\wedge\ldots\wedge e_N\right)$, the general result is

$$\begin{aligned}\text{UB}\left(\Pr\left(e_1\wedge\ldots\wedge e_N\right)\right) \;=\; \min\left\{\Pr\left(e_1\right),\ldots,\Pr\left(e_N\right)\right\} \;\geq\; \Pr\left(e_1\wedge\ldots\wedge e_N\right) \;\geq\\ \max\left\{\sum_{n=1}^{N}\Pr\left(e_n\right)-\left(N-1\right),0\right\} \;=\; \text{LB}\left(\Pr\left(e_1\wedge\ldots\wedge e_N\right)\right)\end{aligned}$$

[29] For the M=1 continuous-outcome case discussed in footnote 19, the bounds on the PTE $2\Pr\left(y_1>y_0\right)-1$ are (using notation in Mullahy, 2018a, eq. (21)):

$$\text{UB}\left(2\Pr\left(y_1>y_0\right)-1\right)=1-2D_{10} \quad\text{and}\quad \text{LB}\left(2\Pr\left(y_1>y_0\right)-1\right)=2D_{01}-1$$

$\text{PTE}_{\text{j}\succ\text{k,C}}$ is point identified by the ATE in (6) under unconfounded treatment assignment, i.e.

$$\text{PTE}_{\text{j}\succ\text{k,C}} \;=\; \mathcal{P}_{\text{j}\succ\text{k,C}} - \mathcal{P}_{\text{k}\succ\text{j,C}} \;=\; \pi_{01} - \pi_{10} \;=\; \pi_{\text{k}} - \pi_{\text{j}} \;=\; \text{E}\Big[\text{y}_{\text{k}} - \text{y}_{\text{j}}\Big]. \tag{29}$$

Bounding (Set Identification of) $\mathcal{P}_{j\succ k,C\bullet}$ and $\mathcal{P}_{k\succ j,C\bullet}$ with M>1

For M>1, consider first C2, C3, C5, and C6 as these characterizations' preference-probability and PTE bounds are straightforward to obtain. For these $\text{C}\bullet$ there are only two marginal events to consider in obtaining bounds on the treatment-preference probabilities. Using (11), (13), (18), and (21), for each of these $\text{C}\bullet$ the marginal events represent the "good" and "bad" outcomes corresponding to each treatment. Thus,

$$\text{UB}\Big(\mathcal{P}_{\text{j}\succ\text{k,C2}}\Big) = \min\Big\{\Pr\Big(\mathbf{y}_{\text{j}} = \mathbf{0}\Big), 1 - \Pr\Big(\mathbf{y}_{\text{k}} = \mathbf{0}\Big)\Big\} \tag{30}$$

$$\text{LB}\Big(\mathcal{P}_{\text{j}\succ\text{k,C2}}\Big) = \max\Big\{\Pr\Big(\mathbf{y}_{\text{j}} = \mathbf{0}\Big) - \Pr\Big(\mathbf{y}_{\text{k}} = \mathbf{0}\Big), 0\Big\} \tag{31}$$

$$\text{UB}\Big(\mathcal{P}_{\text{j}\succ\text{k,C3}}\Big) = \min\Big\{1 - \Pr\Big(\mathbf{y}_{\text{j}} = \mathbf{1}\Big), \Pr\Big(\mathbf{y}_{\text{k}} = \mathbf{1}\Big)\Big\} \tag{32}$$

$$\text{LB}\Big(\mathcal{P}_{\text{j}\succ\text{k,C3}}\Big) = \max\Big\{\Pr\Big(\mathbf{y}_{\text{k}} = \mathbf{1}\Big) - \Pr\Big(\mathbf{y}_{\text{j}} = \mathbf{1}\Big), 0\Big\}. \tag{33}$$

$$\text{UB}\Big(\mathcal{P}_{\text{j}\succ\text{k,C5}}\Big) = \min\Big\{\Pr\Big(\mathbf{y}_{\text{j}} = \mathbf{0}\Big), \Pr\Big(\mathbf{y}_{\text{k}} = \mathbf{1}\Big)\Big\} \tag{34}$$

$$\text{LB}\Big(\mathcal{P}_{\text{j}\succ\text{k,C5}}\Big) = \max\Big\{\Pr\Big(\mathbf{y}_{\text{j}} = \mathbf{0}\Big) + \Pr\Big(\mathbf{y}_{\text{k}} = \mathbf{1}\Big) - 1, 0\Big\}. \tag{35}$$

$$\text{UB}\Big(\mathcal{P}_{\text{j}\succ\text{k,C6(Z)}}\Big) = \min\Big\{\Pr\Big(\mathbf{y}_{\text{j}} \in \text{Z}\Big), \Pr\Big(\mathbf{y}_{\text{k}} \in \text{Z}^{\text{c}}\Big)\Big\} \tag{36}$$

$$\text{LB}\Big(\mathcal{P}_{\text{j}\succ\text{k,C6(Z)}}\Big) = \max\Big\{\Pr\Big(\mathbf{y}_{\text{j}} \in \text{Z}\Big) + \Pr\Big(\mathbf{y}_{\text{k}} \in \text{Z}^{\text{c}}\Big) - 1, 0\Big\}. \tag{37}$$

Exhibit 3 allows one to visualize how the quantities (30)-(37) bound $\mathcal{P}_{\text{j}\succ\text{k,C}\bullet}$ for C2, C3, and C5. Note that for both C2 and C3 one of the lower bounds $\text{LB}\Big(\mathcal{P}_{\text{j}\succ\text{k,C}\bullet}\Big)$ or $\text{LB}\Big(\mathcal{P}_{\text{k}\succ\text{j,C}\bullet}\Big)$ will necessarily be positive while the other will be zero (assuming no ties).

Since it is defined by a disjunction of conjunctions one approach to computing bounds on $\mathcal{P}_{\text{j}\succ\text{k,C4}}$ involves, in essence, bounding the bounds. Several approaches could be used along these lines, not all of which will result in identical numerical results (see footnote 32 below).[30] For the LB

[30] Corresponding to footnote 28, Boole's upper and lower bounds for probabilities $\Pr\Big(\text{e}_1 \vee \ldots \vee \text{e}_{\text{N}}\Big)$ of disjunctions of possibly non-disjoint events or sets e_{n} are:

(cont.)

a direct approach based on this idea gives

$$\mathrm{LB}\left(\mathcal{P}_{\mathrm{j}\succ\mathrm{k},\mathrm{C4}}\right)=\max\left\{\max\left\{\Pr\left(\mathbf{y}_{\mathrm{j}}=\mathbf{0}\right)-\Pr\left(\mathbf{y}_{\mathrm{k}}=\mathbf{0}\right),0\right\},\ \max\left\{\Pr\left(\mathbf{y}_{\mathrm{k}}=\mathbf{1}\right)-\Pr\left(\mathbf{y}_{\mathrm{j}}=\mathbf{1}\right),0\right\}\right\}. \quad (38)$$

For $\mathrm{UB}\left(\mathcal{P}_{\mathrm{j}\succ\mathrm{k},\mathrm{C4}}\right)$ reference to exhibit 3 suggests an alternative approach that generally gives a bound at least as tight if not tighter than Boole's generic UB, specifically:[31]

$$\mathrm{UB}\left(\mathcal{P}_{\mathrm{j}\succ\mathrm{k},\mathrm{C4}}\right)=\min\left\{\Pr\left(\mathbf{y}_{\mathrm{j}}=\mathbf{0}\right)+\Pr\left(\mathbf{y}_{\mathrm{k}}=\mathbf{1}\right),\ 1-\Pr\left(\mathbf{y}_{\mathrm{j}}=\mathbf{1}\right),\ 1-\Pr\left(\mathbf{y}_{\mathrm{k}}=\mathbf{0}\right)\right\}. \quad (39)$$

Each term in braces in (39) is a legitimate UB, so the best UB is the smallest of the three.

For C2-C5 note that only the probabilities involving the joint marginal distributions evaluated at **0** and/or **1** are required to define the bounds since in each instance the treatment-preference characterization is binary (C2 and C3) or based on two binary characterizations (C4 and C5). For C6 the relevant probabilities in (36) and (37) are the sums of the probabilities of the elements of the particular Z and Z^{c} definitions being used, i.e. $\Pr\left(\mathbf{y}_{\bullet}\in\mathrm{Z}\right)=\sum_{\mathbf{q}\in\mathrm{Z}}\Pr\left(\mathbf{y}_{\bullet}=\mathbf{q}\right)$.

Computation is more complicated for C0 and C1, for which each treatment-preference probability is the sum of the probabilities of disjoint events. Unlike C2-C6 wherein "good" and "bad" outcomes are defined by particular values of the $\mathbf{y}_{\bullet}$, C0 and C1 rely on comparisons across all $2^{2\mathrm{M}}$ values of $\mathbf{Y}$ to determine treatment preference; as noted earlier, $2^{2\mathrm{M}-1}-\frac{(2\mathrm{M}-1)!}{\mathrm{M}!(\mathrm{M}-1)!}$ values of $\mathbf{Y}$ correspond to $\mathcal{P}_{\mathrm{j}\succ\mathrm{k},\mathrm{C0}}$, while $3^{\mathrm{M}}-2^{\mathrm{M}}$ values of $\mathbf{Y}$ correspond to $\mathcal{P}_{\mathrm{j}\succ\mathrm{k},\mathrm{C1}}$. Bounding these probabilities entails computing the relevant upper and lower bounds at each such $\mathbf{Y}$ value and then aggregating those bounds in some manner across relevant values of $\mathbf{Y}$. As discussed by Boole, 1854

(cont.)

$$\mathrm{UB}\left(\Pr\left(\mathrm{e}_1\vee\ldots\vee\mathrm{e}_{\mathrm{N}}\right)\right)\ =\ \min\left\{\sum_{\mathrm{n}=1}^{\mathrm{N}}\Pr\left(\mathrm{e}_{\mathrm{n}}\right),1\right\}\ \geq\ \Pr\left(\mathrm{e}_1\vee\ldots\vee\mathrm{e}_{\mathrm{N}}\right)\ \geq$$
$$\max\left\{\Pr\left(\mathrm{e}_1\right),\ldots,\Pr\left(\mathrm{e}_{\mathrm{N}}\right)\right\}\ =\ \mathrm{LB}\left(\Pr\left(\mathrm{e}_1\vee\ldots\vee\mathrm{e}_{\mathrm{N}}\right)\right)$$

[31] In the notation of exhibit 3, the set whose probability is to be bounded is $\mathbb{Y}_{0,\sim}\bigcup\mathbb{Y}_{0,1}\bigcup\mathbb{Y}_{\sim,1}$. The expression in (39) thus translates to

$$\min\left\{\Pr\left(\mathbb{Y}_{0,0}\bigcup\mathbb{Y}_{0,\sim}\bigcup\mathbb{Y}_{0,1}\right)\ +\ \Pr\left(\mathbb{Y}_{0,1}\bigcup\mathbb{Y}_{\sim,1}\bigcup\mathbb{Y}_{1,1}\right),\ \Pr\left(\mathbb{Y}_{0,0}\bigcup\mathbb{Y}_{\sim,0}\bigcup\mathbb{Y}_{0,\sim}\bigcup\mathbb{Y}_{\sim,\sim}\bigcup\mathbb{Y}_{0,1}\bigcup\mathbb{Y}_{\sim,1}\right),\right.$$
$$\left.\Pr\left(\mathbb{Y}_{0,\sim}\bigcup\mathbb{Y}_{\sim,\sim}\bigcup\mathbb{Y}_{1,\sim}\bigcup\mathbb{Y}_{0,1}\bigcup\mathbb{Y}_{\sim,1}\bigcup\mathbb{Y}_{1,1}\right)\right\}$$

(chapter XIX), to avoid double-counting aggregation is more complicated than simply adding up the respective upper and lower bounds of the $\Pr\left(\mathbf{Y}\right)$ across the relevant $\mathbf{Y}$. Although the C0 and C1 bounds are computable, simple representation as in (30)-(39) is not possible. See Appendix A.

For C0 and C1 it should also be noted that computation of the bounds using the joint marginal distributions at the relevant $\mathbf{y}_\bullet$ values requires that each of these joint marginal probabilities be identified by the available data. While in principle not a concern for identification, empirical implementation may encounter sparse data where some—or, in the case of large M, possibly many—of the joint outcomes are not observed in the sample. Empirically these would be zero-probability events by method-of-moments or analog principles (see section 5).[32]

Finally, with reference to (26) note that a legitimate LB on any of the terms in the inequality chain will be a legitimate LB on any terms leftward of that term, whereas a legitimate UB on any of the terms in the inequality chain will be a legitimate UB on any terms rightward of that term. In some instances this result may be helpful as a computational shortcut.

Point and Set Identification of the $PTE_{j\succ k,C\bullet}$

Once the bounds on $\mathcal{P}_{j\succ k,C\bullet}$ and $\mathcal{P}_{k\succ j,C\bullet}$ are defined for any of the $C\bullet$ it is straightforward to obtain bounds on the corresponding PTEs as:

$$\mathrm{UB}\left(\mathrm{PTE}_{j\succ k,C\bullet}\right)=\mathrm{UB}\left(\mathcal{P}_{j\succ k,C\bullet}\right)-\mathrm{LB}\left(\mathcal{P}_{k\succ j,C\bullet}\right) \tag{40}$$

and

$$\mathrm{LB}\left(\mathrm{PTE}_{j\succ k,C\bullet}\right)=\mathrm{LB}\left(\mathcal{P}_{j\succ k,C\bullet}\right)-\mathrm{UB}\left(\mathcal{P}_{k\succ j,C\bullet}\right). \tag{41}$$

While (40) and (41) hold generally, for C2, C3, and C6 point identification of $\mathrm{PTE}_{j\succ k,C\bullet}$ is possible given knowledge of $\Pr\left(\mathbf{y}_\bullet\right)$ at $\mathbf{0}$ (C2), at $\mathbf{1}$ (C3), or over Z (C6). Referring to exhibit 1, note that for C2 $1\left(\mathbf{y}_\bullet\neq\mathbf{0}\right)$ plays the same role as does $\mathrm{y}_\bullet$ when M=1, for C3 $1\left(\mathbf{y}_\bullet=\mathbf{1}\right)$ plays that role, and for C6 $1\left(\mathbf{y}_\bullet\in\mathrm{Z}^{\mathrm{c}}\right)$ plays that role. Thus, analogous to $\mathrm{E}\left[\mathrm{y}_\mathrm{k}-\mathrm{y}_\mathrm{j}\right]$ in (5) one has for C2

[32] Actual computation of the UBs and LBs may result in bounds that do not obey the ordering relationships in (26). While the respective probabilities must obey the ordering in (26), formulae used to compute those bounds are not necessarily so ordered since in some cases there are multiple legitimate ways to compute the bounds. Obtaining tightest bounds in such cases would require a search across the set of legitimate bounds; such considerations are beyond this paper's scope.

$$
\begin{aligned}
\mathrm{E}\Big[1\big(\mathbf{y}_k \neq \mathbf{0}\big) - 1\big(\mathbf{y}_j \neq \mathbf{0}\big)\Big] &= \mathrm{E}\Big[\big(1 - 1\big(\mathbf{y}_k = \mathbf{0}\big)\big) - \big(1 - 1\big(\mathbf{y}_j = \mathbf{0}\big)\big)\Big] \\
&= \mathrm{E}\Big[1\big(\mathbf{y}_j = \mathbf{0}\big) - 1\big(\mathbf{y}_k = \mathbf{0}\big)\Big] \\
&= \Big(\Pr\big(\mathbf{y}_j = \mathbf{0} \wedge \mathbf{y}_k = \mathbf{0}\big) + \Pr\big(\mathbf{y}_j = \mathbf{0} \wedge \mathbf{y}_k \neq \mathbf{0}\big)\Big) - \\
&\qquad \Big(\Pr\big(\mathbf{y}_j = \mathbf{0} \wedge \mathbf{y}_k = \mathbf{0}\big) + \Pr\big(\mathbf{y}_j \neq \mathbf{0} \wedge \mathbf{y}_k = \mathbf{0}\big)\Big) \\
&= \Pr\big(\mathbf{y}_j = \mathbf{0} \wedge \mathbf{y}_k \neq \mathbf{0}\big) - \Pr\big(\mathbf{y}_j \neq \mathbf{0} \wedge \mathbf{y}_k = \mathbf{0}\big) \\
&= \Pr\big(\mathrm{T}_j \succ_{\mathrm{C2}} \mathrm{T}_k\big) - \Pr\big(\mathrm{T}_k \succ_{\mathrm{C2}} \mathrm{T}_j\big) \;=\; \mathrm{PTE}_{j\succ k,\mathrm{C2}}
\end{aligned}
\tag{42}
$$

while symmetrically for C3 and C6 (skipping the middle steps used in (42)):

$$
\mathrm{E}\Big[1\big(\mathbf{y}_k = \mathbf{1}\big) - 1\big(\mathbf{y}_j = \mathbf{1}\big)\Big] \;=\; \Pr\big(\mathrm{T}_j \succ_{\mathrm{C3}} \mathrm{T}_k\big) - \Pr\big(\mathrm{T}_k \succ_{\mathrm{C3}} \mathrm{T}_j\big) \;=\; \mathrm{PTE}_{j\succ k,\mathrm{C3}} \tag{43}
$$

and

$$
\mathrm{E}\Big[1\big(\mathbf{y}_k \in \mathrm{Z}^c\big) - 1\big(\mathbf{y}_j \in \mathrm{Z}^c\big)\Big] \;=\; \Pr\big(\mathrm{T}_j \succ_{\mathrm{C6(Z)}} \mathrm{T}_k\big) - \Pr\big(\mathrm{T}_k \succ_{\mathrm{C6(Z)}} \mathrm{T}_j\big) \;=\; \mathrm{PTE}_{j\succ k,\mathrm{C6(Z)}} \tag{44}
$$

Thus even though their component probabilities $\mathcal{P}_{j\succ k,\mathrm{C}\bullet}$ and $\mathcal{P}_{k\succ j,\mathrm{C}\bullet}$ are not separately point identified, the $\mathrm{PTE}_{j\succ k,\mathrm{C}\bullet}$ are point identified for C2, C3, and C6. This is because their outcome-probability structures are essentially the same binary one seen in exhibit 1. It does not appear that the $\mathrm{PTE}_{j\succ k,\mathrm{C}\bullet}$ can be point identified for C0, C1, C4, and C5[33] so set identification must suffice.[34]

Signing $PTE_{j\succ k,C\bullet}$

Among the questions of concern to a decisionmaker, a prominent one might be the sign of a particular $\mathrm{PTE}_{j\succ k,\mathrm{C}\bullet}$ and whether the data are sufficiently informative to determine that one

---

[33] $\mathrm{PTE}_{j\succ k,\mathrm{C4}} + \mathrm{PTE}_{j\succ k,\mathrm{C5}}$ is point identified since $\mathrm{PTE}_{j\succ k,\mathrm{C4}} + \mathrm{PTE}_{j\succ k,\mathrm{C5}} = \mathrm{PTE}_{j\succ k,\mathrm{C2}} + \mathrm{PTE}_{j\succ k,\mathrm{C3}}$. However it is not obvious that such a quantity is likely to be of much interest.

[34] Based on (2), a different TE one might consider with outcomes s_j and s_k is $\mathrm{s}_k - \mathrm{s}_j$. With unconfounded treatment assignment the ATE $\mathrm{E}\big[\mathrm{s}_k - \mathrm{s}_j\big]$ is identifiable. However, since it is defined by s_j and s_k this ATE elicits the same concerns about across-outcome measurement comparability and (implied) weighting as does $\mathrm{PTE}_{j\succ k,\mathrm{C0}}$. These notwithstanding, estimation of regression models $\mathrm{E}\big[\mathrm{s}_{\mathbf{y}} \big| \mathbf{x} = \mathbf{x}_j\big]$ and $\mathrm{E}\big[\mathrm{s}_{\mathbf{y}} \big| \mathbf{x} = \mathbf{x}_k\big]$ as in (2) can be used to consistently estimate $\mathrm{E}\big[\mathrm{s}_k - \mathrm{s}_j\big]$.

treatment is unambiguously superior or not inferior to another. With PTE point identification (C2, C3, and C6) this is trivial. When only set identification is possible, knowing that $\text{LB}\left(\text{PTE}_{\text{j}\succ\text{k,C}\bullet}\right) > 0$ or $\text{UB}\left(\text{PTE}_{\text{j}\succ\text{k,C}\bullet}\right) < 0$ suffices to sign that PTE, up to sampling error.

Two figures illustrate what is involved for some specific cases. Figure 1a shows for the case of C2 the gains from point versus set identification when the question of interest concerns the sign of $\text{PTE}_{\text{j}\succ\text{k,C2}}$. In the figure, all combinations of $\Pr\left(\mathbf{y}_{\text{j}} = \mathbf{0}\right)$ and $\Pr\left(\mathbf{y}_{\text{k}} = \mathbf{0}\right)$ below the 45-degree line are consistent with $\text{PTE}_{\text{j}\succ\text{k,C2}} \geq 0$. Alternatively, determining whether $\text{PTE}_{\text{j}\succ\text{k,C2}} \geq 0$ by reference to whether $\text{LB}\left(\text{PTE}_{\text{j}\succ\text{k,C2}}\right) \geq 0$ relies on only those combinations of the $\Pr\left(\mathbf{y}_{\bullet} = \mathbf{0}\right)$ that figure into the LB computation in (31). These combinations are shown in the darker shaded area.

Figure 1b depicts the combinations of the $\Pr\left(\mathbf{y}_{\bullet} = \mathbf{0}\right)$ consistent with $\text{LB}\left(\text{PTE}_{\text{j}\succ\text{k,C5}}\right) = 0$, using $\text{LB}\left(\text{PTE}_{\text{j}\succ\text{k,C2}}\right)$ from figure 1a as a baseline reference. Since the C5 bounds involve the marginal probabilities at both **0** and **1**, the picture is drawn holding the $\Pr\left(\mathbf{y}_{\bullet} = \mathbf{1}\right)$ at specific values (shown in the figure's legend) and then tracing out the $\Pr\left(\mathbf{y}_{\bullet} = \mathbf{0}\right)$ combinations consistent with $\text{LB}\left(\text{PTE}_{\text{j}\succ\text{k,C5}}\right) = 0$ at those values. Combinations of the $\Pr\left(\mathbf{y}_{\bullet} = \mathbf{0}\right)$ southeast of the positively-sloped line segments are ones where the C2 and C5 $\text{LB}\left(\text{PTE}_{\text{j}\succ\text{k,C}\bullet}\right)$ are strictly positive.

[figures 1a and 1b about here]

A Parametric Example

To see how the bounds described above perform numerically, true probabilities and corresponding bounds are computed under several different assumptions about the degree of cross-component correlation of the elements of $\mathbf{Y}$ and about the PTE magnitudes. The calculations assume that the $\mathbf{y}_{\bullet}$ have elements $\text{y}_{\bullet,\text{m}} = 1\left(\text{y}^{*}_{\bullet,\text{m}} > 0\right)$ with $\mathbf{Y}^{*} \sim \text{MVN}\left(\left[\boldsymbol{\mu}_{\text{j}}, \boldsymbol{\mu}_{\text{k}}\right], \mathbf{R}\right)$. For all m $\mu_{\text{j,m}} = \Phi^{-1}\left(.1\right)$, so that $\Pr\left(\text{y}_{\text{j,m}} = 1\right) = .1$. $\mu_{\text{k,m}} = \Phi^{-1}\left(\Pr\left(\text{y}_{\text{k,m}} = 1\right)\right)$, with $\Pr\left(\text{y}_{\text{k,m}} = 1\right) = .2$ and $\Pr\left(\text{y}_{\text{k,m}} = 1\right) = .5$ giving "small" and "large" TEs, respectively. $\mathbf{R}$ is a $2\text{M} \times 2\text{M}$ correlation matrix with all off-diagonal elements equal to ρ, which is either 0 or .5 in this exercise. The results for M=2, M=3, and M=4 are displayed in tables 2a-2c.[35] In each cell appears the true joint probability (which would be unknowable from observable data) as well as the corresponding LB and UB (which under unconfoundedness would be identifiable in applications from the joint marginal

[35] The probabilities are generated using Mata's *mvnormal* simulator in Stata's version 15.1.

probabilities) or, in the case of the $\text{PTE}_{\text{j}\succ\text{k,C}\bullet}$ for C2 and C3, the point-identified true values.

[tables 2a, 2b, 2c about here]

Several results are noteworthy. First, as a consequence of the particular parameters specified for this exercise the $\text{LB}\left(\mathcal{P}_{\text{j}\succ\text{k,C}\bullet}\right)$ and $\text{LB}\left(\mathcal{P}_{\text{k}\succ\text{j,C}\bullet}\right)$ are the same for C0, C1, C2, and C4. Also, in light of the discussion in the previous subsection, note that in some instances the $\text{LB}\left(\text{PTE}_{\text{j}\succ\text{k,C}\bullet}\right)$ exceed zero so that one can conclude that the PTE is positive even though its specific magnitude is not identifiable. Finally, unlike the probability orderings among the $\mathcal{P}_{\text{j}\succ\text{k,C}\bullet}$ and $\mathcal{P}_{\text{k}\succ\text{j,C}\bullet}$ from (26), there is no such necessary ordering among the $\text{PTE}_{\text{j}\succ\text{k,C}\bullet}$. So while all the results in tables 2a-2c numerically satisfy

$$\text{PTE}_{\text{j}\succ\text{k,C0}} \;\geq\; \text{PTE}_{\text{j}\succ\text{k,C1}} \;\geq\; \text{PTE}_{\text{j}\succ\text{k,C4}} \;\geq\; \left\{ \begin{array}{c} \text{PTE}_{\text{j}\succ\text{k,C2}} \\ \text{PTE}_{\text{j}\succ\text{k,C3}} \end{array} \right\} \;\geq\; \text{PTE}_{\text{j}\succ\text{k,C5}}\,, \tag{45}$$

this is not a general result but owes rather to the particular probability structures assumed here.

5. Estimating Bounds using Composite Outcomes Measures

What Are Composite Outcomes?

Composite outcomes or endpoints are used widely as health status measures in clinical evaluations, and are particularly prominent in studies involving cardiovascular disease outcomes. For example, in a recent three-arm clinical trial comparing cardiovascular health effects of different Mediterranean diets, the primary outcome studied by Estruch et al., 2018, is "a composite of myocardial infarction, stroke, and death from cardiovascular causes." Occurrence of any or all of the three outcomes over the study period indicates treatment failure while experiencing none of the three implies treatment success. While particular components vary across studies, the Estruch et al. approach is typical. Indeed composite-outcome measures are used broadly in clinical and social science research even though such measures might not actually be dubbed composite outcomes in a particular study's report. For instance, a standard measure of chronic obstructive pulmonary disease (COPD) is the presence of emphysema and/or chronic bronchitis. While this corresponds formally to a composite outcome, COPD is often not explicitly referred to as such.

One might define a composite outcome in various ways (U.S. FDA, 2017). In a typical

application there is a set[36] of M>1, often binary, components outcomes across which the composite outcome is deemed to be a success or represent a "good" outcome only when all of its components are "good" outcomes. Defined thusly, composite outcomes have an all-or-nothing character.

As is standard, let the observed outcome data be determined by the potential outcomes and the assigned treatment as

$$\mathbf{y} = 1\left(\mathbf{x} = \mathbf{x}_{\mathrm{j}}\right) \times \mathbf{y}_{\mathrm{j}} + 1\left(\mathbf{x} = \mathbf{x}_{\mathrm{k}}\right) \times \mathbf{y}_{\mathrm{k}}, \tag{46}$$

where $\mathbf{x}$ denotes a k-vector of exogenous covariates characterizing treatment that will take on one of two possible values, $\mathbf{x}_{\mathrm{j}}$ or $\mathbf{x}_{\mathrm{k}}$, corresponding to the treatments T_{j} and T_{k}. For immediate purposes it suffices to consider a generic scalar composite outcome $1\left(\mathbf{y} \in \mathrm{Z}^{\mathrm{c}}\right)$ where as above $\mathrm{Z} \subset \mathrm{Q}$ is a set containing particular values of $\mathbf{y}$ that correspond to a "good" outcome.

Using Composite Model Estimates to Compute Bounds and Point-Identified PTEs

Recall from section 1 and eq. (4) that one general approach to estimation in the presence of multiple outcomes is to define some measure that effectively collapses an M-dimension outcome into a one-dimension outcome. This common empirical strategy is relevant for purposes at hand. Specifically in this case one specifies and estimates a parametric or nonparametric composite-outcome conditional probability model

$$\Pr\left(\mathbf{y} \in \mathrm{Z} \middle| \mathbf{x}\right) = \mathrm{p}_{\mathrm{c}}\left(\mathbf{x}\right). \tag{47}$$

With unconfounded treatment assignment (exogenous $\mathbf{x}$), estimation of (47) yields

$$\widehat{\Pr}\left(\mathbf{y} \in \mathrm{Z} \middle| \mathbf{x} = \mathbf{x}_{\bullet}\right) \rightarrow \Pr\left(\mathbf{y}_{\bullet} \in \mathrm{Z}\right) \tag{48}$$

Recall the bounds for treatment-preference characterizations C2-C5 derived in section 4. For these characterizations the corresponding UBs and LBs are defined in terms of the estimands in (48), each for $\mathrm{Z} = \left\{\mathbf{0}\right\}$ and/or $\mathrm{Z}^{\mathrm{c}} = \left\{\mathbf{1}\right\}$. As such, for C2-C5 standard estimation of particular composite-outcome models yields the information required to estimate the bounds on $\mathcal{P}_{\mathrm{j}\succ\mathrm{k},\mathrm{C}\bullet}$ and $\mathrm{PTE}_{\mathrm{j}\succ\mathrm{k},\mathrm{C}\bullet}$

[36] While M=3 or M=4 are common, there are interesting cases where M is as small as two or much larger than four.

so long as treatments are assigned exogenously. Such estimates also provide the information to point identify the PTEs for C2 and C3 as described in section 4. (Note that the composite-outcome PTEs for a "good" outcome $\mathrm{Z}=\left\{\mathbf{0}\right\}$ (i.e. C2) and a "bad" outcome $\mathrm{Z}^{\mathrm{c}}=\left\{\mathbf{1}\right\}$ (i.e. C3) correspond to the boldface entries in tables 2a-2c.) While such composite-outcome model estimates may be of interest in their own right in capturing some notions of "treatment effect"—perhaps explaining their prominence in applications—a previously unappreciated attribute is that they provide information essential to point identify or set identify PTEs of the sort proposed here.

For C0 and C1, the same basic ideas apply although a potentially large number of Z definitions may be required to estimate the components of the bounds (see section 4). Beyond the algebraic complexities involved in computation of these bounds, a practical issue is whether the available data are sufficiently rich to yield useful estimates $\widehat{\Pr}\left(\mathbf{y}\in\mathrm{Z}\middle|\mathbf{x}=\mathbf{x}_{\bullet}\right)$ for *each* Z defining an estimand (48) that is, in turn, involved in such computations. In particular, note that the estimated probabilities associated with particular Z that are not represented in the available data equal zero by method-of-moments or analog principles.[37]

6. Two Empirical Examples

Moving to Opportunity

Data from the prominent Moving to Opportunity (MTO) experiment provide an illustrative

---

[37] While inference is not a main concern here, it might be noted that for the point-identified PTEs nonparametric inference is straightforward. Letting $\mathcal{I}_{\bullet,\bullet}$ be observation index sets defined in an obvious manner, analog PTE estimates for C2 ($\mathrm{Z}=\left\{\mathbf{0}\right\}$) and C3 ($\mathrm{Z}^{\mathrm{c}}=\left\{\mathbf{1}\right\}$) are given by

$$\begin{aligned}\widehat{\mathrm{PTE}}_{\mathrm{j}\succ\mathrm{k},\mathrm{C}\bullet} &= \tfrac{1}{\#\mathcal{I}_{2,\mathrm{k}}}\textstyle\sum_{\mathrm{n}\in\mathcal{I}_{2,\mathrm{k}}}1\left(\mathbf{y}_{\mathrm{n}}\in\mathrm{Z}^{\mathrm{c}}\right)\;-\;\tfrac{1}{\#\mathcal{I}_{2,\mathrm{j}}}\textstyle\sum_{\mathrm{n}\in\mathcal{I}_{2,\mathrm{j}}}1\left(\mathbf{y}_{\mathrm{n}}\in\mathrm{Z}^{\mathrm{c}}\right)\\ &= \widehat{\Pr}\left(\mathbf{y}\in\mathrm{Z}^{\mathrm{c}}\middle|\mathbf{x}=\mathbf{x}_{\mathrm{k}}\right)-\widehat{\Pr}\left(\mathbf{y}\in\mathrm{Z}^{\mathrm{c}}\middle|\mathbf{x}=\mathbf{x}_{\mathrm{j}}\right)=\;\mathrm{p}_{\bullet,\mathrm{k}}-\mathrm{p}_{\bullet,\mathrm{j}}\end{aligned}$$

with the corresponding binomial variance estimates given by

$$\widehat{\mathrm{var}}\left(\widehat{\mathrm{PTE}}_{\mathrm{j}\succ\mathrm{k},\mathrm{C}\bullet}\right)=\frac{\mathrm{p}_{\bullet,\mathrm{j}}\left(1-\mathrm{p}_{\bullet,\mathrm{j}}\right)}{\#\mathcal{I}_{\bullet,\mathrm{j}}}+\frac{\mathrm{p}_{\bullet,\mathrm{k}}\left(1-\mathrm{p}_{\bullet,\mathrm{k}}\right)}{\#\mathcal{I}_{\bullet,\mathrm{k}}}.$$

Large-sample results can be used to compute CIs from these $\widehat{\mathrm{var}}\left(...\right)$. With set identification inference is also possible but more complicated (see Imbens and Manski, 2004, and Chernozhukov et al., 2013). Imbens and Manski, 2004, note that "researchers face a substantive choice whether to report intervals that cover the entire identification region or intervals that cover the true parameter value with some fixed probability…Which CI is of interest depends on the application."

example. These data and the experiment's results are reported in Ludwig et al., 2011 and 2013. The experiment consisted of two intervention groups and a control group, but for simplicity only the low-poverty voucher and control groups are considered here. The public-use "pseudo-individual" sample, consisting of N=3,273 observations and described at www.nber.org/mtopuf, is used here. After deleting observations with missing data the remaining sample contains N=2,120 subjects, N=1,178 in the intervention group (T_j) and N=942 in the control group (T_k).

Two exercises are conducted here. In the first, M=2 with binary outcomes obesity (BMI≥40) and diabetes (HbA1c≥6.5); these were the outcomes considered in Ludwig et al., 2011. In the second, M=4 with binary outcomes obesity, diabetes, hypertension (SBP≥140 and/or DBP≥90), and depression (DSM-IV major depressive episode in the past year); this is a subset of the outcomes considered in Ludwig et al., 2013. The results are summarized in table 3a. For these data the bounds on the set-identified PTEs are seen to be rather broad, and in no instance unambiguously informative about the sign of any of the PTEs. For C2 the point-identified $\text{PTE}_{j\succ k,C2}$ indicate a positive (i.e. beneficial) effect of the intervention for both the M=2 and M=4 outcome definitions, whereas for C3 the results of the treatment are less clear.

[table 3a about here]

Multiple Chronic Conditions

A sample of N=887,309 adults ages 18-64 from the 2011, 2013, and 2015 Behavioral Risk Factors Surveillance System (BRFSS) surveys is used to explore the determinants of adult chronic-condition outcomes. The observed outcomes $\mathbf{y}$ are seven binary chronic-condition indicators: cardiovascular disease, arthritis, depression, chronic lower-respiratory disease, cancer, diabetes, and kidney disease. For this exercise an age "treatment" and a schooling "treatment" are considered separately (for purposes of this brief illustrative exercise the paper won't dwell on whether unconfoundedness is reasonable here). For age, T_j and T_k correspond to ages 18-44 and ages 45-64; for schooling, T_j and T_k correspond to not being versus being a college graduate.

The estimated bounds are reported in table 3b. The age PTEs for C2-C4 all suggest unambiguously younger age as the preferred treatment, while for C0 and C1 the width of the PTE bounds interval exceeds one. For the schooling PTEs, all the bound intervals straddle zero, while the point-identified results for C2 and C3 suggest college graduation as the preferred treatment.[38]

[38] Stata code and data used to generate the results in tables 2a-3b and figure 1 are available in a 1.6MB .zip file, https://uwmadison.box.com/temo.zip. The readme file in the main directory provides details.

[table 3b about here]

7. Treatment-Preference Characterizations, Probabilities, and PTEs with Multiple Ordered Outcomes

Instead of binary outcome measurement suppose each component of the $\mathbf{y}_\bullet$ is measured in an ordered, categorical manner. That is, each of the M components can assume one of G possible values in $\{0,1,\ldots,G-1\}$.[39] In keeping with the ordering used previously, larger values of the components of the $\mathbf{y}_\bullet$ correspond to increasingly undesirable outcomes. One prominent example in a health-outcome context is the EQ-5D system describing M=5 dimensions of health: mobility, self-care, usual activities, pain/discomfort and anxiety/depression (Devlin et al., 2018). The specific dimensions are measured in G=3 (EQ-5D-3L) or G=5 (EQ-5D-5L) ordered levels. Among other uses, EQ-5D data provide the foundation for a variety of health-related quality of life measures. Determining values associated with the $3^5 = 243$ (3L) or $5^5 = 3125$ (5L) possible health states described by EQ-5D is a major aspect of EQ-5D-related research.[40]

Treatment-Preference Characterizations

Analysis of treatment-preference and PTEs for ordered outcomes can proceed along essentially the same lines as with binary outcomes, albeit with a few additional considerations. Let $C\bullet^*$ denote characterizations of treatment preference relevant in a multiple ordered-outcome context. While various characterizations might be proposed only variants of the $C\bullet$ characterizations already examined are considered here. Specifically let C0*, C1*, C2*, and C6* correspond exactly to C0, C1, C2, and C6 as defined earlier. C3 is redefined to be C3* wherein $T_j \succ_{C3^*} T_k$ if and only if $\left(\mathbf{y}_j \neq \left(\mathbf{G}-\mathbf{1}\right)\right) \wedge \left(\mathbf{y}_k = \left(\mathbf{G}-\mathbf{1}\right)\right)$, where $\mathbf{G}$ is an M-vector whose elements all equal G. That is, in C3* a "good" outcome is one where not all components are at their worst possible levels while a "bad" outcome is one where each component is at its worst possible level. C4* is defined such that $T_j \succ_{C4^*} T_k$ if and only if $\left(\mathbf{y}_j = \mathbf{0} \wedge \mathbf{y}_k \neq \mathbf{0}\right) \vee \left(\mathbf{y}_j \neq \left(\mathbf{G}-\mathbf{1}\right) \wedge \mathbf{y}_k = \left(\mathbf{G}-\mathbf{1}\right)\right)$, and C5* is defined such that $T_j \succ_{C5^*} T_k$ if and only if $\mathbf{y}_j = \mathbf{0} \wedge \mathbf{y}_k = \left(\mathbf{G}-\mathbf{1}\right)$. The sets $\mathbb{Y}_{j\succ k, C\bullet^*}$ and

[39] Of course the binary case considered heretofore is just the special case G=2. Only for notational simplicity is it assumed that each component assumes the same number of possible categorical values; the logic of what follows in no way relies on this.

[40] Another example is the Apgar score, used for assessment of neonates' health (American Academy of Pediatrics, 2006; Hoynes et al., 2015). Apgar scores use M=5 components measured across G=3 categories with larger values representing better health.

$\mathbb{Y}_{k\succ j,C\bullet *}$ are defined in an obvious manner consistent with their definitions in section 3.

Apart from C0* these characterizations do not require across-component measurement comparability for coherence. That is, any particular value among the G categories of (say) the m-th and m'-th components of $\mathbf{y}_j$ and $\mathbf{y}_k$ need not represent equivalence between those components. Since C1*-C6* rely only on element-by-element comparisons *between* each of the M components of the $\mathbf{y}_j$ and $\mathbf{y}_k$ but not on comparisons *across* the M components of each $\mathbf{y}_\bullet$, the M components' G categories may be measured in any manner deemed relevant. For C1*-C6* each component's measure can be changed from $\{0,\ldots,G-1\}$ to different values $\{v_{0,m},\ldots,v_{G-1,m}\}$; the same probability and PTE results obtain if the ordering of the $v_\bullet$ respects the $\{0,\ldots,G-1\}$ ordering and if references to 0 and G-1 in the C2*-C5* definitions become references to $v_{0,m}$ and $v_{G-1,m}$.

With C0*, however, the sums s_j and s_k entail some degree of cross-component comparability just as when the the $\mathbf{y}_\bullet$ are binary. Whether it is reasonable to endow what are essentially ordered component outcomes with interval- or ratio- scale properties will depend on an application's particulars. In any event the summands defining s_j and s_k are typically apples and oranges. While it may sometims be frowned upon, computing indexes, scores, or scales from sums of unlike components—whether the components are binary or ordered—is longstanding practice in psychometric and clinical settings. Without defending this approach it is discussed here as a benchmark because of its prominence in such fields: despite its apples-oranges character the approach is used widely and sometimes rationalized explicitly.[41]

For concreteness table 4 provides an example of the $C\bullet *$ characterizations for M=2 and G=3. There are $G^{2M}=81$ possible values of the 1×4 vector of potential outcomes $\mathbf{Y}$. Of these 81 possible values, 31 correspond to values of $\mathbf{Y}$ for which $T_j \succ_{C\bullet *} T_k$ for at least one of the $C\bullet *$.

[table 4 about here]

Bounding $\mathcal{P}_{j\succ k,C\bullet *}$ and $PTE_{j\succ k,C\bullet *}$ with Multiple Ordered Outcomes

Once a particular $C\bullet *$ has been selected, conceptualizing $\mathcal{P}_{j\succ k,C\bullet *}$, $\mathcal{P}_{k\succ j,C\bullet *}$, and $PTE_{j\succ k,C\bullet *}$ and computing their corresponding bounds proceeds in essentially the same manner as in the binary case. The relevant joint probabilities are determined and the observed data are consulted to estimate the corresponding probability and PTE bounds. The same probability

[41] Kahneman, 2011, chapter 21, offers an interesting assessment of such measurement issues.

ordering as described in (26) obtains here for $\mathcal{P}_{j\succ k,C\bullet *}$ with reference to $C\bullet *$ instead of $C\bullet$.

As before identification depends on the assumption of unconfounded or exogenous treatment assignment. Analogous to the binary case, $\mathcal{P}_{j\succ k,C\bullet *}$ is point identified for C2*, C3*, and C6*. As G increases the number of possible $\mathbf{Y}$ outcomes grows rapidly. However, in deriving the relevant UBs and LBs (or point-identified PTEs), the concern is not with the entire $\mathbf{Y}$ vector but rather with the particular $\mathbf{y}_\bullet$ whose probabilities enter the bound definitions. As noted earlier, the relevant joint marginal probabilities may be technically identified even though small empirical cell sizes may be of concern regarding the data's ability to deliver useful estimates. To frame these issues, define the sets $\mathbb{J}_{j\succ k,C\bullet *}=\left\{\mathbf{y}_j\middle|\mathbf{y}_j\in\mathbb{Y}_{j\succ k,C\bullet *}\right\}$ and $\mathbb{K}_{j\succ k,C\bullet *}=\left\{\mathbf{y}_k\middle|\mathbf{y}_k\in\mathbb{Y}_{j\succ k,C\bullet *}\right\}$. Table 5 displays for several values of M and G the number of elements in the sets $\mathbb{Y}_{j\succ k,C\bullet *}$ and $\mathbb{J}_{j\succ k,C\bullet *}$. $\#\mathbb{J}_{j\succ k,C\bullet *}$ and, symmetrically, $\#\mathbb{K}_{j\succ k,C\bullet *}$ increase rapidly with G and with M.

[table 5 about here]

Some Implications for Multiple Continuous-Outcome Measures

Suppose the component outcomes in the $\mathbf{y}_\bullet$ are measured continuously. Such measurements can be coarsened into G ordered, categorical measures, sometimes known as interval measurement:

$$y_{j,m}^{coarse}=\left\{\begin{array}{c}0\\ \vdots\\ G-1\end{array}\right\}\ \leftrightarrow\ y_{j,m}\in\left\{\begin{array}{c}\left(-\infty,\ t_{0,m}\right]\\ \vdots\\ \left(t_{(G-2),m},\ \infty\right)\end{array}\right\}, \tag{49}$$

where the $t_{g,m}$ are component-specific thresholds or cut points.

Once coarsened, PTEs for the $y_{j,m}^{coarse}$ can be considered using the strategies discussed above. Selecting a useful degree of coarsening (G) involves trading off information loss against computational complexity. It might be noted that coarsened continuous measures are in fact used in multiple-outcome applications. Two examples are composite-outcome component failure times coarsened to binary N-month survival indicators, and continuously-measured allostatic load components coarsened to binary threshold-crossing indicators (Gruenwald et al., 2006).[42]

[42] Though bounds computation may be challenging, one might tackle directly analysis of continuous multiple outcomes in, say, a multivariate-normal framework. Using coarsened continuous outcomes (cont.)

8. Summary

This paper has suggested strategies for defining, identifying, and estimating treatment effects in contexts where understanding determinants of multiple outcomes is the goal. Notions of treatment preference and their corresponding probability structures and PTEs have been proposed as organizing principles within which questions regarding multiple outcomes might be integrated and pursued. While the paper has suggested seven characterizations of treatment preference appropriate to multiple-outcome contexts, more can be imagined and this should prove a useful research agenda. Regardless of the preference characterization chosen for a particular analysis, that choice should be made before the data are revealed: Scientific fairness forbids "preference mining."

In at least three other areas future research might prove valuable. First is the extension to the multiple-outcome case of methods to handle confounded or endogenous treatment assignment. Second is a consideration in multiple-outcome contexts of strategies to enhance the informativeness of set identification, i.e. tighten the bounds (e.g. Frandsen and Lefgren, 2018, Manski and Pepper, 2000). Third is the fundamental consideration of why in applications outcome vectors are specified in the manner they are, i.e. how do particular specifications of $\mathbf{Y}$ or $\mathbf{y}_\bullet$ more or less well describe the outcomes that actually matter to decisionmakers.

Returning in closing to a statement from section 1, it is hoped this paper has stimulated readers to assess with new perspectives their approaches to understanding multiple outcomes and their determinants. If the paper accomplishes only this, it will have served some valuable purpose.

Acknowledgments

I am grateful to Domenico Depalo, Paddy Gillespie, Hyunseung Kang, Chuck Manski, and participants in presentations at Queen's University Belfast, the University of Bergamo, the University of Chicago, and the University of Wisconsin-Madison for helpful comments, suggestions, and guidance on the literature. Partial support was provided by RWJF Evidence for Action Grant 73336 and by the UW-Madison Center for Demography and Ecology under NICHD grant P2CHD047873.

(cont.)
discards some information and can itself present computational challenges, but it is conceptually straightforward and readily accommodates nonparametric identification and estimation.

References

Abadie, A. and M.D. Cattaneo. 2018. "Econometric Methods for Program Evaluation." *Annual Review of Economics* 10: 465-503.

Alkire, S. and J. Foster. 2010. "Counting and Multidimensional Poverty Measurement." Journal of Public Economics 95: 476-487.

American Academy of Pediatrics, Committee on Fetus and Newborn; American College of Obstetricians and Gynecologists and Committee of Obstetric Practice. 2006. "The Apgar Score." *Pediatrics* 117: 1444-1447.

Angrist, J.D. and G.W. Imbens. 1995. "Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity." *Journal of the American Statistical Association*, 90: 431–442.

Athey, S. et al. 2016. "Estimating Treatment Effects using Multiple Surrogates: The Role of the Surrogate Score and the Surrogate Index." arXiv:1603.09326v2 [stat.ME].

Athey, S. and G.W. Imbens. 2017. "The Econometrics of Randomized Experiments." *Handbook of Economic Field Experiments* 1: 73-140.

Boole, G. 1854. *An Investigation of the Laws of Thought, of Which are Founded the Mathematical Theories of Logic and Probabilities*. London: Walton and Maberly.

Buttorff, C. et al. 2017. *Multiple Chronic Conditions in the United States*. RAND Corporation, Report TL-221-PFCD.

Cebul, R.D. et al. 2011. "Electronic Health Records and Quality of Diabetes Care." *NEJM* 365: 825-833.

Chernozhukov, V. et al. 2013. "Intersection Bounds: Estimation and Inference." *Econometrica* 81: 667-737.

Coleman-Jensen, A. et al. 2018. *Household Food Security in the United States in 2017*. Washington: USDA, Economic Research Service.

Devlin, N.J. et al. 2018. "Valuing Health-Related Quality of Life: An EQ-5D-5L Value Set for England." *Health Economics* 27: 7-22.

Dmitrienko, A. and R.B. D'Agostino. 2018. "Multiplicity Considerations in Clinical Trials." *NEJM* 378: 2115-2122.

Dodge, K.A. et al. 2014. "Impact of Early Intervention on Psychopathology, Crime, and Well-being at Age 25." *American Journal of Psychiatry* 172: 59-70.

Estruch, R. et al. 2018. "Primary Prevention of Cardiovascular Disease with a Mediterranean Diet Supplemented with Extra-Virgin Olive Oil or Nuts." *NEJM* 378: e34(1)-e34(14).

Färe, R. and L. Svensson. 1980. "Congestion of Production Factors." *Econometrica* 48: 1745-1753.

Fleishman, J.A. et al. 2014. "Screening for Depression Using the PHQ-2: Changes over Time in Conjunction with Mental Health Treatment." Agency for Healthcare Research and Quality

Working Paper No. 14002.

Frandsen, B.R. and L.J. Lefgren. 2018. "Partial Identification of the Distribution of Treatment Effects with an Application to the Knowledge Is Power Program (KIPP)." NBER w.p. 24616.

Geronimus, A.T. et al. 2006. "'Weathering' and Age Patterns of Allostatic Load Scores Among Blacks and Whites in the United States." *American Journal of Public Health* 96: 826-833.

Gruenwald, T.L. et al. 2006. "Combinations of Biomarkers Predictive of Later Life Mortality." *PNAS* 103: 14158–14163.

Hansen, B. 2018. *Econometrics* (January, 2018, version). www.ssc.wisc.edu/~bhansen/econometrics/

Hanssen, D.J. et al. 2014. "Physical, Lifestyle, Psychological, and Social Determinants of Pain Intensity, Pain Disability, and the Number of Pain Locations in Depressed Older Adults." *Pain* 155: 2088-2096.

Hoynes, H. et al. 2015. "Income, the Earned Income Tax Credit, and Infant Health." *AEJ Economic Policy* 7: 172-211.

Imbens, G.W. and C.F. Manski. 2004. "Confidence Intervals for Partially Identified Parameters." *Econometrica* 72: 1845-1857.

Imbens, G.W. and J.M. Wooldridge. 2009. "Recent Developments in the Econometrics of Program Evaluation." *Journal of Economic Literature* 47: 5-86.

Institute of Medicine. 2006. *Performance Measurement: Accelerating Improvement.* Washington: The National Academies Press.

Jackson, C.K. 2018. "What Do Test Scores Miss? The Importance of Teacher Effects on Non–Test Score Outcomes." *Journal of Political Economy* 126: 2072–2107.

Khan, N. et al. 2008. "Effect of prescription drug coverage on health of the elderly." *Health Services Research* 43: 1576-1597.

Kahneman, D. 2011. *Thinking, Fast and Slow.* New York: Farrar, Straus and Giroux.

Loeb, S., and D. Figlio. 2011. "School Accountability." in E. A. Hanushek et al., Eds. *Handbook of the Economics of Education*, Vol. 3. San Diego: North Holland.

Look AHEAD Research Group. 2013. "Cardiovascular Effects of Intensive Lifestyle Intervention in Type 2 Diabetes." *NEJM* 369:145-154.

Ludwig, J. et al. 2011. "Neighborhoods, Obesity, and Diabetes—A Randomized Social Experiment." *NEJM* 365: 1509-1519.

Ludwig, J. et al. 2013. "Long-Term Neighborhood Effects on Low-income Families: Evidence from Moving to Opportunity." *American Economic Review Papers and Proceedings* 103: 226-231.

Lynn, J. et al. 2015. "Value-Based Payments Require Valuing What Matters to Patients." *JAMA* 314: 1445-1446.

Manski, C.F. 2004. "Statistical Treatment Rules for Heterogeneous Populations." *Econometrica* 72:

1221–1246.

Manski, C.F. 2018. "Reasonable Patient Care under Uncertainty." *Health Economics* 27: 1397-1421.

Manski, C.F. and J.V. Pepper. 2000. "Monotone Instrumental Variables: With an Application to the Returns to Schooling." *Econometrica* 68: 997-1010.

Manski, C.F. and A. Tetenov. 2018. "Trial Size for Near-Optimal Choice between Surveillance and Aggressive Treatment: Reconsidering MSLT-II." Working Paper, Northwestern University.

Mullahy, J. 2018a. "Individual Results May Vary: Inequality-Probability Bounds for Some Health-Outcome Treatment Effects." *Journal of Health Economics* 61: 151-162 .

Mullahy, J. 2018b. "The Anatomy of Composite Health Outcomes." (in progress).

Nissen, S.E. et al. 2016. "Cardiovascular Safety of Celecoxib, Naproxen, or Ibuprofen for Arthritis." *NEJM* 375: 2519-2529.

Nolan, B. and C.T. Whelan. 2010. "Using Non-Monetary Deprivation Indicators to Analyze Poverty and Social Exclusion: Lessons from Europe?" *Journal of Policy Analysis and Management* 29: 305-325.

Parving, H.-H. et al. 2012. "Cardiorenal End Points in a Trial of Aliskiren for Type 2 Diabetes." *NEJM* 367: 2204-2213.

Pesko, M.F. et al. 2016. "The Effect of Potential Electronic Nicotine Delivery System Regulations on Nicotine Product Selection." *Addiction* 111: 734-744.

Porter, M.E. et al. 2015. "Standardizing Patient Outcome Measures." *NEJM* 374: 504-506.

Romano, J.P et al. 2010. "Multiple Testing." in *The New Palgrave Dictionary of Economics*, Online Edition. London: Palgrave Macmillan.

Rosenfield, K. et al. 2016. "Randomized Trial of Stent versus Surgery for Asymptomatic Carotid Stenosis." *NEJM* 374: 1011-1020.

Ruoff, C. and D. Rye. 2016. "The ICSD-3 and DSM-5 Guidelines for Diagnosing Narcolepsy: Clinical Relevance and Practicality." *Current Medical Research and Opinion* 32: 1611-1622.

Sarma, S. et al. 2015. "The Effect of Leisure-Time Physical Activity on Obesity, Diabetes, High BP and Heart Disease among Canadians: Evidence from 2000/2001 to 2005/2006." *Health Economics* 24: 1531-1547.

Shwartz, M. et al. 2015. "Composite Measures of Health Care Provider Performance: A Description of Approaches." *Milbank Quarterly* 93: 788-825.

Siebert, S. et al. 2016. "Characteristics of Rheumatoid Arthritis and Its Association with Major Comorbid Conditions: Cross-sectional Study of 502 649 UK Biobank Participants." *RMD Open* doi: 10.1136/rmdopen-2016-000267.

Stauffer, V.L. et al. 2018. "Evaluation of Galcanezumab for the Prevention of Episodic Migraine—The EVOLVE-1 Randomized Clinical Trial." *JAMA Neurology*: 75: 1080-1088.

Stone, G.W. et al. 2018. "Transcatheter Mitral-Valve Repair in Patients with Heart Failure."

NEJM DOI: 10.1056/NEJMoa1806640.

U.S. Federal Deposit Insurance Corporation. 1997. *Uniform Financial Institutions Rating System.* FDIC Statements of Policy. Washington: FDIC (Retrieved September 27, 2018, at www.fdic.gov/ regulations/laws/rules/5000-900.html).

U.S. Food and Drug Administration. 2017. *Multiple Endpoints in Clinical Trials—Guidance for Industry.* USDHHS/FDA CBER/CDER.

U.S. National Heart, Lung, and Blood Institute. 2018. "Metabolic Syndrome." www.nhlbi.nih.gov/ health-topics/metabolic-syndrome (retrieved Oct. 16, 2018).

Walitt, B. et al. 2015. "The Prevalence and Characteristics of Fibromyalgia in the 2012 National Health Interview Survey." *PLoS One* doi: 10.1371/journal.pone.0138024.

Williamson, P.R. et al. 2017. "The COMET Handbook: version 1.0." *Trials* 18: 280.

Appendix A: Computing Bounds on $\mathcal{P}_{\mathbf{j\succ k,C1}}$ and $\mathcal{P}_{\mathbf{j\succ k,C0}}$

This appendix considers approaches to computing lower and upper bounds on $\mathcal{P}_{\mathrm{j\succ k,C1}}$. The same basic logic can be used to obtain bounds on $\mathcal{P}_{\mathrm{j\succ k,C0}}$, so those details will not be considered here. In either case the results for bounds on $\mathcal{P}_{\mathrm{k\succ j,C\bullet}}$ follow symmetrically. Combined, these bounds can then be used to obtain bounds on the corresponding $\mathrm{PTE}_{\mathrm{j\succ k,C\bullet}}$ as in section 4.

Using the notation of section 3, (10) can be written as

$$\mathcal{P}_{\mathrm{j\succ k,C1}} \;=\; \Pr\left(\mathbf{y}_{\mathrm{k}} \geq \mathbf{y}_{\mathrm{j}}\right) \;=\; \Pr\left(\mathbf{Y} \in \mathbb{Y}_{\mathrm{j\succ k,C1}}\right) \;=\; \sum\nolimits_{\mathbf{Y}\in\mathbb{Y}_{\mathrm{j\succ k,C1}}} \Pr\left(\mathbf{Y}\right). \tag{A.1}$$

This follows since the events $\mathbf{Y} \in \mathbb{Y}_{\mathrm{j\succ k,C1}}$ are disjoint events. The computations for $\mathrm{UB}\left(\mathcal{P}_{\mathrm{j\succ k,C1}}\right)$ and $\mathrm{LB}\left(\mathcal{P}_{\mathrm{j\succ k,C1}}\right)$ are more complicated than those for C2-C5 shown in section 4. Well-defined bounds can be obtained, although they may or may not turn out to be informative.

Computing Upper Bounds on $\mathcal{P}_{j\succ k,C1}$

From Boole's general approach, an upper bound on the probability of the union of N arbitrary and possibly non-disjoint events e_{n} is given by

$$\mathrm{UB}\left(\Pr\left(\bigcup\nolimits_{\mathrm{n}=1}^{\mathrm{N}} \mathrm{e}_{\mathrm{n}}\right)\right) = \min\left\{\sum\nolimits_{\mathrm{n}=1}^{\mathrm{N}} \Pr\left(\mathrm{e}_{\mathrm{n}}\right), 1\right\}. \tag{A.2}$$

To be useful, information on the marginal probabilities $\Pr\left(\mathrm{e}_{\mathrm{n}}\right)$ must be available. This general approach can be used to compute a legitimate upper bound on $\mathcal{P}_{\mathrm{j\succ k,C1}}$ as

$$\mathrm{UB}\left(\mathcal{P}_{\mathrm{j\succ k,C1}}\right) \;=\; \min\left\{\left[\sum\nolimits_{\mathbf{Y}\in\mathbb{Y}_{\mathrm{j\succ k,C1}}} \min\left\{\Pr\left(\mathbf{y}_{\mathrm{j}}\right), \Pr\left(\mathbf{y}_{\mathrm{k}}\right)\right\}\right], 1\right\} \;\geq\; \min\left\{\sum\nolimits_{\mathbf{Y}\in\mathbb{Y}_{\mathrm{j\succ k,C1}}} \Pr\left(\mathbf{Y}\right), 1\right\}. \tag{A.3}$$

This follows since the $3^{\mathrm{M}} - 2^{\mathrm{M}}$ events $\mathbf{Y} \in \mathbb{Y}_{\mathrm{j\succ k,C1}}$ that jointly define $\mathcal{P}_{\mathrm{j\succ k,C1}}$, though themselves mutually disjoint, are not fully observable. As such their respective upper bounds

$$\mathrm{UB}\left(\Pr\left(\mathbf{Y}\right)\right) = \min\left\{\Pr\left(\mathbf{y}_{\mathrm{j}}\right), \Pr\left(\mathbf{y}_{\mathrm{k}}\right)\right\}, \text{ for each } \mathbf{Y} \in \mathbb{Y}_{\mathrm{j\succ k,C1}}, \tag{A.4}$$

are required to determine an overall $\mathrm{UB}\left(\mathcal{P}_{\mathrm{j\succ k,C1}}\right)$.

However (A.4) will generally result in an unnecessarily large $\mathrm{UB}\left(\mathcal{P}_{\mathrm{j\succ k,C1}}\right)$. Drawing on Boole (1854, chapter XIX), if any single marginal event probability $\Pr\left(\mathbf{y}_{\bullet}\right)$ appears more than once in the sum $\mathrm{S}_{\mathbf{Y}} = \sum_{\mathbf{Y}\in\mathbb{Y}_{\mathrm{j\succ k,C1}}} \min\left\{\Pr\left(\mathbf{y}_{\mathrm{j}}\right), \Pr\left(\mathbf{y}_{\mathrm{k}}\right)\right\}$ in the RHS of (A.3) then that sum should be redefined as a different sum, say $\tilde{\mathrm{S}}_{\mathbf{Y}}$, that includes that particular marginal event probability as a

summand only once, thus resulting in a smaller (i.e. tighter) upper bound.[43]

Such considerations are relevant here since the probabilities of the marginal events $\mathbf{y}_\bullet$ for each $\mathbf{Y} \in \mathbb{Y}_{j\succ k,C1}$ may appear multiple times in $S_{\mathbf{Y}}$. Consider for concreteness the seven cells in the C1 column of table 1 corresponding to the seven $\mathbf{Y}$ events that involve the marginal event $\mathbf{y}_k = [1,1,1]$ (rows 1-7 of the table). If $\Pr(\mathbf{y}_k)$ is less than the corresponding $\Pr(\mathbf{y}_j)$ for two or more of those seven events then that $\Pr(\mathbf{y}_k)$ should appear only once as a summand in $\tilde{S}_{\mathbf{Y}}$. Indeed, that $\Pr(\mathbf{y}_k)$ will appear in the sum $\tilde{S}_{\mathbf{Y}}$ that defines $\mathrm{UB}(\mathcal{P}_{j\succ k,C1})$ either exactly once (if it is smaller than any one or more of the seven corresponding $\Pr(\mathbf{y}_j)$) or not at all (if it is larger than all the corresponding $\Pr(\mathbf{y}_j)$).

Thus to compute the UB, let

$$a_{\mathbf{Y}} \;=\; \min\{\Pr(\mathbf{y}_j), \Pr(\mathbf{y}_k)\} \;=\; \mathrm{UB}(\Pr(\mathbf{Y})), \quad \text{for each } \mathbf{Y} \in \mathbb{Y}_{j\succ k,C1}. \tag{A.5}$$

Define $A_{\mathbf{Y}} = \{a_{\mathbf{Y}}\}$, $A_\bullet = \{\mathbf{y}_\bullet \,|\, \Pr(\mathbf{y}_\bullet) \in A_{\mathbf{Y}}\}$. Importantly the standard definition of a set as a collection of distinct objects[44] is used in defining A_j and A_k. Then, supposing there are no ties where $\Pr(\mathbf{y}_j) = \Pr(\mathbf{y}_k)$,[45]

$$\mathrm{UB}(\mathcal{P}_{j\succ k,C1}) = \min\left\{\left(\sum_{\mathbf{y}_j \in A_j} \Pr(\mathbf{y}_j) + \sum_{\mathbf{y}_k \in A_k} \Pr(\mathbf{y}_k)\right), 1\right\}. \tag{A.6}$$

Computing Lower Bounds on $\mathcal{P}_{j\succ k,C1}$

In the general case where events may or may not be disjoint an $\mathrm{LB}(\mathcal{P}_{j\succ k,C1})$ is given by

$$\mathrm{LB}(\mathcal{P}_{j\succ k,C1}) = \max_{\mathbf{Y} \in \mathbb{Y}_{j\succ k,C1}} \left\{\max\{\Pr(\mathbf{y}_j) + \Pr(\mathbf{y}_k) - 1, 0\}\right\}. \tag{A.7}$$

When the events $\mathbf{Y}$ are known to be disjoint a lower bound at least as tight if not tighter than (A.7) is obtained as

$$\mathrm{LB}(\mathcal{P}_{j\succ k,C1}) = \sum_{\mathbf{Y} \in \mathbb{Y}_{j\succ k,C1}} \max\{\Pr(\mathbf{y}_j) + \Pr(\mathbf{y}_k) - 1, 0\}. \tag{A.8}$$

Given the particular structure of the outcomes examined here, however, a still-tighter LB is possible to obtain without requiring any additional assumptions. The motivation for this approach

---

[43] Discerning this result from Boole's mid-nineteenth-century prose was, for the author of this paper at least, a challenging exercise.

[44] Formally $\{2,2,1,3\}$ and $\{b,c,a,c\}$ are not sets but $\{2,1,3\}$ and $\{b,c,a\}$ are.

[45] Ties can be handled at the cost of additional computations.

is that in applications—particularly where M is large—it may be found that all the summands $\max\left\{\Pr\left(\mathbf{y}_j\right)+\Pr\left(\mathbf{y}_k\right)-1,0\right\}$ in (A.8) are zero thus resulting in the uninformative $\mathrm{LB}\left(\mathcal{P}_{j\succ k,C1}\right)=0$.

To circumvent this, note that there are $2^{M+1}-3$ events $\mathbf{Y}$ in $\mathbb{Y}_{j\succ k,C1}$ where $\mathbf{y}_j=\mathbf{0}$ or $\mathbf{y}_k=\mathbf{1}$ or both; let this subset of $\mathbb{Y}_{j\succ k,C1}$ be denoted $\mathbb{Y}^{A}_{j\succ k,C1}$. For the $\mathbf{Y}$ in $\mathbb{Y}^{A}_{j\succ k,C1}$ the corresponding inequality events $\mathbf{y}_k\geq\mathbf{y}_j$ can be aggregated across the elements of $\mathbb{Y}^{A}_{j\succ k,C1}$ to define two disjoint sets $\mathbb{Y}^{A1}_{j\succ k,C1}=\left\{\mathbf{Y}\middle|\mathbf{y}_j=\mathbf{0}\wedge\mathbf{y}_k\neq\mathbf{0}\right\}$ and $\mathbb{Y}^{A2}_{j\succ k,C1}=\left\{\mathbf{Y}\middle|\mathbf{y}_j\neq\mathbf{0}\wedge\mathbf{y}_k=\mathbf{1}\right\}$ with $\mathbb{Y}^{A1}_{j\succ k,C1}\bigcup\mathbb{Y}^{A2}_{j\succ k,C1}=\mathbb{Y}^{A}_{j\succ k,C1}$ (note that $\mathbb{Y}^{A1}_{j\succ k,C1}$ coincides with $\mathbb{Y}_{j\succ k,C2}$). In the M=3 example, for instance, in the C1 column of table 1 $\mathbb{Y}^{A1}_{j\succ k,C1}$ corresponds to rows 1 plus 8-13 while $\mathbb{Y}^{A2}_{j\succ k,C1}$ corresponds to rows 2-7. The LBs on the probabilities of these aggregated events are

$$\begin{aligned}\mathrm{LB}\left(\Pr\left(\mathbf{Y}\in\mathbb{Y}^{A1}_{j\succ k,C1}\right)\right)&=\mathrm{LB}\left(\Pr\left(\mathbf{y}_j=\mathbf{0}\wedge\mathbf{y}_k\neq\mathbf{0}\right)\right)\\&=\max\left\{\Pr\left(\mathbf{y}_j=\mathbf{0}\right)+\Pr\left(\mathbf{y}_k\neq\mathbf{0}\right)-1,0\right\}\\&=\max\left\{\Pr\left(\mathbf{y}_j=\mathbf{0}\right)-\Pr\left(\mathbf{y}_k=\mathbf{0}\right),0\right\}\end{aligned}\tag{A.9}$$

$$\begin{aligned}\mathrm{LB}\left(\Pr\left(\mathbf{Y}\in\mathbb{Y}^{A2}_{j\succ k,C1}\right)\right)&=\mathrm{LB}\left(\Pr\left(\mathbf{y}_j\neq\mathbf{0}\wedge\mathbf{y}_k=\mathbf{1}\right)\right)\\&=\max\left\{\Pr\left(\mathbf{y}_j\neq\mathbf{0}\right)+\Pr\left(\mathbf{y}_k=\mathbf{1}\right)-1,0\right\}\\&=\max\left\{\Pr\left(\mathbf{y}_k=\mathbf{1}\right)-\Pr\left(\mathbf{y}_j=\mathbf{0}\right),0\right\}\end{aligned}\tag{A.10}$$

Thus

$$\begin{aligned}\mathrm{LB}\left(\mathcal{P}_{j\succ k,C1}\right)\;=\;&\mathrm{LB}\left(\Pr\left(\mathbf{Y}\in\mathbb{Y}^{A1}_{j\succ k,C1}\right)\right)\;+\;\mathrm{LB}\left(\Pr\left(\mathbf{Y}\in\mathbb{Y}^{A2}_{j\succ k,C1}\right)\right)\;+\\&\sum\nolimits_{\mathbf{Y}\in\mathbb{Y}_{j\succ k,C1}\setminus\mathbb{Y}^{A}_{j\succ k,C1}}\max\left\{\Pr\left(\mathbf{y}_j\right)+\Pr\left(\mathbf{y}_k\right)-1,0\right\}\end{aligned}\tag{A.11}$$

The formulation in (A.11) will be no smaller than (A.8) but in some instances will be larger, i.e. the bound in (A.11) will be at least as tight as the bound in (A.8).[46]

[46] It should be noted that aggregations other than those used to define e_1 and e_2 are possible and potentially more informative. The particular e_1 and e_2 aggregations described here were selected because they are informative with respect to the data structures described in section 4 and summarized in tables 2a-2c.

Appendix B: Treatment-Preference Characterizations, Probabilities, and PTEs with Multiple Binary Outcomes and More than Two Treatments

The relevant comparison of multiple treatment outcomes sometimes involves more than two treatments. In a study of diet-related cardiovascular disease prevention Estruch et al., 2018, consider three alternative treatments: Mediterranean diet supplemented with extra-virgin olive oil; Mediterranean diet supplemented with mixed nuts; and control diet (advice to reduce dietary fat). Nissen et al., 2016, compare in a three-arm trial the cardiovascular safety profiles of celecoxib, ibuprofen, and naproxen for patients with osteoarthritis or rheumatoid arthritis.[47] In both studies, composite cardiovascular endpoints are of primary interest.

This appendix considers briefly extensions of the results in sections 3 and 4 to such treatment settings, focusing on the three-treatment case for simplicity. Extending the treatment preference characterizations from section 3 to the three-treatment case is largely straightforward.

Treatment Preference and PTEs with More than Two Treatments

To show the results generically, let $\mathrm{j,k},\ell \in \{0,1,2\}$, $\mathrm{j} \neq \mathrm{k} \neq \ell \neq \mathrm{j}$, index the three distinct treatments, $\mathrm{T_j}$, $\mathrm{T_k}$, and T_ℓ, and corresponding potential outcomes, $\mathbf{y}_\mathrm{j}$, $\mathbf{y}_\mathrm{k}$, and $\mathbf{y}_\ell$, so that the $1\times 3\mathrm{M}$ vector of potential outcomes is $\mathbf{Y} = \left[\, \mathbf{y}_\mathrm{j} \,\middle|\, \mathbf{y}_\mathrm{k} \,\middle|\, \mathbf{y}_\ell \,\right]$. Characterizations $\mathrm{C}\bullet^{+}$ correspond to the $\mathrm{C}\bullet$ in defined in section 3:

$$\begin{aligned}\mathcal{P}_{\mathrm{j}\succ\{\mathrm{k},\ell\},\mathrm{C0}^{+}} &= \Pr\Big(\big(\mathrm{T_j} \succ_{\mathrm{C0}} \mathrm{T_k}\big) \wedge \big(\mathrm{T_j} \succ_{\mathrm{C0}} \mathrm{T}_\ell\big)\Big) \\ &= \Pr\Big(\big(\mathrm{s_k} > \mathrm{s_j}\big) \wedge \big(\mathrm{s}_\ell > \mathrm{s_j}\big)\Big)\end{aligned} \tag{B.1}$$

$$\begin{aligned}\mathcal{P}_{\mathrm{j}\succ\{\mathrm{k},\ell\},\mathrm{C1}^{+}} &= \Pr\Big(\big(\mathrm{T_j} \succ_{\mathrm{C1}} \mathrm{T_k}\big) \wedge \big(\mathrm{T_j} \succ_{\mathrm{C1}} \mathrm{T}_\ell\big)\Big) \\ &= \Pr\Big(\big(\mathbf{y}_\mathrm{k} \geq \mathbf{y}_\mathrm{j}\big) \wedge \big(\mathbf{y}_\ell \geq \mathbf{y}_\mathrm{j}\big)\Big)\end{aligned} \tag{B.2}$$

$$\begin{aligned}\mathcal{P}_{\mathrm{j}\succ\{\mathrm{k},\ell\},\mathrm{C2}^{+}} &= \Pr\Big(\big(\mathrm{T_j} \succ_{\mathrm{C2}} \mathrm{T_k}\big) \wedge \big(\mathrm{T_j} \succ_{\mathrm{C2}} \mathrm{T}_\ell\big)\Big) \\ &= \Pr\Big(\big(\mathbf{y}_\mathrm{j} = \mathbf{0}\big) \wedge \big(\mathbf{y}_\mathrm{k} \neq \mathbf{0}\big) \wedge \big(\mathbf{y}_\ell \neq \mathbf{0}\big)\Big)\end{aligned} \tag{B.3}$$

$$\begin{aligned}\mathcal{P}_{\mathrm{j}\succ\{\mathrm{k},\ell\},\mathrm{C3}^{+}} &= \Pr\Big(\big(\mathrm{T_j} \succ_{\mathrm{C3}} \mathrm{T_k}\big) \wedge \big(\mathrm{T_j} \succ_{\mathrm{C3}} \mathrm{T}_\ell\big)\Big) \\ &= \Pr\Big(\big(\mathbf{y}_\mathrm{j} \neq \mathbf{1}\big) \wedge \big(\mathbf{y}_\mathrm{k} = \mathbf{1}\big) \wedge \big(\mathbf{y}_\ell = \mathbf{1}\big)\Big)\end{aligned} \tag{B.4}$$

$$\begin{aligned}\mathcal{P}_{\mathrm{j}\succ\{\mathrm{k},\ell\},\mathrm{C4}^{+}} &= \Pr\Big(\big(\mathrm{T_j} \succ_{\mathrm{C4}} \mathrm{T_k}\big) \wedge \big(\mathrm{T_j} \succ_{\mathrm{C4}} \mathrm{T}_\ell\big)\Big) \\ &= \Pr\Big(\Big(\big(\mathbf{y}_\mathrm{j} = \mathbf{0}\big) \wedge \big(\mathbf{y}_\mathrm{k} \neq \mathbf{0}\big) \wedge \big(\mathbf{y}_\ell \neq \mathbf{0}\big)\Big) \vee \Big(\big(\mathbf{y}_\mathrm{j} \neq \mathbf{1}\big) \wedge \big(\mathbf{y}_\mathrm{k} = \mathbf{1}\big) \wedge \big(\mathbf{y}_\ell = \mathbf{1}\big)\Big)\Big)\end{aligned} \tag{B.5}$$

[47] See Mullahy, 2018a (p. 158) for a related discussion in the M=1 case with continuous outcomes.

$$\begin{aligned}\mathcal{P}_{\mathrm{j}\succ\{\mathrm{k},\ell\},\mathrm{C5}^{+}} &= \Pr\Big(\Big(\mathrm{T}_{\mathrm{j}} \succ_{\mathrm{C5}} \mathrm{T}_{\mathrm{k}}\Big) \wedge \Big(\mathrm{T}_{\mathrm{j}} \succ_{\mathrm{C5}} \mathrm{T}_{\ell}\Big)\Big) \\ &= \Pr\Big(\Big(\mathbf{y}_{\mathrm{j}} = \mathbf{0} \wedge \mathbf{y}_{\mathrm{k}} = \mathbf{1}\Big) \vee \Big(\mathbf{y}_{\mathrm{j}} = \mathbf{0} \wedge \mathbf{y}_{\ell} = \mathbf{1}\Big)\Big)\end{aligned} \tag{B.6}$$

$$\begin{aligned}\mathcal{P}_{\mathrm{j}\succ\{\mathrm{k},\ell\},\mathrm{C6(Z)}^{+}} &= \Pr\Big(\Big(\mathrm{T}_{\mathrm{j}} \succ_{\mathrm{C6(Z)}} \mathrm{T}_{\mathrm{k}}\Big) \wedge \Big(\mathrm{T}_{\mathrm{j}} \succ_{\mathrm{C6(Z)}} \mathrm{T}_{\ell}\Big)\Big) \\ &= \Pr\Big(\Big(\mathbf{y}_{\mathrm{j}} \in \mathrm{Z} \wedge \mathbf{y}_{\mathrm{k}} \in \mathrm{Z}^{\mathrm{c}}\Big) \vee \Big(\mathbf{y}_{\mathrm{j}} \in \mathrm{Z} \wedge \mathbf{y}_{\ell} \in \mathrm{Z}^{\mathrm{c}}\Big)\Big)\end{aligned} \tag{B.7}$$

In (B.1)-(B.7), $\mathcal{P}_{\mathrm{j}\succ\{\mathrm{k},\ell\},\mathrm{C}\bullet^{+}}$ is shorthand for "the probability that treatment T_{j} is preferred to both treatment T_{k} and treatment T_{ℓ} given characterization $\mathrm{C}\bullet^{+}$." Analogous to (26) it can be shown that

$$\mathcal{P}_{\mathrm{j}\succ\{\mathrm{k},\ell\},\mathrm{C0}^{+}} \;\geq\; \mathcal{P}_{\mathrm{j}\succ\{\mathrm{k},\ell\},\mathrm{C1}^{+}} \;\geq\; \mathcal{P}_{\mathrm{j}\succ\{\mathrm{k},\ell\},\mathrm{C4}^{+}} \;\geq\; \left\{ \begin{array}{l} \mathcal{P}_{\mathrm{j}\succ\{\mathrm{k},\ell\},\mathrm{C2}^{+}} \\ \mathcal{P}_{\mathrm{j}\succ\{\mathrm{k},\ell\},\mathrm{C3}^{+}} \end{array} \right\} \;\geq\; \mathcal{P}_{\mathrm{j}\succ\{\mathrm{k},\ell\},\mathrm{C5}^{+}} \tag{B.8}$$

How to characterize PTEs in these cases is not immediately evident. Specifically, given the characterizations of treatment preference $\mathrm{C}\bullet^{+}$, what are the relevant comparators or contrasts? In the two-treatment case the appropriate contrast is obvious, but when three or more treatments are involved several options might reasonably be considered.[48] Two possible PTE definitions based on (B.1)-(B.7) are discussed here, but others might certainly be considered; in any event the specification of any particular PTE should be tied to the evaluation question or treatment decision at hand.

In the first the PTE is defined by the difference between any of (B.1)-(B.7) and the probability that *both* T_{k} and T_{ℓ} are preferred to T_{j} given characterization $\mathrm{C}\bullet^{+}$. This latter probability is given by

$$\mathcal{P}_{\{\mathrm{k},\ell\}\succ\mathrm{j},\mathrm{C}\bullet^{+}} = \Pr\Big(\Big(\mathrm{T}_{\mathrm{k}} \succ_{\mathrm{C}\bullet} \mathrm{T}_{\mathrm{j}}\Big) \wedge \Big(\mathrm{T}_{\ell} \succ_{\mathrm{C}\bullet} \mathrm{T}_{\mathrm{j}}\Big)\Big), \tag{B.9}$$

so that

$$\begin{aligned}\mathrm{PTE}^{\mathrm{A}}_{\mathrm{j}\succ\{\mathrm{k},\ell\},\mathrm{C}\bullet^{+}} &= \Pr\Big(\Big(\mathrm{T}_{\mathrm{j}} \succ_{\mathrm{C}\bullet} \mathrm{T}_{\mathrm{k}}\Big) \wedge \Big(\mathrm{T}_{\mathrm{j}} \succ_{\mathrm{C}\bullet} \mathrm{T}_{\ell}\Big)\Big) - \Pr\Big(\Big(\mathrm{T}_{\mathrm{k}} \succ_{\mathrm{C}\bullet} \mathrm{T}_{\mathrm{j}}\Big) \wedge \Big(\mathrm{T}_{\ell} \succ_{\mathrm{C}\bullet} \mathrm{T}_{\mathrm{j}}\Big)\Big) \\ &= \mathcal{P}_{\mathrm{j}\succ\{\mathrm{k},\ell\},\mathrm{C}\bullet^{+}} - \mathcal{P}_{\{\mathrm{k},\ell\}\succ\mathrm{j},\mathrm{C}\bullet^{+}}\end{aligned} \tag{B.10}$$

The second PTE definition is given by the difference between any of (B.1)-(B.7) and the probability that T_{k} *and/or* T_{ℓ} is preferred to T_{j} under $\mathrm{C}\bullet$. This latter probability is

$$\Pr\Big(\Big(\mathrm{T}_{\mathrm{k}} \succ_{\mathrm{C}\bullet} \mathrm{T}_{\mathrm{j}}\Big) \vee \Big(\mathrm{T}_{\ell} \succ_{\mathrm{C}\bullet} \mathrm{T}_{\mathrm{j}}\Big)\Big) \tag{B.11}$$

so that

$$\mathrm{PTE}^{\mathrm{B}}_{\mathrm{j}\succ\{\mathrm{k},\ell\},\mathrm{C}\bullet^{+}} = \mathcal{P}_{\mathrm{j}\succ\{\mathrm{k},\ell\},\mathrm{C}\bullet^{+}} - \Pr\Big(\Big(\mathrm{T}_{\mathrm{k}} \succ_{\mathrm{C}\bullet} \mathrm{T}_{\mathrm{j}}\Big) \vee \Big(\mathrm{T}_{\ell} \succ_{\mathrm{C}\bullet} \mathrm{T}_{\mathrm{j}}\Big)\Big) \tag{B.12}$$

[48] For instance, Nissen et al., 2016, consider pairwise comparisons.

Bounds on Treatment-Preference Probabilities and PTEs

With three or more treatments, computing bounds for the $C0^+$ and $C1^+$ characterizations is quite complicated and won't be pursued here. For $C2^+ - C6^+$ the bounds are straightforward to derive and, under unconfoundedness, estimate. These are derived here for $C2^+$ but the same logic applies directly for $C3^+$ (a mirror image of $C2^+$), $C5^+$, $C6^+$, and (with a bit more algebra) for $C4^+$. (Unlike the two-treatment setting, the PTEs for $C2^+$, $C3^+$, and $C6^+$ are no longer point identified.)

From (B.3), the probability to be bounded is $\Pr\Big(\Big(\mathbf{y}_{\mathrm{j}}=\mathbf{0}\Big)\wedge\Big(\mathbf{y}_{\mathrm{k}}\neq\mathbf{0}\Big)\wedge\Big(\mathbf{y}_{\ell}\neq\mathbf{0}\Big)\Big)$. Using Boole's general results, it follows that

$$\begin{aligned}\mathrm{UB}\Big(\mathcal{P}_{\mathrm{j}\succ\{\mathrm{k},\ell\},C2^+}\Big) &= \min\Big\{\Pr\Big(\mathbf{y}_{\mathrm{j}}=\mathbf{0}\Big),\Pr\Big(\mathbf{y}_{\mathrm{k}}\neq\mathbf{0}\Big),\Pr\Big(\mathbf{y}_{\ell}\neq\mathbf{0}\Big)\Big\}\\ &= \min\Big\{\Pr\Big(\mathbf{y}_{\mathrm{j}}=\mathbf{0}\Big),1-\Pr\Big(\mathbf{y}_{\mathrm{k}}=\mathbf{0}\Big),1-\Pr\Big(\mathbf{y}_{\ell}=\mathbf{0}\Big)\Big\}\end{aligned} \tag{B.13}$$

and

$$\begin{aligned}\mathrm{LB}\Big(\mathcal{P}_{\mathrm{j}\succ\{\mathrm{k},\ell\},C2^+}\Big) &= \max\Big\{\Pr\Big(\mathbf{y}_{\mathrm{j}}=\mathbf{0}\Big)+\Pr\Big(\mathbf{y}_{\mathrm{k}}\neq\mathbf{0}\Big)+\Pr\Big(\mathbf{y}_{\ell}\neq\mathbf{0}\Big)-2,0\Big\}\\ &= \max\Big\{\Pr\Big(\mathbf{y}_{\mathrm{j}}=\mathbf{0}\Big)-\Big(\Pr\Big(\mathbf{y}_{\mathrm{k}}=\mathbf{0}\Big)+\Pr\Big(\mathbf{y}_{\ell}=\mathbf{0}\Big)\Big),0\Big\}\end{aligned} \tag{B.14}$$

Whether the LBs are informative and, if so, how such information could be used to discern the preferred treatment are, of course, relevant questions in applications.[49] In practice one can compute these bounds for the permutations of $\mathrm{j},\mathrm{k},\ell$ based (as in section 5) on estimates of the composite-outcome probabilities $\Pr\Big(\mathbf{y}=\mathbf{0}\Big|\mathbf{x}=\mathbf{x}_{\bullet}\Big)$ under unconfoundedness assumptions and given observed outcome data $\mathbf{y}=1\Big(\mathbf{x}=\mathbf{x}_{\mathrm{j}}\Big)\mathbf{y}_{\mathrm{j}}+1\Big(\mathbf{x}=\mathbf{x}_{\mathrm{k}}\Big)\mathbf{y}_{\mathrm{k}}+1\Big(\mathbf{x}=\mathbf{x}_{\ell}\Big)\mathbf{y}_{\ell}$.

Bounds on the PTEs defined in (B.10) and (B.12) follow from (B.13)-(B.14) and from bounds on the subtrahends in (B.10) and (B.12). The latter are:

$$\begin{aligned}\mathrm{UB}_{\mathrm{A}}=\mathrm{UB}\Big(\mathcal{P}_{\{\mathrm{k},\ell\}\succ\mathrm{j},C2^+}\Big) &= \mathrm{UB}\Big(\Big(\mathbf{y}_{\mathrm{j}}\neq\mathbf{0}\Big)\wedge\Big(\mathbf{y}_{\mathrm{k}}=\mathbf{0}\Big)\wedge\Big(\mathbf{y}_{\ell}=\mathbf{0}\Big)\Big)\\ &= \min\Big\{1-\Pr\Big(\mathbf{y}_{\mathrm{j}}=\mathbf{0}\Big),\Pr\Big(\mathbf{y}_{\mathrm{k}}=\mathbf{0}\Big),\Pr\Big(\mathbf{y}_{\ell}=\mathbf{0}\Big)\Big\}\end{aligned} \tag{B.15}$$

$$\begin{aligned}\mathrm{LB}_{\mathrm{A}}=\mathrm{LB}\Big(\mathcal{P}_{\{\mathrm{k},\ell\}\succ\mathrm{j},C2^+}\Big) &= \mathrm{LB}\Big(\Big(\mathbf{y}_{\mathrm{j}}\neq\mathbf{0}\Big)\wedge\Big(\mathbf{y}_{\mathrm{k}}=\mathbf{0}\Big)\wedge\Big(\mathbf{y}_{\ell}=\mathbf{0}\Big)\Big)\\ &= \max\Big\{\Pr\Big(\mathbf{y}_{\mathrm{k}}=\mathbf{0}\Big)+\Pr\Big(\mathbf{y}_{\ell}=\mathbf{0}\Big)-\Pr\Big(\mathbf{y}_{\mathrm{j}}=\mathbf{0}\Big)-1,0\Big\}\end{aligned} \tag{B.16}$$

[49] For any of the $\mathrm{j},\mathrm{k},\ell$ arrangements, it is straightforward to see that at most one of the three LBs, $\max\Big\{\Pr\Big(\mathbf{y}_{\mathrm{j}}=\mathbf{0}\Big)-\Big(\Pr\Big(\mathbf{y}_{\mathrm{k}}=\mathbf{0}\Big)+\Pr\Big(\mathbf{y}_{\ell}=\mathbf{0}\Big)\Big),0\Big\}$, can be nonzero.

$$\begin{aligned}\mathrm{UB_B} &= \mathrm{UB}\Big(\Pr\Big(\Big(\mathrm{T_k} \succ_{\mathrm{C2}} \mathrm{T_j}\Big) \vee \Big(\mathrm{T_\ell} \succ_{\mathrm{C2}} \mathrm{T_j}\Big)\Big)\Big) \\ &= \min\Big\{\Pr\Big(\mathbf{y}_{\mathrm{k}} = \mathbf{0}\Big) + \Pr\Big(\mathbf{y}_{\ell} = \mathbf{0}\Big), 1 - \Pr\Big(\mathbf{y}_{\mathrm{j}} = \mathbf{0}\Big)\Big\}\end{aligned} \tag{B.17}$$

$$\begin{aligned}\mathrm{LB_B} &= \mathrm{LB}\Big(\Pr\Big(\Big(\mathrm{T_k} \succ_{\mathrm{C2}} \mathrm{T_j}\Big) \vee \Big(\mathrm{T_\ell} \succ_{\mathrm{C2}} \mathrm{T_j}\Big)\Big)\Big) \\ &= \max\Big\{\max\Big\{\Pr\Big(\mathbf{y}_{\mathrm{k}} = \mathbf{0}\Big), \Pr\Big(\mathbf{y}_{\ell} = \mathbf{0}\Big)\Big\} - \Pr\Big(\mathbf{y}_{\mathrm{j}} = \mathbf{0}\Big), 0\Big\}\end{aligned} \tag{B.18}$$

It follows that:

$$\mathrm{UB}\Big(\mathrm{PTE}^{\mathrm{A}}_{\mathrm{j}\succ\{\mathrm{k},\ell\},\mathrm{C2}^{+}}\Big) = \min\Big\{\Pr\Big(\mathbf{y}_{\mathrm{j}} = \mathbf{0}\Big), 1 - \Pr\Big(\mathbf{y}_{\mathrm{k}} = \mathbf{0}\Big), 1 - \Pr\Big(\mathbf{y}_{\ell} = \mathbf{0}\Big)\Big\} - \mathrm{LB_A} \tag{B.19}$$

$$\mathrm{LB}\Big(\mathrm{PTE}^{\mathrm{A}}_{\mathrm{j}\succ\{\mathrm{k},\ell\},\mathrm{C2}^{+}}\Big) = \max\Big\{\Pr\Big(\mathbf{y}_{\mathrm{j}} = \mathbf{0}\Big) - \Big(\Pr\Big(\mathbf{y}_{\mathrm{k}} = \mathbf{0}\Big) + \Pr\Big(\mathbf{y}_{\ell} = \mathbf{0}\Big)\Big), 0\Big\} - \mathrm{UB_A} \tag{B.20}$$

$$\mathrm{UB}\Big(\mathrm{PTE}^{\mathrm{B}}_{\mathrm{j}\succ\{\mathrm{k},\ell\},\mathrm{C2}^{+}}\Big) = \min\Big\{\Pr\Big(\mathbf{y}_{\mathrm{j}} = \mathbf{0}\Big), 1 - \Pr\Big(\mathbf{y}_{\mathrm{k}} = \mathbf{0}\Big), 1 - \Pr\Big(\mathbf{y}_{\ell} = \mathbf{0}\Big)\Big\} - \mathrm{LB_B} \tag{B.21}$$

$$\mathrm{LB}\Big(\mathrm{PTE}^{\mathrm{B}}_{\mathrm{j}\succ\{\mathrm{k},\ell\},\mathrm{C2}^{+}}\Big) = \max\Big\{\Pr\Big(\mathbf{y}_{\mathrm{j}} = \mathbf{0}\Big) - \Big(\Pr\Big(\mathbf{y}_{\mathrm{k}} = \mathbf{0}\Big) + \Pr\Big(\mathbf{y}_{\ell} = \mathbf{0}\Big)\Big), 0\Big\} - \mathrm{UB_B} \tag{B.22}$$

All terms in these bounds can be identified and estimated given estimates of the relevant composite-outcome joint probabilities.

Figure 1a: $\Pr\left(\mathbf{y}_j = \mathbf{0}\right)$ and $\Pr\left(\mathbf{y}_k = \mathbf{0}\right)$ Combinations Consistent with $\text{PTE}_{j\succ k,C2} \geq 0$ — Point Identification (light and dark shading) and $\text{LB}\left(\text{PTE}_{j\succ k,C2}\right) \geq 0$ (dark shading)

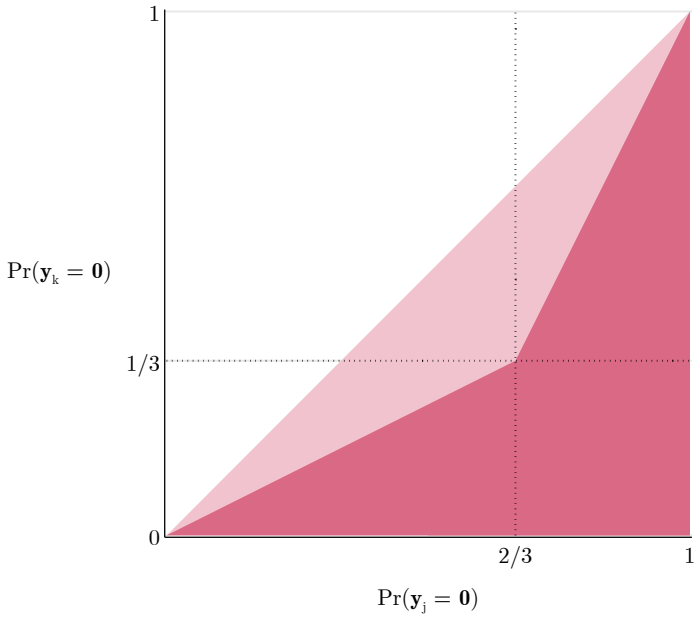

Figure 1b: $\Pr\left(\mathbf{y}_j = \mathbf{0}\right)$ and $\Pr\left(\mathbf{y}_k = \mathbf{0}\right)$ Combinations Consistent with $\text{LB}\left(\text{PTE}_{j\succ k,C2}\right) = 0$ and $\text{LB}\left(\text{PTE}_{j\succ k,C5}\right) = 0$ at Selected Values of $\Pr\left(\mathbf{y}_j = \mathbf{1}\right)$ and $\Pr\left(\mathbf{y}_k = \mathbf{1}\right)$

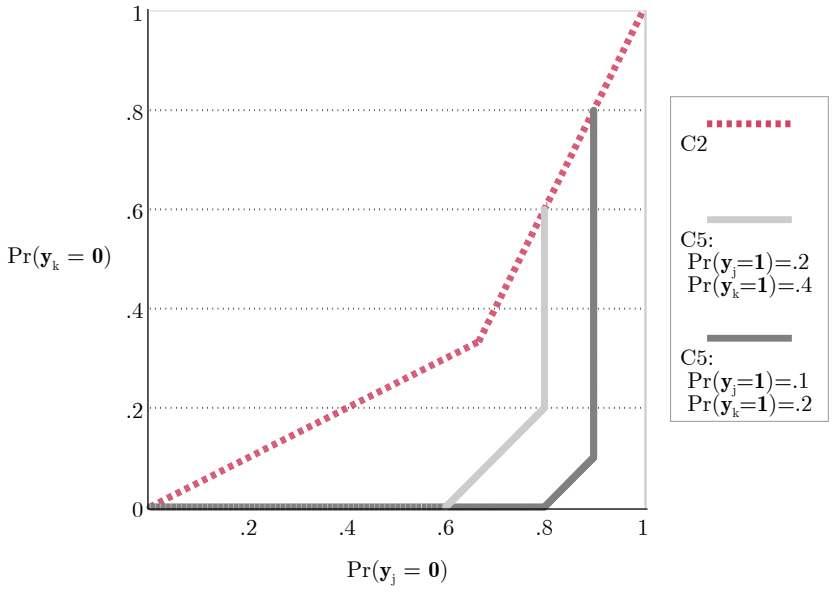

Table 1: Comparison across $C\bullet$ of $\mathbf{Y}$ Values Consistent with $T_j \succ_{C\bullet} T_k$, M=3

(22 of $2^{2M} = 64$ possible $\mathbf{Y}$ values correspond to $T_j \succ_{C\bullet} T_k$ for at least one $C\bullet$)

	$\mathbf{Y} = [\, \mathbf{y}_j \mid \mathbf{y}_k \,]$	$T_j \succ_{C\bullet} T_k$					
		C0	C1	C2	C3	C4	C5
1	[0 0 0 | 1 1 1]	•	•	•	•	•	•
2	[1 0 0 | 1 1 1]	•	•		•	•	
3	[0 1 0 | 1 1 1]	•	•		•	•	
4	[0 0 1 | 1 1 1]	•	•		•	•	
5	[1 1 0 | 1 1 1]	•	•		•	•	
6	[1 0 1 | 1 1 1]	•	•		•	•	
7	[0 1 1 | 1 1 1]	•	•		•	•	
8	[0 0 0 | 1 0 0]	•	•	•		•	
9	[0 0 0 | 0 1 0]	•	•	•		•	
10	[0 0 0 | 0 0 1]	•	•	•		•	
11	[0 0 0 | 1 1 0]	•	•	•		•	
12	[0 0 0 | 1 0 1]	•	•	•		•	
13	[0 0 0 | 0 1 1]	•	•	•		•	
14	[1 0 0 | 1 1 0]	•	•				
15	[0 1 0 | 1 1 0]	•	•				
16	[1 0 0 | 1 0 1]	•	•				
17	[0 0 1 | 1 0 1]	•	•				
18	[0 1 0 | 0 1 1]	•	•				
19	[0 0 1 | 0 1 1]	•	•				
20	[0 0 1 | 1 1 0]	•					
21	[0 1 0 | 1 0 1]	•					
22	[1 0 0 | 0 1 1]	•					

Table 2a: $\mathcal{P}_{j\succ k,C\bullet}$, $\mathcal{P}_{k\succ j,C\bullet}$ and $\text{PTE}_{j\succ k,C\bullet}$: True Values and Boole Bounds, M=2

(Cell Entries: True Probability {LB, UB}; $\text{PTE}_{j\succ k,C\bullet}$ are Point Identified for C2 and C3)

ρ	$\Pr\left(y_{k,m}=1\right)$	C•	$\mathcal{P}_{j\succ k,C\bullet}$	$\mathcal{P}_{k\succ j,C\bullet}$	$\text{PTE}_{j\succ k,C\bullet}$
0	.2	C0	.30 {.17, .36}	.12 {0, .19}	.17 {-.02, .36}
		C1	.30 {.17, .36}	.12 {0, .19}	.17 {-.02, .36}
		C2	.29 {**.17**, .36}	.12 {0, .19}	**.17**
		C3	.04 {**.03**, .04}	.01 {0, .01}	**.03**
		C4	.30 {.17, .36}	.12 {0, .19}	.17 {-.02, .36}
		C5	.03 {0, .04}	.01 {0, .01}	.03 {-.01, .04}
	.5	C0	.65 {.56, .93}	.05 {0, .19}	.60 {.37, .93}
		C1	.65 {.56, .93}	.05 {0, .19}	.60 {.37, .93}
		C2	.61 {**.56**, .75}	.05 {0, .19}	**.56**
		C3	.25 {**.24**, .25}	.01 {0, .01}	**.24**
		C4	.65 {.56, .75}	.05 {0, .19}	.60 {.37, .75}
		C5	.20 {.06, .25}	.003 {0, .01}	.20 {.05, .25}
.5	.2	C0	.23 {.15, .45}	.07 {0, .17}	.17 {-.02, .45}
		C1	.23 {.15, .45}	.07 {0, .17}	.17 {-.02, .45}
		C2	.20 {**.15**, .31}	.05 {0, .17}	**.15**
		C3	.07 {**.06**, .09}	.02 {0, .03}	**.06**
		C4	.23 {.15, .31}	.07 {0, .17}	.17 {-.02, .31}
		C5	.04 {0, .09}	.005 {0, .03}	.03 {-.03, .09}
	.5	C0	.60 {.50, .80}	.01 {0, .17}	.58 {.33, .80}
		C1	.60 {.50, .80}	.01 {0, .17}	.58 {.33, .80}
		C2	.51 {**.50**, .67}	.01 {0, .17}	**.50**
		C3	.31 {**.30**, .33}	.004 {0, .03}	**.30**
		C4	.60 {.50, .67}	.01 {0, .17}	.58 {.33, .67}
		C5	.22 {.17, .33}	.004 {0, .03}	.22 {.13, .33}

(cont.)

(cont.)

Note to table 2a: For m=1,…,M the calculations assume that: the $\mathbf{y}_j$ and $\mathbf{y}_k$ M-vectors have elements $y_{j,m} = 1\left(y^*_{j,m} > 0\right)$, $y_{k,m} = 1\left(y^*_{k,m} > 0\right)$, with $\mathbf{Y}^* = \left[\, \mathbf{y}^*_j \,\middle|\, \mathbf{y}^*_k \,\right] \sim \text{MVN}\left(\left[\, \boldsymbol{\mu}_j \,\middle|\, \boldsymbol{\mu}_k \,\right], \mathbf{R}\right)$; $\mu_{j,m} = \Phi^{-1}(.1)$, so that $\Pr\left(y_{j,m} = 1\right) = .1$; $\mu_{k,m} = \Phi^{-1}\left(\Pr\left(y_{k,m} = 1\right)\right)$, with the specific values given in the table; and $\mathbf{R}$ is a $2M \times 2M$ correlation matrix with all off-diagonal elements equal to ρ. $\Pr\left(y_{k,m} = 1\right) = .2$ and $\Pr\left(y_{k,m} = 1\right) = .5$ represent "small" and "large" TEs, respectively. LBs for $\mathcal{P}_{j \succ k, C_\bullet}$ for C2 and C3, shown in bold, correspond to corresponding point-identified $\text{PTE}_{j \succ k, C_\bullet}$. For all entries UB>True Probability>LB, although these strict inequalities may be obscured by rounding.

Table 2b: $\mathcal{P}_{j\succ k,C\bullet}$, $\mathcal{P}_{k\succ j,C\bullet}$ and $\text{PTE}_{j\succ k,C\bullet}$: True Values and Boole Bounds, M=3

(Cell Entries: True Probability {LB, UB}; $\text{PTE}_{j\succ k,C\bullet}$ are Point Identified for C2 and C3)

ρ	$\Pr\left(y_{k,m}=1\right)$	C•	$\mathcal{P}_{j\succ k,C\bullet}$	$\mathcal{P}_{k\succ j,C\bullet}$	$\text{PTE}_{j\succ k,C\bullet}$
0	.2	C0	.38 {.22, .49}	.15 {0, .27}	.23 {-.05, .49}
		C1	.37 {.22, .49}	.15 {0, .27}	.23 {-.05, .49}
		C2	.36 {**.22**, .49}	.14 {0, .27}	**.22**
		C3	.01 {**.01**, .01}	.001 {0, .001}	**.01**
		C4	.36 {.22, .49}	.14 {0, .27}	.22 {-.05, .49}
		C5	.01 {0, .01}	.001 {0, .001}	.01 {-.001, .01}
	.5	C0	.76 {.60, 1}	.04 {0, .27}	.72 {.33, 1}
		C1	.73 {.60, 1}	.04 {0, .27}	.69 {.33, 1}
		C2	.64 {**.60**, .73}	.03 {0, .13}	**.60**
		C3	.13 {**.12**, .13}	.001 {0, .001}	**.12**
		C4	.67 {.60, .85}	.03 {0, .13}	.64 {.48, .85}
		C5	.09 {0, .13}	.0001 {0, .001}	.09 {-.001, .13}
.5	.2	C0	.29 {.17, .44}	.07 {0, .22}	.22 {-.05, .44}
		C1	.28 {.17, .44}	.07 {0, .22}	.21 {-.05, .44}
		C2	.22 {**.17**, .39}	.05 {0, .22}	**.17**
		C3	.04 {**.03**, .05}	.01 {0, .02}	**.03**
		C4	.25 {.17, .39}	.06 {0, .22}	.19 {-.05, .39}
		C5	.01 {0, .05}	.001 {0, .02}	.01 {-.02, .05}
	.5	C0	.70 {.53, .95}	.01 {0, .22}	.69 {.31, .95}
		C1	.68 {.53, .95}	.01 {0, .22}	.68 {.31, .95}
		C2	.54 {**.53**, .75}	.01 {0, .22}	**.53**
		C3	.24 {**.23**, .25}	.002 {0, .02}	**.23**
		C4	.65 {.53, .75}	.01 {0, .22}	.64 {.31, .75}
		C5	.12 {.03, .25}	.00001 {0, .02}	.12 {.02, .25}

Note: See note to table 2a.

Table 2c: $\mathcal{P}_{j\succ k,C\bullet}$, $\mathcal{P}_{k\succ j,C\bullet}$ and $\text{PTE}_{j\succ k,C\bullet}$: True Values and Boole Bounds, M=4

(Cell Entries: True Probability {LB, UB}; $\text{PTE}_{j\succ k,C\bullet}$ are Point Identified for C2 and C3)

ρ	$\Pr(y_{k,m}=1)$	C•	$\mathcal{P}_{j\succ k,C\bullet}$	$\mathcal{P}_{k\succ j,C\bullet}$	$\text{PTE}_{j\succ k,C\bullet}$
0	.2	C0	.44 {.25, .59}	.16 {0, .34}	.28 {-.10, .59}
		C1	.42 {.25, .59}	.15 {0, .34}	.26 {-.10, .59}
		C2	.39 {**.25**, .59}	.14 {0, .34}	**.25**
		C3	.002 {**.002**, .002}	.0001 {0, .0001}	**.002**
		C4	.39 {.25, .59}	.14 {0, .34}	.25 {-.10, .59}
		C5	.001 {0, .002}	.0004 {0, .0001}	.001 {-.0001, .002}
	.5	C0	.83 {.59, .99}	.04 {0, .11}	.80 {.48, .99}
		C1	.75 {.59, .99}	.03 {0, .11}	.72 {.48, .99}
		C2	.62 {**.59**, .66}	.02 {0, .06}	**.59**
		C3	.06 {**.06**, .06}	.0001 {0, .0001}	**.06**
		C4	.64 {.59, .72}	.02 {0, .06}	.62 {.53, .72}
		C5	.04 {0, .06}	.00001 {0, .0001}	.04 {-.0001, .06}
.5	.2	C0	.34 {.19, .53}	.08 {0, .26}	.26 {-.07, .53}
		C1	.30 {.19, .53}	.07 {0, .26}	.24 {-.07, .53}
		C2	.24 {**.19**, .44}	.05 {0, .26}	**.19**
		C3	.03 {**.02**, .03}	.004 {0, .01}	**.02**
		C4	.26 {.19, .44}	.06 {0, .26}	.21 {-.07, .44}
		C5	.004 {0, .03}	.0001 {0, .009}	.004 {-.01, .03}
	.5	C0	.76 {.54, 1}	.01 {0, .26}	.75 {.28, 1}
		C1	.73 {.54, 1}	.01 {0, .26}	.72 {.28, 1}
		C2	.55 {**.54**, .74}	.01 {0, .20}	**.54**
		C3	.19 {**.19**, .20}	.001 {0, .01}	**.19**
		C4	.67 {.54, .80}	.01 {0, .21}	.66 {.33, .80}
		C5	.07 {0, .20}	.0000003 {0, .009}	.07 {-.01, .20}

Note: See note to table 2a.

Table 3a: Estimated Bounds on $\mathcal{P}_{j\succ k,C\bullet}$, $\mathcal{P}_{k\succ j,C\bullet}$, and $\text{PTE}_{j\succ k,C\bullet}$—MTO Sample
(T_j is intervention, T_k is control; $\text{PTE}_{j\succ k,C\bullet}$ is point identified for C2 and C3);

C•	M=2 (Obesity, Diabetes)			M=4 (Obesity, Diabetes, Depression, Hypertension)		
	$\mathcal{P}_{j\succ k,C\bullet}$	$\mathcal{P}_{k\succ j,C\bullet}$	$\text{PTE}_{j\succ k,C\bullet}$	$\mathcal{P}_{j\succ k,C\bullet}$	$\mathcal{P}_{k\succ j,C\bullet}$	$\text{PTE}_{j\succ k,C\bullet}$
C0	{.02, .62}	{0, .59}	{-.57, .62}	{.02, .85}	{.004, .95}	{-.93, .85}
C1	{.02, .62}	{0, .59}	{-.57, .62}	{.02, .85}	{.004, .95}	{-.93, .85}
C2	{**.01**, .37}	{0, .36}	**.01**	{**.02**, .23}	{0, .21}	**.02**
C3	{**.02**, .13}	{0, .10}	**.02**	{0, .03}	{**.004** .03}	**-.004**
C4	{.02, .49}	{0, .46}	{-.44, .49}	{.02, .26}	{.004, .24}	{-.22, .25}
C5	{0, .13}	{0, .10}	{-.10, .13}	{0, .03}	{0, .03}	{-.03, .03}

Table 3b: Estimated Bounds on $\mathcal{P}_{j\succ k,C\bullet}$, $\mathcal{P}_{k\succ j,C\bullet}$, and $\text{PTE}_{j\succ k,C\bullet}$—BRFSS Sample (M=7)
(Age and Schooling treatments; $\text{PTE}_{j\succ k,C\bullet}$ is point identified for C2 and C3)

C•	Age (T_j=18-44; T_k=45-64)			Schooling (T_j=Not college grad.; T_k=college grad.)		
	$\mathcal{P}_{j\succ k,C\bullet}$	$\mathcal{P}_{k\succ j,C\bullet}$	$\text{PTE}_{j\succ k,C\bullet}$	$\mathcal{P}_{j\succ k,C\bullet}$	$\mathcal{P}_{k\succ j,C\bullet}$	$\text{PTE}_{j\succ k,C\bullet}$
C0	{.29, 1}	{0, .64}	{-.35, 1}	{0, .89}	{.05, 1}	{-1, .84}
C1	{.29, .97}	{0, .56}	{-.26, .97}	{0, 75}	{.05, .88}	{-.88, .70}
C2	{**.29**, .51}	{0, .22}	**.29**	{0, .32}	{**.05**, .37}	**-.05**
C3	{**.0005**, .001}	{0, .0001}	**.0005**	{0, .0002}	{**.0003**, .0005}	**-.0003**
C4	{.29, .51}	{0, .22}	{.08, .51}	{0, .32}	{.05, .37}	{-.37, .26}
C5	{0, .0006}	{0, .0001}	{-.0001, .0006}	{0, .0002}	{0, .0005}	{-.0005, .0002}

Table 4: Comparison across $C\bullet*$ of $\mathbf{Y}$ Values Consistent with $T_j \succ_{C\bullet} T_k$, M=2, G=3

(31 of $G^{2M} = 81$ possible $\mathbf{Y}$ values correspond to $T_j \succ_{C\bullet} T_k$ for at least one $C\bullet*$)

	$\mathbf{Y} = \left[\mathbf{y}_j \middle\vdots \mathbf{y}_k \right]$	$T_j \succ_{C\bullet*} T_k$					
		C0*	C1*	C2*	C3*	C4*	C5*
1	$\left[\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right]$	●	●	●		●	
2	$\left[\begin{smallmatrix} 0 & 0 \\ 0 & 2 \end{smallmatrix}\right]$	●	●	●		●	
3	$\left[\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right]$	●	●	●		●	
4	$\left[\begin{smallmatrix} 0 & 0 \\ 1 & 1 \end{smallmatrix}\right]$	●	●	●		●	
5	$\left[\begin{smallmatrix} 0 & 0 \\ 1 & 2 \end{smallmatrix}\right]$	●	●	●		●	
6	$\left[\begin{smallmatrix} 0 & 0 \\ 2 & 0 \end{smallmatrix}\right]$	●	●	●		●	
7	$\left[\begin{smallmatrix} 0 & 0 \\ 2 & 1 \end{smallmatrix}\right]$	●	●	●		●	
8	$\left[\begin{smallmatrix} 0 & 0 \\ 2 & 2 \end{smallmatrix}\right]$	●	●	●	●	●	●
9	$\left[\begin{smallmatrix} 0 & 1 \\ 0 & 2 \end{smallmatrix}\right]$	●	●				
10	$\left[\begin{smallmatrix} 0 & 1 \\ 1 & 1 \end{smallmatrix}\right]$	●	●				
11	$\left[\begin{smallmatrix} 0 & 1 \\ 1 & 2 \end{smallmatrix}\right]$	●	●				
12	$\left[\begin{smallmatrix} 0 & 1 \\ 2 & 0 \end{smallmatrix}\right]$	●					
13	$\left[\begin{smallmatrix} 0 & 1 \\ 2 & 1 \end{smallmatrix}\right]$	●	●				
14	$\left[\begin{smallmatrix} 0 & 1 \\ 2 & 2 \end{smallmatrix}\right]$	●	●		●	●	
15	$\left[\begin{smallmatrix} 0 & 2 \\ 1 & 2 \end{smallmatrix}\right]$	●	●				
16	$\left[\begin{smallmatrix} 0 & 2 \\ 2 & 1 \end{smallmatrix}\right]$	●					
17	$\left[\begin{smallmatrix} 0 & 2 \\ 2 & 2 \end{smallmatrix}\right]$	●	●		●	●	
18	$\left[\begin{smallmatrix} 1 & 0 \\ 0 & 2 \end{smallmatrix}\right]$	●					
19	$\left[\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right]$	●	●				
20	$\left[\begin{smallmatrix} 1 & 0 \\ 1 & 2 \end{smallmatrix}\right]$	●	●				
21	$\left[\begin{smallmatrix} 1 & 0 \\ 2 & 0 \end{smallmatrix}\right]$	●	●				
22	$\left[\begin{smallmatrix} 1 & 0 \\ 2 & 1 \end{smallmatrix}\right]$	●	●				
23	$\left[\begin{smallmatrix} 1 & 0 \\ 2 & 2 \end{smallmatrix}\right]$	●	●		●	●	
24	$\left[\begin{smallmatrix} 1 & 1 \\ 1 & 2 \end{smallmatrix}\right]$	●	●				
25	$\left[\begin{smallmatrix} 1 & 1 \\ 2 & 1 \end{smallmatrix}\right]$	●	●				
26	$\left[\begin{smallmatrix} 1 & 1 \\ 2 & 2 \end{smallmatrix}\right]$	●	●		●	●	
27	$\left[\begin{smallmatrix} 1 & 2 \\ 2 & 2 \end{smallmatrix}\right]$	●	●		●	●	
28	$\left[\begin{smallmatrix} 2 & 0 \\ 1 & 2 \end{smallmatrix}\right]$	●					
29	$\left[\begin{smallmatrix} 2 & 0 \\ 2 & 1 \end{smallmatrix}\right]$	●	●				
30	$\left[\begin{smallmatrix} 2 & 0 \\ 2 & 2 \end{smallmatrix}\right]$	●	●		●	●	
31	$\left[\begin{smallmatrix} 2 & 1 \\ 2 & 2 \end{smallmatrix}\right]$	●	●		●	●	

Table 5: $\#\mathbb{Y}_{j\succ k,C\bullet *}$ and $\#\mathbb{J}_{j\succ k,C\bullet *}$ for $M\in\{2,3,4,5\}$ and $G\in\{2,3,4\}$

			G								
			2			3			4		
			G^{2M}	$\#\mathbb{Y}_{j\succ k,C\bullet *}$	$\#\mathbb{J}_{j\succ k,C\bullet *}$	G^{2M}	$\#\mathbb{Y}_{j\succ k,C\bullet *}$	$\#\mathbb{J}_{j\succ k,C\bullet *}$	G^{2M}	$\#\mathbb{Y}_{j\succ k,C\bullet *}$	$\#\mathbb{J}_{j\succ k,C\bullet *}$
M	2	C0*	16	5	3	81	31	8	256	106	15
		C1*		5	3		27	8		84	15
		C2*		3	1		8	1		15	1
		C3*		3	3		8	8		15	15
		C4*		5	3		15	8		29	15
		C5*		1	1		1	1		1	1
	3	C0*	64	22	7	729	294	26	4096	1758	63
		C1*		19	7		189	26		936	63
		C2*		7	1		26	1		63	1
		C3*		7	7		26	26		63	63
		C4*		13	7		51	26		125	63
		C5*		1	1		1	1		1	1
	4	C0*	256	93	15	6561	2727	80	65536	28722	255
		C1*		65	15		1215	80		9744	255
		C2*		15	1		80	1		255	1
		C3*		15	15		80	80		255	255
		C4*		29	15		159	80		509	255
		C5*		1	1		1	1		1	1
	5	C0*	1024	386	31	59049	25048	242	1048576	466136	1023
		C1*		211	31		7533	242		98976	1023
		C2*		31	1		242	1		1023	1
		C3*		31	31		242	242		1023	1023
		C4*		61	31		483	242		2045	1023
		C5*		1	1		1	1		1	1

(cont.)

(cont.)

Note to table 5: The general formulae on which this table's entries are based are:

$\#\mathbb{Y}_{j\succ k,C0^*}$: See oeis.org A000346, A212730, A213465 for the basic idea.

$\#\mathbb{Y}_{j\succ k,C1^*}$: $\left(.5G\left(G+1\right)\right)^M - G^M$

$\#\mathbb{Y}_{j\succ k,C2^*}$ and $\#\mathbb{Y}_{j\succ k,C3^*}$: $G^M - 1$

$\#\mathbb{Y}_{j\succ k,C4^*}$: $2G^M - 3$

$\#\mathbb{J}_{j\succ k,C0^*}$, $\#\mathbb{J}_{j\succ k,C1^*}$, $\#\mathbb{J}_{j\succ k,C3^*}$, and $\#\mathbb{J}_{j\succ k,C4^*}$: $G^M - 1$

Not displayed in the table, $\#\mathbb{K}_{j\succ k,C\bullet^*}$ is the same as $\#\mathbb{J}_{j\succ k,C\bullet^*}$ for C0*, C1*, and C4* but switches with $\#\mathbb{J}_{j\succ k,C\bullet^*}$ between C2* and C3*.