Supplemental Appendix for
"Optimal Dynamic Treatment Regimes
and Partial Welfare Ordering"

Sukjin Han

Department of Economics

University of Bristol

First Draft: August 8, 2019

This Draft: December 19, 2022

Abstract

In Section A, the analysis with binary outcomes and discrete covariates is extended to continuous outcomes and covariates, and time-varying covariates and stochastic regimes are discussed. Section B provides the detailed definition of the matrices used in Section 3.3 of the main text. Section C contains most proofs of theorems and lemmas. Section D shows how to formally incorporate additional identifying assumptions in the framework. Section E presents numerical exercises. Finally, Section F contains discussions on topological sorts, cardinality reduction, and inference, among others.

A Extensions

A.1 Continuous Y_t and X

Suppose the outcomes Y_t's and pre-treatment covariate vector X are continuously distributed on $[y_l, y_u]$ and $\mathcal{X}$, respectively. Consider the treatment allocation $\tilde{\delta}_t$ with continuous $y_t \in [y_l, y_u]$ and binary $d_t \in \{0,1\}$:

$$\tilde{\delta}_t(\boldsymbol{y}^{t-1}, \boldsymbol{d}^{t-1}) = d_t \in \{0,1\} \tag{A.1}$$

for $t = 2, ..., T$ and $\tilde{\delta}_1(x) = d_1 \in \{0,1\}$ with continuous x. This rule may not be a feasible or practical strategy considering the cost of incrementally customizing the allocation based on continuous characteristics $\boldsymbol{y}^{t-1}$. Instead, the planner may want to employ a regime that is only discretely adaptive to the continuous outcomes. This can be achieved by a threshold-crossing allocation rule: for each $t = 2, ..., T$,

$$\tilde{\delta}_t(\boldsymbol{y}^{t-1}, \boldsymbol{d}^{t-1}; \boldsymbol{\gamma}^{t-1}) = \delta_t(1\{y_1 \geq \gamma_1\}, ..., 1\{y_{t-1} \geq \gamma_{t-1}\}, \boldsymbol{d}^{t-1}), \tag{A.2}$$

$$\tilde{\delta}_1(x; \gamma_0) = \delta_1(1\{\gamma_{01}' x \geq \gamma_{02}\}) \tag{A.3}$$

where $\boldsymbol{\gamma}^{t-1} \equiv (\gamma_1, ..., \gamma_{t-1})$ and $\gamma_0 \equiv (\gamma_{01}', \gamma_{02})$ are threshold parameter vectors and $\delta_t(\cdot)$ is the original treatment allocation rule (2.1)–(2.2) based on discrete outcomes and covariates. This threshold-crossing rule is a popular decision rule in practice due to its intuitive form and is considered in earlier theoretical studies such as in Murphy (2003) and Kitagawa and Tetenov (2018). Note that the full regime $(\tilde{\delta}_1(\cdot; \gamma_0), \tilde{\delta}_2(\cdot; \gamma_1), ..., \tilde{\delta}_T(\cdot; \boldsymbol{\gamma}^{T-1}))$ can be characterized by $(\boldsymbol{\delta}(\cdot), \boldsymbol{\gamma})$ where $\boldsymbol{\delta}(\cdot)$ the original regime with discrete outcomes and $\boldsymbol{\gamma} \equiv (\gamma_0, \gamma_1, ..., \gamma_{T-1})$. Therefore, we proceed with latter in the following analysis.

Based on $(\boldsymbol{\delta}, \boldsymbol{\gamma}) \in \mathcal{D} \times \Gamma$, we define the welfare $W_{\boldsymbol{\delta},\boldsymbol{\gamma}}$ analogous to (2.5). For example, $W_{\boldsymbol{\delta},\boldsymbol{\gamma}} = E[Y_T(\boldsymbol{\delta}, \boldsymbol{\gamma})]$ where $Y_T(\boldsymbol{\delta}, \boldsymbol{\gamma})$ is defined as (2.3)–(2.4) but each $\delta_t(\cdot)$ and $\delta_1(\cdot)$ replaced with $\tilde{\delta}_t(\cdot; \boldsymbol{\gamma}^{t-1})$ and $\tilde{\delta}_1(x; \gamma_0)$ defined above, respectively. We wish to find $(\boldsymbol{\delta}^*, \boldsymbol{\gamma}^*)$ that maximize welfare $W_{\boldsymbol{\delta},\boldsymbol{\gamma}}$:

$$(\boldsymbol{\delta}^*,\boldsymbol{\gamma}^*)=\arg\max_{\boldsymbol{\delta}(\cdot)\in\mathcal{D},\boldsymbol{\gamma}\in\Gamma}W_{\boldsymbol{\delta},\boldsymbol{\gamma}}.$$

This maximization problem, equivalently the identification of $(\boldsymbol{\delta}^*,\boldsymbol{\gamma}^*)$, is challenging because $W_{\boldsymbol{\delta},\boldsymbol{\gamma}}$ may not be point identified. Therefore, analogous to the partial identification approach in the main text, we proceed as follows. For a given pair $(\boldsymbol{\delta},\boldsymbol{\gamma})$ and $(\boldsymbol{\delta}',\boldsymbol{\gamma}')$ in $\mathcal{D}\times\Gamma$, let $L(\boldsymbol{\delta},\boldsymbol{\gamma},\boldsymbol{\delta}',\boldsymbol{\gamma}')$ be the lower bound on the welfare gap

$$W_{\boldsymbol{\delta},\boldsymbol{\gamma}}-W_{\boldsymbol{\delta}',\boldsymbol{\gamma}'}.$$

Then, the identified set for $(\boldsymbol{\delta}^*,\boldsymbol{\gamma}^*)$ can be characterized as

$$\{(\boldsymbol{\delta}',\boldsymbol{\gamma}'):L(\boldsymbol{\delta},\boldsymbol{\gamma},\boldsymbol{\delta}',\boldsymbol{\gamma}')\leq 0\text{ for all }(\boldsymbol{\delta},\boldsymbol{\gamma})\in\mathcal{D}\times\Gamma\text{ and }(\boldsymbol{\delta},\boldsymbol{\gamma})\neq(\boldsymbol{\delta}',\boldsymbol{\gamma}')\}.\tag{A.4}$$

Note that, for given $\boldsymbol{\gamma}\in\Gamma$, the maximization of $W_{\boldsymbol{\delta},\boldsymbol{\gamma}}$ with respect to $\boldsymbol{\delta}$ can be solved by establishing the partial ordering of $W_{\boldsymbol{\delta},\boldsymbol{\gamma}}$ with respect to $\boldsymbol{\delta}$. Therefore, for policy, it would also be useful to inspect the partial ordering of $W_{\boldsymbol{\delta},\boldsymbol{\gamma}}$ for any given $\boldsymbol{\gamma}$. This analysis can be done by constructing the DAG for $W_{\boldsymbol{\delta},\boldsymbol{\gamma}}$ using the lower bound $L(\boldsymbol{\delta},\boldsymbol{\delta}';\boldsymbol{\gamma})$ on the welfare gap $W_{\boldsymbol{\delta},\boldsymbol{\gamma}}-W_{\boldsymbol{\delta}',\boldsymbol{\gamma}}$.

We first consider the calculation of $L(\boldsymbol{\delta},\boldsymbol{\delta}';\boldsymbol{\gamma})$ for given $\boldsymbol{\gamma}$, which can be done by solving a sequence of LPs. The challenge is that the continuous outcome variables generate infinite-dimensional programs, which are infeasible to solve in practice. We overcome this challenge by means of approximation. Let $\tilde{\boldsymbol{Y}}_t\equiv\{Y_t(\boldsymbol{d}^t)\}_{\boldsymbol{d}^t}\in[y_l,y_u]^{2^t}$ and $\tilde{\boldsymbol{D}}_t\equiv\{D_t(\boldsymbol{z}^t)\}_{\boldsymbol{z}^t}\in\{0,1\}^{2^t}$ be vectors that constitute $\tilde{S}_t\equiv(\tilde{\boldsymbol{Y}}_t,\tilde{\boldsymbol{D}}_t)$, which is defined analogous to that in the text, and let $y_t(\boldsymbol{d}^t)$ and $d_t(\boldsymbol{z}^t)$ be the realized mappings of $Y_t(\boldsymbol{d}^t)$ and $D_t(\boldsymbol{z}^t)$. Also, define $\tilde{\boldsymbol{Y}}\equiv(\tilde{\boldsymbol{Y}}_1,...,\tilde{\boldsymbol{Y}}_T)$ and $\tilde{\boldsymbol{D}}\equiv(\tilde{\boldsymbol{D}}_1,...,\tilde{\boldsymbol{D}}_T)$. The key element in the formulation is the following conditional cumulative distribution function:

$$q(\tilde{\boldsymbol{y}}, \tilde{\boldsymbol{d}}, x) \equiv \Pr[\tilde{\boldsymbol{Y}} \leq \tilde{\boldsymbol{y}} | \tilde{\boldsymbol{D}} = \tilde{\boldsymbol{d}}, X = x]$$
$$\equiv \Pr[Y_t(\boldsymbol{d}^t) \leq y_t(\boldsymbol{d}^t) \ \forall \boldsymbol{d}^t \text{ and } t | D_t(\boldsymbol{z}^t) = d_t(\boldsymbol{z}^t) \ \forall \boldsymbol{z}^t \text{ and } t, X = x],$$

where "$\leq$" between vectors is understood as element-wise inequalities. The infinite-dimensional object $q(\cdot)$ is the decision variable in the optimization. Let $\mathcal{Q}$ be the infinite-dimensional space of all $q(\cdot,\cdot,\cdot)$'s.

To construct the constraints of the program, consider the distribution of the data:

$$\begin{aligned}
&\Pr[\boldsymbol{Y} \leq \boldsymbol{y}, \boldsymbol{D} = \boldsymbol{d} | \boldsymbol{Z} = \boldsymbol{z}, X = x] \\
&= \Pr[Y_t(\boldsymbol{d}^t) \leq y_t, D_t(\boldsymbol{z}^t) = d_t \ \forall t | X = x] \\
&= \Pr[D_t(\boldsymbol{z}^t) = d_t \ \forall t | X = x] \Pr[Y_t(\boldsymbol{d}^t) \leq y_t \ \forall t | D_t(\boldsymbol{z}^t) = d_t \ \forall t, X = x] \\
&= \sum_{x \in \mathcal{X}} p(x) \sum_{\tilde{\boldsymbol{d}}: d_t(\boldsymbol{z}^t) = d_t \forall t} \Pr[\tilde{\boldsymbol{D}} = \tilde{\boldsymbol{d}} | X = x] \times \\
&\qquad \times \int_{y_l}^{y_u} \cdots \int_{y_l}^{y_u} \left\{ \int_{y_l}^{y_1} \cdots \int_{y_l}^{y_T} q(\tilde{\boldsymbol{y}}, \tilde{\boldsymbol{d}}, x) dy_1(d_1) \cdots d\boldsymbol{y}_T(\boldsymbol{d}^T) \right\} d\tilde{\boldsymbol{y}}_1^- \cdots d\tilde{\boldsymbol{y}}_T^- \\
&\equiv T_{\boldsymbol{y},\boldsymbol{d}|\boldsymbol{z}} \circ q,
\end{aligned}$$

where $\tilde{\boldsymbol{y}}_t^-$ is $\tilde{\boldsymbol{y}}_t$ without $y_t(\boldsymbol{d}^t)$ (with some ambiguity of notation), $\int_{y_l}^{y_u} (\cdot) d\tilde{\boldsymbol{y}}_t^-$ is the corresponding multivariate integral, and $T_{\boldsymbol{y},\boldsymbol{d}|\boldsymbol{z}} : \mathcal{Q} \to \mathbb{R}$ is the operator of $q(\cdot,\cdot,\cdot)$. Then, the continuum of constraints can be written as

$$(T_{\boldsymbol{y}} \circ q)(x) = p(\boldsymbol{y}, x) \qquad \forall (\boldsymbol{y}, x) \in [y_l, y_u]^T \times \mathcal{X},$$

where $T_{\boldsymbol{y}}$ is a vector of operators $T_{\boldsymbol{y},\boldsymbol{d}|\boldsymbol{z}}$'s across $(\boldsymbol{d}, \boldsymbol{z})$ for $q(\cdot,\cdot,x)$ and $p(\boldsymbol{y}, x)$ is a d_p-vector of $\Pr[\boldsymbol{Y} \leq \boldsymbol{y}, \boldsymbol{D} = \boldsymbol{d} | \boldsymbol{Z} = \boldsymbol{z}, X = x]$'s across $(\boldsymbol{d}, \boldsymbol{z})$. Fix $\boldsymbol{\gamma} \in \Gamma$. Since the welfare is also an integral of $q(\cdot,\cdot,\cdot)$, we can write

$$W_{\boldsymbol{\delta},\gamma} = T_{\boldsymbol{\delta},\gamma} \circ q$$

for an operator $T_{\boldsymbol{\delta},\gamma} : \mathcal{Q} \to \mathbb{R}$. Consequently, for $\boldsymbol{\delta}, \boldsymbol{\delta}' \in \mathcal{D}$, we have the following program:

$$\begin{aligned} U(\boldsymbol{\delta},\boldsymbol{\delta}';\boldsymbol{\gamma}) &= \max\nolimits_{q\in\mathcal{Q}}(T_{\boldsymbol{\delta},\gamma} - T_{\boldsymbol{\delta}',\gamma}) \circ q, \\ L(\boldsymbol{\delta},\boldsymbol{\delta}';\boldsymbol{\gamma}) &= \min\nolimits_{q\in\mathcal{Q}}(T_{\boldsymbol{\delta},\gamma} - T_{\boldsymbol{\delta}',\gamma}) \circ q, \end{aligned} \quad s.t. \quad (T_{\boldsymbol{y}} \circ q)(x) = p(\boldsymbol{y}, x) \quad \forall(\boldsymbol{y}, x) \in [y_l, y_u]^T \times \mathcal{X}. \tag{A.5}$$

Because $q(\cdot,\cdot,\cdot) \in \mathcal{Q}$ is an infinite-dimensional object (unlike q in the case of discrete Y_t) and the constraints are also infinite dimensional, the program (A.5) is infinite-dimensional. To gain feasibility, we transform this infinite-dimensional program into a (finite-dimensional) linear program as follows. First, we approximate $q(\cdot,\tilde{\boldsymbol{d}},\cdot)$ using the method of sieve. In particular, the Bernstein polynomial is a suitable choice for sieve basis, because equality and inequality constraints on $q(\cdot,\cdot,\cdot)$ can be easily imposed as equality and inequality constraints on the coefficients of the basis functions. Consider

$$q(\tilde{\boldsymbol{y}}, \tilde{\boldsymbol{d}}, x) \approx \sum_{\boldsymbol{k}=1}^{K} \theta_{\boldsymbol{k}}^{\tilde{\boldsymbol{d}}} b_{\boldsymbol{k}}(\tilde{\boldsymbol{y}}, x),$$

where $b_{\boldsymbol{k}}(\tilde{\boldsymbol{y}}, x) \equiv b_{\boldsymbol{k},K}(\tilde{\boldsymbol{y}}, x)$ is a multivariate Bernstein polynomial with its coefficient $\theta_{\boldsymbol{k}}^{\tilde{\boldsymbol{d}}} \equiv \theta_{\boldsymbol{k},K}^{\tilde{\boldsymbol{d}}} \equiv q(\boldsymbol{k}_1/K, ..., \boldsymbol{k}_T/K, \tilde{\boldsymbol{d}}, k_x/K)$ with the following definition: $\boldsymbol{k} \equiv (\boldsymbol{k}_1, ..., \boldsymbol{k}_T, k_x)$ is a vector of indices where $\boldsymbol{k}_t \equiv \{k_t(\boldsymbol{d}^t)\}_{\boldsymbol{d}^t}$, $\boldsymbol{k}_t/K$ stands for elementwise devision, and $\sum_{\boldsymbol{k}=1}^{K}$ stands for multiple summations, each of which is the sum from each element of $\boldsymbol{k}$ up to K. By replacing $q(\cdot,\cdot,\cdot)$ with this Bernstein expansion in (A.5), we obtain a semi-infinite linear program where the decision variables are simply $\theta_{\boldsymbol{k}}^{\tilde{\boldsymbol{d}}}$ for all $\boldsymbol{k}, \tilde{\boldsymbol{d}}$ and there are the continuum of constraints. Next, we combine the continuum of constraints using the following result: for any measurable function $h : [y_l, y_u]^T \times \mathcal{X} \to \mathbb{R}^{d_p}$, $E\,\|h(\boldsymbol{Y}, X)\| = 0$ if and only if $h(\boldsymbol{y}, x) = 0$ almost everywhere in $[y_l, y_u]^T \times \mathcal{X}$. Therefore, the constraints can be replaced with $E\left\|(T_{\boldsymbol{Y}} \circ q)(X) - p(\boldsymbol{Y}, X)\right\| = 0$. Consequently, we obtain a (finite-dimensional) linear program. We refer the reader to Section 7 of Han and Yang (2022) for the full details of the Bernstein approximation and the transformation of constraints. Finally, an analogous approach can be used to calculate $L(\boldsymbol{\delta}, \boldsymbol{\gamma}, \boldsymbol{\delta}', \boldsymbol{\gamma}')$ for each pair of $(\boldsymbol{\delta}, \boldsymbol{\gamma})$ and $(\boldsymbol{\delta}', \boldsymbol{\gamma}')$ in $\mathcal{D} \times \Gamma$. In practice, we can use grid $\bar{\Gamma} \subseteq \Gamma$ for Γ to characterize the identified set (A.4).

A.2 Time-Varying Covariates

Earlier, we assume for simplicity that potentially endogenous covariates are time-invariant and determined before treatments. Extending the setting to time-varying covariates is straightforward. When covariates are discrete, the allocation rule (2.1) can simply be modified to $\delta_t(\boldsymbol{y}^{t-1}, \boldsymbol{d}^{t-1}, \boldsymbol{x}^{t-1})$ and $\delta_1(x_0)$ where x_t for $t = 2, ..., T$ is time-varying covariates and x_0 is pre-treatment covariates. When time-varying covariates are continuous, the threshold-crossing rule introduced in (A.2) may be modified to $1\{\gamma_{t1} y_t + \gamma_{t2}' x_t \geq \gamma_{t3}\}$ for each $t = 2, ..., T$. That is, for each $t = 2, ..., T$,

$$\tilde{\delta}_t(\boldsymbol{y}^{t-1}, \boldsymbol{d}^{t-1}, \boldsymbol{x}^{t-1}; \boldsymbol{\gamma}^{t-1}) = \delta_t(1\{y_1 + \gamma_{11}' x_1 \geq \gamma_{12}\}, ..., 1\{y_t + \gamma_{t-1,1}' x_t \geq \gamma_{t-1,2}\}, \boldsymbol{d}^{t-1}), \quad \text{(A.6)}$$

$$\tilde{\delta}_1(x_0; \gamma_0) = \delta_1(1\{\gamma_{01}' x_0 \geq \gamma_{02}\}), \quad \text{(A.7)}$$

where $\boldsymbol{\gamma}^{t-1} \equiv (\gamma_1, ..., \gamma_{t-1})$ with $\gamma_t \equiv (\gamma_{t1}', \gamma_{t2})$ and $\gamma_0 \equiv (\gamma_{01}', \gamma_{02})$. With time-varying covariates, the main assumption (Assumption SX) may be modified as follows: $\boldsymbol{Z} \perp (\boldsymbol{Y}(\boldsymbol{d}), \boldsymbol{D}(\boldsymbol{z}))|\boldsymbol{X}, X_0$ where $\boldsymbol{X} = (X_1, ..., X_T)$. The construction of the linear program is very similar to the ones in the earlier cases and therefore omitted.

A.3 Stochastic Regimes

For each $t = 2, ..., T$, define an adaptive *stochastic* treatment rule $\rho_t : \{0,1\}^{t-1} \times \{0,1\}^{t-1} \rightarrow [0,1]$ that allocates the probability of treatment:

$$\rho_t(\boldsymbol{y}^{t-1}, \boldsymbol{r}^{t-1}) = r_t \in [0,1] \tag{A.8}$$

and $\rho_1(x) = r_1 \in [0,1]$. Then, the vector of these ρ_t's is a *dynamic stochastic regime* $\boldsymbol{\rho}(\cdot) \equiv \boldsymbol{\rho}^T(\cdot) \in \mathcal{D}_{stoch}$ where $\mathcal{D}_{stoch}$ is the set of all possible stochastic regimes. Dynamic stochastic regimes are considered in, e.g., Murphy et al. (2001), Murphy (2003), and Manski (2004). A deterministic regime is a special case where $\rho_t(\cdot)$ takes the extreme values of 1 and 0. Therefore, $\mathcal{D} \subset \mathcal{D}_{stoch}$ where $\mathcal{D}$ is the set of deterministic regimes. We define $Y_T(\boldsymbol{\rho}(\cdot))$ with $\boldsymbol{\rho}(\cdot) \in \mathcal{D}_{stoch}$ as the counterfactual outcome $Y_T(\boldsymbol{\delta}(\cdot))$ where the deterministic rule $\delta_t(\cdot) = 1$ is randomly assigned with probability $\rho_t(\cdot)$ and $\delta_t(\cdot) = 0$ otherwise for all $t \leq T$. Finally, define

$$W_{\boldsymbol{\rho}} \equiv \mathbb{E}[Y_T(\boldsymbol{\rho}(\cdot))],$$

where $\mathbb{E}$ denotes an expectation over the counterfactual outcome and the random mechanism defining a rule, and define $\boldsymbol{\rho}^*(\cdot) \equiv \arg\max_{\boldsymbol{\rho}(\cdot) \in \mathcal{D}_{stoch}} W_{\boldsymbol{\rho}}$. The following theorem show that a deterministic regime is achieved as being optimal even though stochastic regimes are allow.

Theorem A.1. *Suppose* $W_{\boldsymbol{\rho}} \equiv \mathbb{E}[Y_T(\boldsymbol{\rho}(\cdot))]$ *for* $\boldsymbol{\rho}(\cdot) \in \mathcal{D}_{stoch}$ *and* $W_{\boldsymbol{\delta}} \equiv E[Y_T(\boldsymbol{\delta}(\cdot))]$ *for* $\boldsymbol{\delta}(\cdot) \in \mathcal{D}$. *It satisfies that*

$$\boldsymbol{\delta}^*(\cdot) \equiv \arg\max_{\boldsymbol{\delta}(\cdot) \in \mathcal{D}} W_{\boldsymbol{\delta}} = \arg\max_{\boldsymbol{\rho}(\cdot) \in \mathcal{D}_{stoch}} W_{\boldsymbol{\rho}}.$$

By the law of iterative expectation, we have

$$\mathbb{E}[Y_T(\boldsymbol{\rho}(\cdot))] = \mathbb{E}\left[\mathbb{E}\left[\mathbb{E}\left[\cdots \mathbb{E}\left[\mathbb{E}[Y_T(\boldsymbol{r})|\boldsymbol{Y}^{T-1}(\boldsymbol{r}^{T-1})]\middle| \boldsymbol{Y}^{T-2}(\boldsymbol{r}^{T-2})\right] \cdots \middle| Y_1(r_1)\right]\middle| X\right]\right], \tag{A.9}$$

where the bridge variables $\boldsymbol{r} = (r_1, ..., r_T)$ satisfy

$$\begin{aligned}
r_1 &= \rho_1(x), \\
r_2 &= \rho_2(Y_1(\rho_1), \rho_1), \\
r_3 &= \rho_3(\boldsymbol{Y}^2(\boldsymbol{\rho}^2), \boldsymbol{\rho}^2), \\
&\vdots \\
r_T &= \rho_T(\boldsymbol{Y}^{T-1}(\boldsymbol{\rho}^{T-1}), \boldsymbol{\rho}^{T-1}).
\end{aligned}$$

Given (A.9), we prove the theorem by showing that the solution $\boldsymbol{\rho}^*(\cdot)$ can be justified by backward induction in a finite-horizon dynamic programming. To illustrate this with deterministic regimes when $T = 2$, we have

$$\delta_2^*(y_1, d_1) = \arg\max_{d_2} E[Y_2(\boldsymbol{d})|Y_1(d_1) = y_1], \tag{A.10}$$

and, by defining $V_2(y_1, d_1) \equiv \max_{d_2} E[Y_2(\boldsymbol{d})|Y_1(d_1) = y_1]$,

$$\delta_1^*(x) = \arg\max_{d_1} E[V_2(Y_1(d_1), d_1)|X = x]. \tag{A.11}$$

Then, $\boldsymbol{\delta}^*(\cdot)$ is equal to the collection of these solutions: $\boldsymbol{\delta}^*(\cdot) = (\delta_1^*, \delta_2^*(\cdot))$.

Proof. First, given (A.9), the optimal stochastic rule in the final period can be defined as

$$\rho_T^*(\boldsymbol{y}^{T-1}, \boldsymbol{r}^{T-1}) \equiv \arg\max_{r_T \in [0,1]} \mathbb{E}[Y_T(\boldsymbol{r})|\boldsymbol{Y}^{T-1}(\boldsymbol{r}^{T-1}) = \boldsymbol{y}^{T-1}].$$

Define a value function at period T as $V_T(\boldsymbol{y}^{T-1}, \tilde{\boldsymbol{d}}^{T-1}) \equiv \max_{r_T} \mathbb{E}[Y_T(\boldsymbol{r})|\boldsymbol{Y}^{T-1}(\boldsymbol{r}^{T-1}) = \boldsymbol{y}^{T-1}]$. Similarly, for each $t = 1, ..., T-1$, let

$$\rho_t^*(\boldsymbol{y}^{t-1}, \boldsymbol{r}^{t-1}) \equiv \arg\max_{r_t \in [0,1]} \mathbb{E}[V_{t+1}(\boldsymbol{Y}^t(\boldsymbol{r}^t), \boldsymbol{r}^t)|\boldsymbol{Y}^{t-1}(\boldsymbol{r}^{t-1}) = \boldsymbol{y}^{t-1}]$$

and $V_t(\boldsymbol{y}^{t-1}, \boldsymbol{r}^{t-1}) \equiv \max_{r_t} \mathbb{E}[V_{t+1}(\boldsymbol{Y}^t(\boldsymbol{r}^t), \boldsymbol{r}^t)|\boldsymbol{Y}^{t-1}(\boldsymbol{r}^{t-1}) = \boldsymbol{y}^{t-1}]$. Finally, let

$$\rho_1^*(x) \equiv \arg\max_{r_1 \in [0,1]} \mathbb{E}[V_2(Y_1(r_1), r_1)|X = x].$$

Then, $\boldsymbol{\rho}^*(\cdot) = (\rho_1^*(\cdot), ..., \rho_T^*(\cdot))$. Since $\{0,1\} \subset [0,1]$, the same argument can apply for the deterministic regime using the current framework but each maximization domain being $\{0,1\}$. This analogously defines $\delta_t^*(\cdot) \in \{0,1\}$ for all t, and then $\boldsymbol{\delta}^*(\cdot) = (\delta_1^*(\cdot), ..., \delta_T^*(\cdot))$, similarly as in Murphy (2003).

Now, for the maximization problems above, let $\tilde{W}_t(\boldsymbol{r}^t, \boldsymbol{y}^{t-1})$ represent the objective function at t for $2 \leq t \leq T$ with $\tilde{W}_1(r_1)$ for $t = 1$. By the definition of the stochastic regime, it satisfies that

$$\begin{aligned}
\tilde{W}_t(\boldsymbol{r}^t, \boldsymbol{y}^{t-1}) &= r_t W_t(1, \boldsymbol{r}^{t-1}, \boldsymbol{y}^{t-1}) + (1 - r_t) W_t(0, \boldsymbol{r}^{t-1}, \boldsymbol{y}^{t-1}) \\
&= r_t \left\{ W_t(1, \boldsymbol{r}^{t-1}, \boldsymbol{y}^{t-1}) - W_t(0, \boldsymbol{r}^{t-1}, \boldsymbol{y}^{t-1}) \right\} + W_t(0, \boldsymbol{r}^{t-1}, \boldsymbol{y}^{t-1}).
\end{aligned}$$

Therefore, $W_t(1, \boldsymbol{r}^{t-1}, \boldsymbol{y}^{t-1}) \geq W_t(0, \boldsymbol{r}^{t-1}, \boldsymbol{y}^{t-1})$ or $1 = \arg\max_{r_t \in \{0,1\}} \tilde{W}_t(\boldsymbol{r}^t, \boldsymbol{y}^{t-1})$ if and only if $1 = \arg\max_{r_t \in [0,1]} \tilde{W}_t(\boldsymbol{r}^t, \boldsymbol{y}^{t-1})$. Symmetrically, $0 = \arg\max_{r_t \in \{0,1\}} \tilde{W}_t(\boldsymbol{r}^t, \boldsymbol{y}^{t-1})$ if and only if $0 = \arg\max_{r_t \in [0,1]} \tilde{W}_t(\boldsymbol{r}^t, \boldsymbol{y}^{t-1})$. This implies that $\rho_t^*(\cdot) = \delta_t^*(\cdot)$ for all $t = 1, ..., T$, which proves the theorem. □

B Matrices in Section 3.3

We show how to construct matrices A_k and B in (3.2) and (3.4) for the linear programming (3.6). The construction of A_k and B uses the fact that any linear functional of $\Pr[\boldsymbol{Y}(\boldsymbol{d}) = \boldsymbol{y}|X = x]$ or $\Pr[\boldsymbol{Y}(\boldsymbol{d}) = \boldsymbol{y}, \boldsymbol{D}(\boldsymbol{z}) = \boldsymbol{d}|X = x]$ can be characterized as a linear combination of $q_s(x)$. Although the notation of this section can be somewhat heavy, if one is committed to the use of linear programming instead of an analytic solution, most of the derivation can be systematically reproduced in a standard software, such as MATLAB and Python.

Consider B first. By Assumption SX, we have

$$\begin{aligned} p_{\boldsymbol{y},\boldsymbol{d}|\boldsymbol{z},x} &= \Pr[\boldsymbol{Y}(\boldsymbol{d}) = \boldsymbol{y}, \boldsymbol{D}(\boldsymbol{z}) = \boldsymbol{d} | X = x] \\ &= \Pr[S : Y_t(\boldsymbol{d}^t) = y_t, D_t(\boldsymbol{z}^t) = d_t \ \forall t | X = x] \\ &= \sum_{s \in \mathcal{S}_{\boldsymbol{y},\boldsymbol{d}|\boldsymbol{z}}} q_s(x), \end{aligned} \tag{B.1}$$

where $\mathcal{S}_{\boldsymbol{y},\boldsymbol{d}|\boldsymbol{z}} \equiv \{S = \beta(\tilde{\boldsymbol{S}}) : Y_t(\boldsymbol{d}^t) = y_t, D_t(\boldsymbol{z}^t) = d_t \ \forall t\}$, $\tilde{\boldsymbol{S}} \equiv (\tilde{S}_1, ..., \tilde{S}_T)$ with $\tilde{S}_t \equiv (\{Y_t(\boldsymbol{d}^t)\}_{\boldsymbol{d}^t}, \{D_t(\boldsymbol{z}^t)\}_{\boldsymbol{z}^t})$, and $\beta(\cdot)$ is a one-to-one map that transforms a binary sequence into a decimal value. Then, for a $1 \times dim(q(x))$ vector $B_{\boldsymbol{y},\boldsymbol{d}|\boldsymbol{z}}$ of ones and zeros,

$$p_{\boldsymbol{y},\boldsymbol{d}|\boldsymbol{z},x} = \sum_{s \in \mathcal{S}_{\boldsymbol{y},\boldsymbol{d}|\boldsymbol{z}}} q_s(x) = B_{\boldsymbol{y},\boldsymbol{d}|\boldsymbol{z}} q(x)$$

and the $dim(p_x) \times dim(q(x))$ matrix B_0 vertically stacks $B_{\boldsymbol{y},\boldsymbol{d}|\boldsymbol{z}}$ so that $p_x = B_0 q(x)$ where $p_x \equiv \{p_{\boldsymbol{y},\boldsymbol{d}|\boldsymbol{z},x}\}_{\boldsymbol{y},\boldsymbol{d},\boldsymbol{z}}$ except redundant elements. Finally, we have $p = Bq$ where $p \equiv (p'_{x_1}, ..., p'_{x_L})'$, $B \equiv \begin{bmatrix} B_0 & & \\ & \ddots & \\ & & B_0 \end{bmatrix}$, and $q = (q(x_1)', ..., q(x_L)')'$ with $\mathcal{X} \equiv \{x_1, ..., x_L\}$.

For A_k, recall $W_{\boldsymbol{\delta}_k}$ is a linear functional of $q_{\boldsymbol{\delta}_k}(\boldsymbol{y}) \equiv \Pr[\boldsymbol{Y}(\boldsymbol{\delta}_k(\cdot)) = \boldsymbol{y}]$. For given $\boldsymbol{\delta}(\cdot)$, by repetitively applying the law of iterated expectation, we can show

$$\begin{aligned} &\Pr[\boldsymbol{Y}(\boldsymbol{\delta}(\cdot)) = \boldsymbol{y}] \\ =& \Pr[Y_T(\boldsymbol{d}) = y_T | \boldsymbol{Y}^{T-1}(\boldsymbol{d}^{T-1}) = \boldsymbol{y}^{T-1}] \\ &\times \Pr[Y_{T-1}(\boldsymbol{d}^{T-1}) = y_{T-1} | \boldsymbol{Y}^{T-2}(\boldsymbol{d}^{T-2}) = \boldsymbol{y}^{T-2}] \times \cdots \times \Pr[Y_1(d_1) = y_1], \end{aligned} \tag{B.2}$$

where, because of the appropriate conditioning in (B.2), the bridge variables $\boldsymbol{d} = (d_1, ..., d_T)$

satisfies

$$\begin{aligned}
d_1 &= \delta_1, \\
d_2 &= \delta_2(y_1, d_1), \\
d_3 &= \delta_3(\boldsymbol{y}^2, \boldsymbol{d}^2), \\
&\vdots \\
d_T &= \delta_T(\boldsymbol{y}^{T-1}, \boldsymbol{d}^{T-1}).
\end{aligned}$$

Therefore, (B.2) can be viewed as a linear functional of $\Pr[\boldsymbol{Y}(\boldsymbol{d}) = \boldsymbol{y}]$. To illustrate, when $T = 2$, the welfare defined as the average counterfactual terminal outcome satisfies

$$\begin{aligned}
E[Y_T(\boldsymbol{\delta}(\cdot))] &= \sum_{y_1} \Pr[Y_2(\delta_1, \delta_2(Y_1(\delta_1), \delta_1)) = 1 | Y_1(\delta_1) = y_1] \Pr[Y_1(\delta_1) = y_1] \\
&= \sum_{y_1} \Pr[Y_2(\delta_1, \delta_2(y_1, \delta_1)) = 1, Y_1(\delta_1) = y_1]. \qquad \text{(B.3)}
\end{aligned}$$

Then, for a chosen $\boldsymbol{\delta}(\cdot)$, the values $\delta_1 = d_1$ and $\delta_2(y_1, \delta_1) = d_2$ at which $Y_2(\delta_1, \delta_2(y_1, \delta_1))$ and $Y_1(\delta_1)$ are defined is given in Table 1 as shown in the main text. Therefore, $E[Y_2(\boldsymbol{\delta}(\cdot))]$ can be written as a linear functional of $\Pr[Y_2(d_1, d_2) = y_2, Y_1(d_1) = y_1]$.

Now, define a linear functional $h_k(\cdot)$ that maps $\Pr[\boldsymbol{Y}(\boldsymbol{d}) = \boldsymbol{y}]$ into $\Pr[\boldsymbol{Y}(\boldsymbol{\delta}_k(\cdot)) = \boldsymbol{y}]$ according to (B.2). But note that $\Pr[\boldsymbol{Y}(\boldsymbol{d}) = \boldsymbol{y}] = \sum_{s \in \mathcal{S}_{\boldsymbol{y},\boldsymbol{d}}} q_s$ by

$$\begin{aligned}
&\Pr[\boldsymbol{Y}(\boldsymbol{d}) = \boldsymbol{y}] \\
&= \Pr[S : Y_t(\boldsymbol{d}^t) = y_t \quad \forall t] \\
&= \sum_{s \in \mathcal{S}_{\boldsymbol{y},\boldsymbol{d}}} q_s, \qquad \text{(B.4)}
\end{aligned}$$

where $\mathcal{S}_{\boldsymbol{y},\boldsymbol{d}} \equiv \{S = \beta(\tilde{\boldsymbol{S}}) : Y_t(\boldsymbol{d}^t) = y_t \ \forall t\}$. Consequently, we have

$$\begin{aligned} W_k = f(q_{\boldsymbol{\delta}_k}) &= f(\Pr[\boldsymbol{Y}(\boldsymbol{\delta}_k(\cdot)) = \cdot]) \\ &= f \circ h_k(\Pr[\boldsymbol{Y}(\cdot) = \cdot, \boldsymbol{D}(\boldsymbol{z}) = \cdot]), \\ &= f \circ h_k\left(\sum_{s \in \mathcal{S}_{\cdot,\cdot|\boldsymbol{z}}} q_s\right) \equiv A_k q. \end{aligned}$$

To continue the illustration (3.3) in the main text, note that

$$\Pr[\boldsymbol{Y}(1,1) = (1,1)] = \Pr[S : Y_1(1) = 1, Y_2(1,1) = 1] = \sum_{s \in \mathcal{S}_{11}} q_s,$$

where $\mathcal{S}_{11} \equiv \{S = \beta(\tilde{S}_1, \tilde{S}_2) : Y_1(1) = 1, Y_2(1,1) = 1\}$. Similarly, we have

$$\Pr[\boldsymbol{Y}(1,1) = (0,1)] = \Pr[S : Y_1(1) = 0, Y_2(1,1) = 1] = \sum_{s \in \mathcal{S}_{01}} q_s,$$

where $\mathcal{S}_{01} \equiv \{S = \beta(\tilde{S}_1, \tilde{S}_2) : Y_1(1) = 0, Y_2(1,1) = 1\}$.

C Proofs

C.1 Proof of Theorem 3.1

Let $\mathcal{Q}_p \equiv \{q : Bq = p\} \cap \mathcal{Q}$ be the feasible set. To prove part (i), first note that the sharp DAG can be explicitly defined as $G(\mathcal{K}, \mathcal{E}_p)$ with

$$\mathcal{E}_p \equiv \{(k, k') \in \mathcal{K} : A_k q > A_{k'} q \text{ for all } q \in \mathcal{Q}_p\}.$$

Here, $A_k q > A_{k'} q$ for all $q \in \mathcal{Q}_p$ if and only if $L_{k,k'} > 0$ as $L_{k,k'}$ is the sharp lower bound of $(A_k - A_{k'})q$ in (3.6). The latter is because the feasible set $\{q : Bq = p \text{ and } q \in \mathcal{Q}\}$ is convex and thus $\{\Delta_{k,k'} q : Bq = p \text{ and } q \in \mathcal{Q}\}$ is convex, which implies that any point between

$[L_{k,k'}, U_{k,k'}]$ is attainable.

To prove part (ii), it is helpful to note that $\mathcal{D}_p^*$ in (3.5) can be equivalently defined as

$$\begin{aligned}\mathcal{D}_p^* &\equiv \{\boldsymbol{\delta}_{k'}(\cdot) : \nexists k \in \mathcal{K} \text{ such that } A_k q > A_{k'} q \text{ for all } q \in \mathcal{Q}_p\} \\ &= \{\boldsymbol{\delta}_{k'}(\cdot) : A_k q \leq A_{k'} q \text{ for all } k \in \mathcal{K} \text{ and some } q \in \mathcal{Q}_p\}.\end{aligned}$$

Let $\tilde{\mathcal{D}}_p^* \equiv \{\boldsymbol{\delta}_{k'}(\cdot) : \nexists k \in \mathcal{K}$ such that $L_{k,k'} > 0$ and $k \neq k'\}$. First, we prove that $\mathcal{D}_p^* \subset \tilde{\mathcal{D}}_p^*$. Note that

$$\mathcal{D} \backslash \tilde{\mathcal{D}}_p^* = \{\boldsymbol{\delta}_{k'} : L_{k,k'} > 0 \text{ for some } k \neq k'\}.$$

Suppose $\boldsymbol{\delta}_{k'} \in \mathcal{D} \backslash \tilde{\mathcal{D}}_p^*$. Then, for some $k \neq k'$, $(A_k - A_{k'})q \geq L_{k,k'} > 0$ for all $q \in \mathcal{Q}_p$. Therefore, for such k, $A_k q > A_{k'} q$ for all $q \in \mathcal{Q}_p$, and thus $\boldsymbol{\delta}_{k'} \notin \mathcal{D}_p^* \equiv \{\arg\max_{\boldsymbol{\delta}_k} A_k q : q \in \mathcal{Q}_p\}$.

Now, we prove that $\tilde{\mathcal{D}}_p^* \subset \mathcal{D}_p^*$. Suppose $\boldsymbol{\delta}_{k'} \in \tilde{\mathcal{D}}_p^*$. Then $\nexists k \neq k'$ such that $L_{k,k'} > 0$. Equivalently, for any given $k \neq k'$, either (a) $U_{k,k'} \leq 0$ or (b) $L_{k,k'} < 0 < U_{k,k'}$. Consider (a), which is equivalent to $\max_{q \in \mathcal{Q}_p} (A_k - A_{k'})q \leq 0$. This implies that $A_k q \leq A_{k'} q$ for all $q \in \mathcal{Q}_p$. Consider (b), which is equivalent to $\min_{q \in \mathcal{Q}_p} (A_k - A_{k'})q < 0 < \max_{q \in \mathcal{Q}_p} (A_k - A_{k'})q$. This implies that $\exists q \in \mathcal{Q}_p$ such that $A_k q = A_{k'} q$. Combining these implications of (a) and (b), it should be the case that $\exists q \in \mathcal{Q}_p$ such that, for all $k \neq k'$, $A_{k'} q \geq A_k q$. Therefore, $\boldsymbol{\delta}_k \in \mathcal{D}_p^*$.
□

C.2 Alternative Characterization of the Identified Set

Given the DAG, the identified set of $\boldsymbol{\delta}^*(\cdot)$ can also be obtained as the collection of initial vertices of all the directed paths of the DAG. For a DAG $G(\mathcal{K}, \mathcal{E})$, a *directed path* is a subgraph $G(\mathcal{K}_j, \mathcal{E}_j)$ $(1 \leq j \leq J \leq 2^{|\mathcal{K}|})$ where $\mathcal{K}_j \subset \mathcal{K}$ is a totally ordered set with initial vertex $\tilde{k}_{j,1}$.[1] In stating our main theorem, we make it explicit that the DAG calculated by the linear programming is a function of the data distribution p.

Theorem C.1. *Suppose Assumptions SX and B hold. Then, $\mathcal{D}_p^*$ defined in (3.5) satisfies*

$$\mathcal{D}_p^* = \{\boldsymbol{\delta}_{\tilde{k}_{j,1}}(\cdot) \in \mathcal{D} : 1 \le j \le J\}, \tag{C.1}$$

where $\tilde{k}_{j,1}$ is the initial vertex of the directed path $G(\mathcal{K}_{p,j}, \mathcal{E}_{p,j})$ of $G(\mathcal{K}, \mathcal{E}_p)$.

Proof. Let $\tilde{\mathcal{D}}^* \equiv \{\boldsymbol{\delta}_{\tilde{k}_{j,1}}(\cdot) \in \mathcal{D} : 1 \le j \le J\}$. First, note that since $\tilde{k}_{j,1}$ is the initial vertex of directed path j, it should be that $W_{\tilde{k}_{j,1}} \ge W_{\tilde{k}_{j,m}}$ for any $\tilde{k}_{j,m}$ in that path by definition. We begin by supposing $\mathcal{D}_p^* \supset \tilde{\mathcal{D}}^*$. Then, there exist $\boldsymbol{\delta}^*(\cdot; q) = \arg\max_{\boldsymbol{\delta}_k(\cdot) \in \mathcal{D}} A_k q$ for some q that satisfies $Bq = p$ and $q \in \mathcal{Q}$, but which is not the initial vertex of any directed path. Such $\boldsymbol{\delta}^*(\cdot; q)$ cannot be other (non-initial) vertices of any paths as it is contradiction by the definition of $\boldsymbol{\delta}^*(\cdot; q)$. But the union of all directed paths is equal to the original DAG, therefore there cannot exist such $\boldsymbol{\delta}^*(\cdot; q)$.

Now suppose $\mathcal{D}_p^* \subset \tilde{\mathcal{D}}^*$. Then, there exists $\boldsymbol{\delta}_{\tilde{k}_{j,1}}(\cdot) \ne \boldsymbol{\delta}^*(\cdot; q) = \arg\max_{\boldsymbol{\delta}_k(\cdot) \in \mathcal{D}} A_k q$ for some q that satisfies $Bq = p$ and $q \in \mathcal{Q}$. This implies that $W_{\tilde{k}_{j,1}} < W_{\tilde{k}}$ for some $\tilde{k}$. But $\tilde{k}$ should be a vertex of the same directed path (because $W_{\tilde{k}_{j,1}}$ and $W_{\tilde{k}}$ are ordered), but then it is contradiction as $\tilde{k}_{j,1}$ is the initial vertex. Therefore, $\mathcal{D}_p^* = \tilde{\mathcal{D}}^*$. □

C.3 Proof of Theorem F.1

Given Theorem C.1, proving $\tilde{\mathcal{D}}^* = \{\boldsymbol{\delta}_{k_{l,1}}(\cdot) : 1 \le l \le L_G\}$ will suffice. Recall $\tilde{\mathcal{D}}^* \equiv \{\boldsymbol{\delta}_{\tilde{k}_{j,1}}(\cdot) \in \mathcal{D} : 1 \le j \le J\}$ where $\tilde{k}_{j,1}$ is the initial vertex of the directed path $G(\mathcal{K}_{p,j}, \mathcal{E}_{p,j})$. When all topological sorts are singletons, the proof is trivial so we rule out this possibility. Suppose $\tilde{\mathcal{D}}^* \supset \{\boldsymbol{\delta}_{k_{l,1}}(\cdot) : 1 \le l \le L_G\}$. Then, for some l, there should exist $\boldsymbol{\delta}_{k_{l,m}}(\cdot)$ for some $m \ne 1$ that is contained in $\tilde{\mathcal{D}}^*$ but not in $\{\boldsymbol{\delta}_{k_{l,1}}(\cdot) : 1 \le l \le L_G\}$, i.e., that satisfies either (i) $W_{k_{l,1}} > W_{k_{l,m}}$ or (ii) $W_{k_{l,1}}$ and $W_{k_{l,m}}$ are incomparable and thus either $W_{k_{l',1}} > W_{k_{l,m}}$ for some $l' \ne l$ or

[1] For example, in Figure 1(a), there are two directed paths ($J = 2$) with $V_1 = \{1, 2, 3\}$ ($\tilde{k}_{1,1} = 1$) and $V_2 = \{2, 3, 4\}$ ($\tilde{k}_{2,1} = 4$).

$W_{k_{l,m}}$ is a singleton in another topological sort. Consider case (i). If $\boldsymbol{\delta}_{k_{l,1}}(\cdot) \in \mathcal{D}_j$ for some j, then it should be that $\boldsymbol{\delta}_{k_{l,m}}(\cdot) \in \mathcal{D}_j$ as $\boldsymbol{\delta}_{k_{l,1}}(\cdot)$ and $\boldsymbol{\delta}_{k_{l,m}}(\cdot)$ are comparable in terms of welfare, but then $\boldsymbol{\delta}_{k_{l,m}}(\cdot) \in \tilde{\mathcal{D}}^*$ contradicts the fact that $\boldsymbol{\delta}_{k_{l,1}}(\cdot)$ the initial vertex of the topological sort. Consider case (ii). The singleton case is trivially rejected since if the topological sort a singleton, then $\boldsymbol{\delta}_{k_{l,m}}(\cdot)$ should have been already in $\{\boldsymbol{\delta}_{k_{l,1}}(\cdot) : 1 \leq l \leq L_G\}$. In the other case, since the two welfares are not comparable, it should be that $\boldsymbol{\delta}_{k_{l,m}}(\cdot) \in \mathcal{D}_{j'}$ for $j' \neq j$. But $\boldsymbol{\delta}_{k_{l,m}}(\cdot)$ cannot be the one that delivers the largest welfare since $W_{k_{l',1}} > W_{k_{l,m}}$ where $\boldsymbol{\delta}_{k_{l',1}}(\cdot)$. Therefore $\boldsymbol{\delta}_{k_{l,m}}(\cdot) \in \tilde{\mathcal{D}}^*$ is contradiction. Therefore there is no element in $\tilde{\mathcal{D}}^*$ that is not in $\{\boldsymbol{\delta}_{k_{l,1}}(\cdot) : 1 \leq l \leq L_G\}$.

Now suppose $\tilde{\mathcal{D}}^* \subset \{\boldsymbol{\delta}_{k_{l,1}}(\cdot) : 1 \leq l \leq L_G\}$. Then for l such that $\boldsymbol{\delta}_{k_{l,1}}(\cdot) \notin \tilde{\mathcal{D}}^*$, either $W_{k_{l,1}}$ is a singleton or $W_{k_{l,1}}$ is an element in a non-singleton topological sort. But if it is a singleton, then it is trivially totally ordered and is the maximum welfare, and thus $\boldsymbol{\delta}_{k_{l,1}}(\cdot) \notin \tilde{\mathcal{D}}^*$ is contradiction. In the other case, if $W_{k_{l,1}}$ is a maximum welfare, then $\boldsymbol{\delta}_{k_{l,1}}(\cdot) \notin \tilde{\mathcal{D}}^*$ is contradiction. If it is not a maximum welfare, then it should be a maximum in another topological sort, which is contradiction in either case of being contained in $\{\boldsymbol{\delta}_{k_{l,1}}(\cdot) : 1 \leq l \leq L_G\}$ or not. This concludes the proof that $\tilde{\mathcal{D}}^* = \{\boldsymbol{\delta}_{k_{l,1}}(\cdot) : 1 \leq l \leq L_G\}$. $\square$

D Incorporating Additional Identifying Assumptions

To incorporate additional identifying assumptions in Section 3.5, we extend the main framework of Sections 3.3–3.4. Suppose h is a $d_q \times 1$ vector of ones and zeros, where zeros are imposed by given identifying assumptions. Introduce $d_q \times d_q$ diagonal matrix $H = diag(h)$. Then, we can define a space for $\bar{q} \equiv Hq$ as

$$\bar{\mathcal{Q}} \equiv \{\bar{q} : \sum_s \bar{q}_s(x) = 1 \ \forall x \text{ and } \bar{q}_s(x) \geq 0 \ \forall s, x\}. \tag{D.1}$$

Note that the dimension of this space is smaller than the dimension of $\mathcal{Q}$ if h contains zeros. Then we can modify (3.2) and (3.4) as

$$B\bar{q} = p,$$

$$W_k = A_k\bar{q},$$

respectively. Let $\boldsymbol{\delta}^*(\cdot;\bar{q}) \equiv \arg\max_{\boldsymbol{\delta}_k(\cdot)\in\mathcal{D}} W_k = A_k\bar{q}$. Then, the identified set with the identifying assumptions coded in h is defined as

$$\bar{\mathcal{D}}_p^* \equiv \{\boldsymbol{\delta}^*(\cdot;\bar{q}) : B\bar{q} = p \text{ and } \bar{q} \in \mathcal{Q}\} \subset \mathcal{D}, \tag{D.2}$$

which is assumed to be empty when $B\bar{q} \neq p$. Importantly, the latter occurs when any of the identifying assumptions are misspecified. Note that H is idempotent. Define $\bar{\Delta} \equiv \Delta H$ and $\bar{B} \equiv BH$. Then $\Delta\bar{q} = \bar{\Delta}\bar{q}$ and $B\bar{q} = \bar{B}\bar{q}$. Therefore, to generate the DAG and characterize the identified set, Theorem 3.1 can be modified by replacing q, B and Δ with $\bar{q}$, $\bar{B}$ and $\bar{\Delta}$, respectively.

Then, for example, we can incorporate Assumption M1 by choosing appropriate h. Recall $\tilde{S}_t \equiv (\{Y_t(\boldsymbol{d}^t)\}, \{D_t(\boldsymbol{z}^t)\}) \in \{0,1\}^{2^t} \times \{0,1\}^{2^t}$ and $\mathcal{S}_{\boldsymbol{y},\boldsymbol{d}|\boldsymbol{z}} \equiv \{S = \beta(\tilde{\boldsymbol{S}}) : Y_t(\boldsymbol{d}^t) = y_t, D_t(\boldsymbol{z}^t) = d_t \ \forall t\}$ given $(\boldsymbol{y}, \boldsymbol{d}, \boldsymbol{z})$. For example, the no-defier assumption can be incorporated in h by having $h_s = 0$ for $s \in \{S \in \mathcal{S}_{\boldsymbol{y},\boldsymbol{d}|\boldsymbol{z}} : D_t(\boldsymbol{z}^{t-1}, 1) = 0 \text{ and } D_t(\boldsymbol{z}^{t-1}, 0) = 1 \ \forall t\}$ and $h_s = 1$ otherwise.

The following lemmas establish the equivalence between Assumptions M1 and M2 and corresponding threshold-crossing models.

Lemma D.1. *Suppose Assumption SX holds and* $\Pr[D_t = 1|\boldsymbol{Y}^{t-1}, \boldsymbol{D}^{t-1}, \boldsymbol{Z}^t, X]$ *is a non-trivial function of* Z_t. *Assumption M1 is equivalent to (3.9) being satisfied conditional on* $(\boldsymbol{Y}^{t-1}, \boldsymbol{D}^{t-1}, \boldsymbol{Z}^{t-1}, X)$ *for each* t.

Lemma D.2. *Suppose Assumption SX holds,* $\Pr[D_t = 1|\boldsymbol{Y}^{t-1}, \boldsymbol{D}^{t-1}, \boldsymbol{Z}^t, X]$ *is a non-trivial function of* Z_t, *and* $\Pr[Y_t = 1|\boldsymbol{Y}^{t-1}, \boldsymbol{D}^t, X]$ *is a non-trivial function of* D_t. *Assumption M2* *is equivalent to (3.10)–(3.11) being satisfied conditional on* $(\boldsymbol{Y}^{t-1}, \boldsymbol{D}^{t-1}, \boldsymbol{Z}^{t-1}, X)$ *for each* t.

D.1 Proof of Lemma D.1

Conditional on $(\boldsymbol{Y}^{t-1}, \boldsymbol{D}^{t-1}, \boldsymbol{Z}^{t-1}, X) = (\boldsymbol{y}^{t-1}, \boldsymbol{d}^{t-1}, \boldsymbol{z}^{t-1}, x)$, it is easy to show that (3.9) implies Assumption M1. Suppose $\pi_t(\boldsymbol{y}^{t-1}, \boldsymbol{d}^{t-1}, \boldsymbol{z}^{t-1}, 1, x) > \pi_t(\boldsymbol{y}^{t-1}, \boldsymbol{d}^{t-1}, \boldsymbol{z}^{t-1}, 1, x)$ as $\pi_t(\cdot)$ is a nontrivial function of Z_t. Then, we have

$$1\{\pi_t(\boldsymbol{y}^{t-1}, \boldsymbol{d}^{t-1}, \boldsymbol{z}^{t-1}, 1, x) \geq V_t\} \geq 1\{\pi_t(\boldsymbol{y}^{t-1}, \boldsymbol{d}^{t-1}, \boldsymbol{z}^{t-1}, 0, x) \geq V_t\}$$

w.p.1, or equivalently, $D_t(\boldsymbol{z}^{t-1}, 1) \geq D_t(\boldsymbol{z}^{t-1}, 0)$ w.p.1. Suppose $\pi_t(\boldsymbol{y}^{t-1}, \boldsymbol{d}^{t-1}, \boldsymbol{z}^{t-1}, 1, x) < \pi_t(\boldsymbol{y}^{t-1}, \boldsymbol{d}^{t-1}, \boldsymbol{z}^{t-1}, 1, x)$. Then, by a parallel argument, $D_t(\boldsymbol{z}^{t-1}, 1) \leq D_t(\boldsymbol{z}^{t-1}, 0)$ w.p.1.

Now, we show that Assumption M1 implies (3.9) conditional on $(\boldsymbol{Y}^{t-1}, \boldsymbol{D}^{t-1}, \boldsymbol{Z}^{t-1}, X)$. For each t, Assumption SX implies $Y_t(\boldsymbol{d}^t), D_t(\boldsymbol{z}^t) \perp \boldsymbol{Z}^t | (\boldsymbol{Y}^{t-1}(\boldsymbol{d}^{t-1}), \boldsymbol{D}^{t-1}(\boldsymbol{z}^{t-1}), \boldsymbol{Z}^{t-1}, X)$, which in turn implies the following conditional independence:

$$Y_t(\boldsymbol{d}^t), D_t(\boldsymbol{z}^t) \perp \boldsymbol{Z}^t | (\boldsymbol{Y}^{t-1}, \boldsymbol{D}^{t-1}, \boldsymbol{Z}^{t-1}, X). \tag{D.3}$$

Conditional on $(\boldsymbol{Y}^{t-1}, \boldsymbol{D}^{t-1}, \boldsymbol{Z}^{t-1}, X)$, (3.9) and (D.3) correspond to Assumption S-1 in Vytlacil (2002). Assumption R(i) and (D.3) correspond to Assumption L-1, and Assumption M1 corresponds to Assumption L-2 in Vytlacil (2002). Therefore, the desired result follows by Theorem 1 of Vytlacil (2002). □

D.2 Proof of Lemma D.2

We are remained to prove that, conditional on $(\boldsymbol{Y}^{t-1}, \boldsymbol{D}^{t-1}, X)$, (3.10) is equivalent to the second part of Assumption M2. But this proof is analogous to the proof of Lemma D.1 by replacing the roles of D_t and Z_t with those of Y_t and D_t, respectively. Therefore, we have the desired result. □

E Numerical Studies

We conduct numerical exercises to illustrate (i) the theoretical results developed in Sections 3.1–3.4, (ii) the role of the assumptions introduced in Section 3.5, and (iii) the overall computational scale of the problem. For $T = 2$, we consider the following data-generating process:

$$D_{i1} = 1\{\pi_1 Z_{i1} + \alpha_i + v_{i1} \geq 0\}, \tag{E.1}$$

$$Y_{i1} = 1\{\mu_1 D_{i1} + \alpha_i + e_{i1} \geq 0\}, \tag{E.2}$$

$$D_{i2} = 1\{\pi_{21} Y_{i1} + \pi_{22} D_{i1} + \pi_{23} Z_{i2} + \alpha_i + v_{i2} \geq 0\}, \tag{E.3}$$

$$Y_{i2} = 1\{\mu_{21} Y_{i1} + \mu_{22} D_{i2} + \alpha_i + e_{i2} \geq 0\}, \tag{E.4}$$

where $(v_1, e_1, v_2, e_2, \alpha)$ are mutually independent and jointly normally distributed, the endogeneity of D_{i1} and D_{i2} as well as the serial correlation of the unobservables are captured by the individual effect α_i, and (Z_1, Z_2) are Bernoulli, independent of $(v_1, e_1, v_2, e_2, \alpha)$. Notice that the process is intended to satisfy Assumptions SX, K, M1, and M2. We consider a data-generating process where all the coefficients in (E.1)–(E.4) take positive values. In this exercise, we consider the welfare $W_k = E[Y_2(\boldsymbol{\delta}_k(\cdot))]$.

We consider eight possible regimes shown in Table 1 (i.e., $|\mathcal{D}| = |\mathcal{K}| = 8$). We calculate the lower and upper bounds $(L_{k,k'}, U_{k,k'})$ on the welfare gap $W_k - W_{k'}$ for all pairs $k, k' \in \{1, ..., 8\}$ $(k < k')$. This is to illustrate the role of assumptions in improving the bounds. We conduct the bubble sort, which makes $\binom{8}{2} = 28$ pair-wise comparisons, resulting in 28×2 linear programs to run.[2] As the researcher, we maintain Assumption K. Then, for each linear program, the dimension of q is $|\mathcal{Q}| + 1 = |\mathcal{S}| = |\mathcal{S}_1| \times |\mathcal{S}_2| = 2^2 \times 2^2 \times 2^8 \times 2^4 = 65,536$. Note that the dimension is reduced with additional identifying assumptions. The number of main constraints is $\dim(p) = 2^{3\times 2} - 2^2 = 60$. There are $1 + 65,536$ additional constraints that

[2]There are more efficient algorithms than the bubble sort, such as the *quick sort*, although they must be modified to incorporate the distinct feature of our problem: the possible incomparability that stems from partial identification. Note that for comparable pairs, transitivity can be applied and thus the total number of comparisons can be smaller.

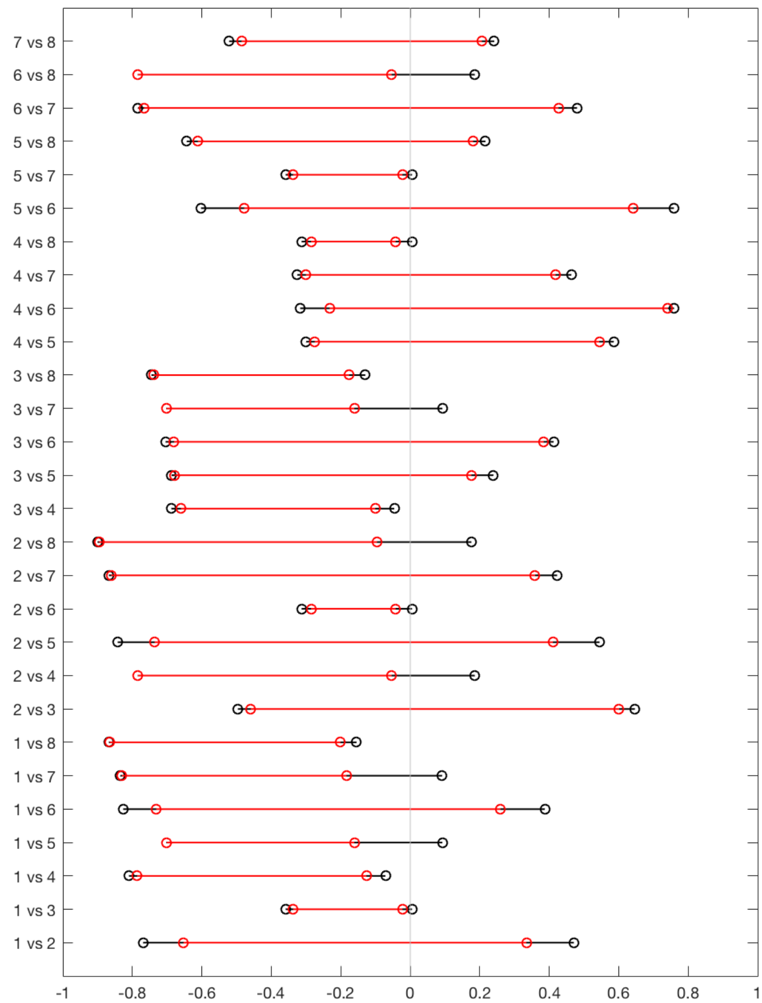

Figure 1: Sharp Bounds on Welfare Gaps under M1 (black) and M2 (red)

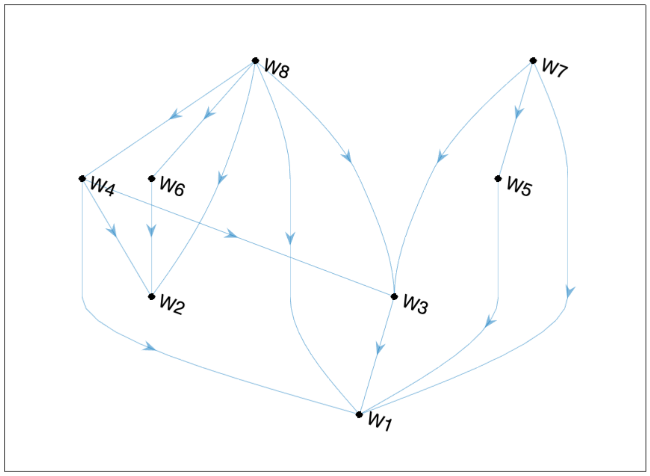

Figure 2: Sharp Directed Acyclic Graph under M2

define the simplex, i.e., $\sum_s q_s = 1$ and $q_s \geq 0$ for all $s \in \mathcal{S}$. Each linear program takes less than a second to calculate $L_{k,k'}$ or $U_{k,k'}$ with a computer with a 2.2 GHz single-core processor and 16 GB memory and with a modern solver such as CPLEX, MOSEK, and GUROBI.

Figure 1 reports the bounds $(L_{k,k'}, U_{k,k'})$ on $W_k - W_{k'}$ for all $(k, k') \in \{1, ..., 8\}$ under Assumption M1 (in black) and Assumption M2 (in red). In the figure, we can determine the sign of the welfare gap for those bounds that exclude zero. The difference between the black and red bounds illustrates the role of Assumption M2 relative to M1. That is, there are more bounds that avoid the zero vertical line with M2, which is consistent with the theory. It is important to note that, because M2 does not assume the direction of monotonicity, the sign of the welfare gap is not imposed by the assumption but recovered from the data.[3] Each set of bounds generates an associated DAGs (produced as an 8×8 adjacency matrix). Given the solutions of the linear programs, the adjacency matrix and thus the graph is simple to produce automatically using a standard software such as MATLAB. We proceed with Assumption M2 for brevity.

Figure 2 depicts the sharp DAG generated from $(L_{k,k'}, U_{k,k'})$'s under Assumption M2,

[3] The direction of the monotonicity in M2 can be estimated directly from the data by using the fact that $\text{sign}(E[Y_t|Z_t = 1, \boldsymbol{Y}^{t-1}, \boldsymbol{D}^{t-1}] - E[Y_t|Z_t = 1, \boldsymbol{Y}^{t-1}, \boldsymbol{D}^{t-1}]) = \text{sign}(E[Y_t(D^{t-1}, 1)|\boldsymbol{Y}^{t-1}, \boldsymbol{D}^{t-1}] - E[Y_t(D^{t-1}, 0)|, \boldsymbol{Y}^{t-1}, \boldsymbol{D}^{t-1}])$ almost surely. This result is an extension of Shaikh and Vytlacil (2011) to our multi-period setting.

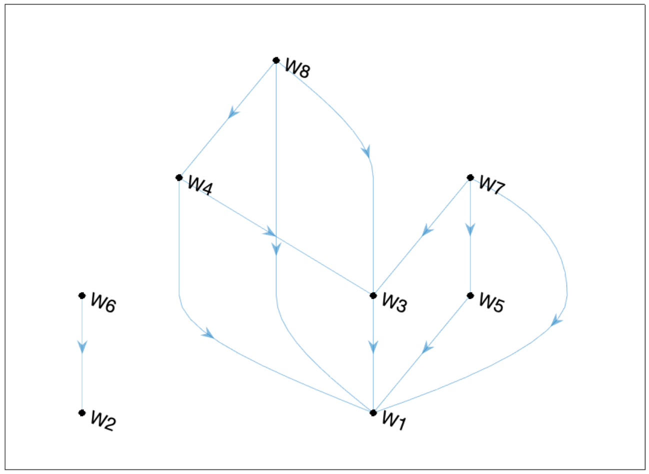

Figure 3: Sharp Directed Acyclic Graph under M2 (with only Z_1)

based on Theorem 3.1(i). Then, by Theorem 3.1(ii), the identified set of $\boldsymbol{\delta}^*(\cdot)$ is

$$\mathcal{D}_p^* = \{\boldsymbol{\delta}_7(\cdot), \boldsymbol{\delta}_8(\cdot)\}.$$

The common feature of the elements in $\mathcal{D}_p^*$ is that it is optimal to allocate $\delta_2 = 1$ for all $y_1 \in \{0, 1\}$. Finally, the following is one of the topological sorts produced from the DAG:

$$(\boldsymbol{\delta}_8(\cdot), \boldsymbol{\delta}_4(\cdot), \boldsymbol{\delta}_7(\cdot), \boldsymbol{\delta}_3(\cdot), \boldsymbol{\delta}_5(\cdot), \boldsymbol{\delta}_1(\cdot), \boldsymbol{\delta}_6(\cdot), \boldsymbol{\delta}_2(\cdot)).$$

We also conducted a parallel analysis but with a slightly different data-generating process, where (a) all the coefficients in (E.1)–(E.4) are positive except $\mu_{22} < 0$ and (b) Z_2 does not exist. In Case (a), we obtain $\mathcal{D}_p^* = \{\boldsymbol{\delta}_2(\cdot)\}$ as a singleton, i.e., we point identify $\boldsymbol{\delta}^*(\cdot) = \boldsymbol{\delta}_2(\cdot)$. The DAG for Case (b) is shown in Figure 3. We still obtain an informative DAG even with a single instrument. In this case, we obtain $\mathcal{D}_p^* = \{\boldsymbol{\delta}_6(\cdot), \boldsymbol{\delta}_7(\cdot), \boldsymbol{\delta}_8(\cdot)\}$.

Finally, we present further simulation results to investigate how the strength of instruments affect the partial ordering. We maintain the same simulation design and data-generating process as above. The original case of Figures 1 and 2 uses $(1, 0.8)$ for the values of the coefficients (π_1, π_{23}) on (Z_1, Z_2). Figure 4 shows the bounds and the DAG when $(\pi_1, \pi_{23}) = (0.5, 0.4)$, that is, the instruments (Z_1, Z_2) have 50% of strength compared to the

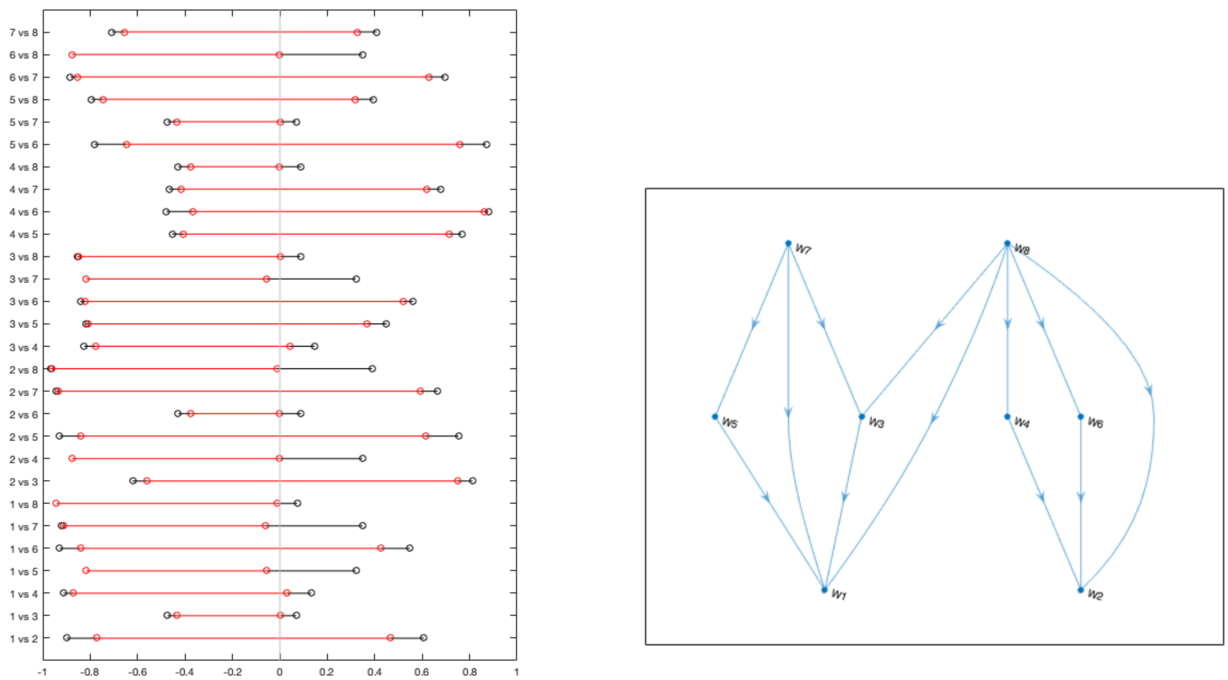

Figure 4: Left: Sharp Bounds on Welfare Gaps under M1 (black) and M2 (red); Right: Sharp DAG under M2 (IV strength: 50% of Figures 1 and 2)

original case. Figure 5 shows the results when $(\pi_1, \pi_{23}) = (0.25, 0.2)$, that is, the instruments (Z_1, Z_2) have only 25% of strength compared to the original case. In both figures, we obtain informative DAGs under M2. However, note that when we do not assume M2, the weaker instruments produce completely uninformative partial orderings as suggested from the bounds on the welfare gaps depicted in black. This exercise suggests the usefulness of M2 when instruments are weak. Finally, Figure 6 presents the results when $(\pi_1, \pi_{23}) = (1.5, 1.2)$, that is, the instruments (Z_1, Z_2) have 150% of strength compared to the original case. Although the DAG under M2 is identical to that in the original case, the informative bounds under M1 implies that the DAG under M1 will be very informative.

F Discussions

In Sections F.1–F.4, we propose some ways to report results of this paper including the partial ordering. These approaches can be useful especially when the obtained partial ordering is

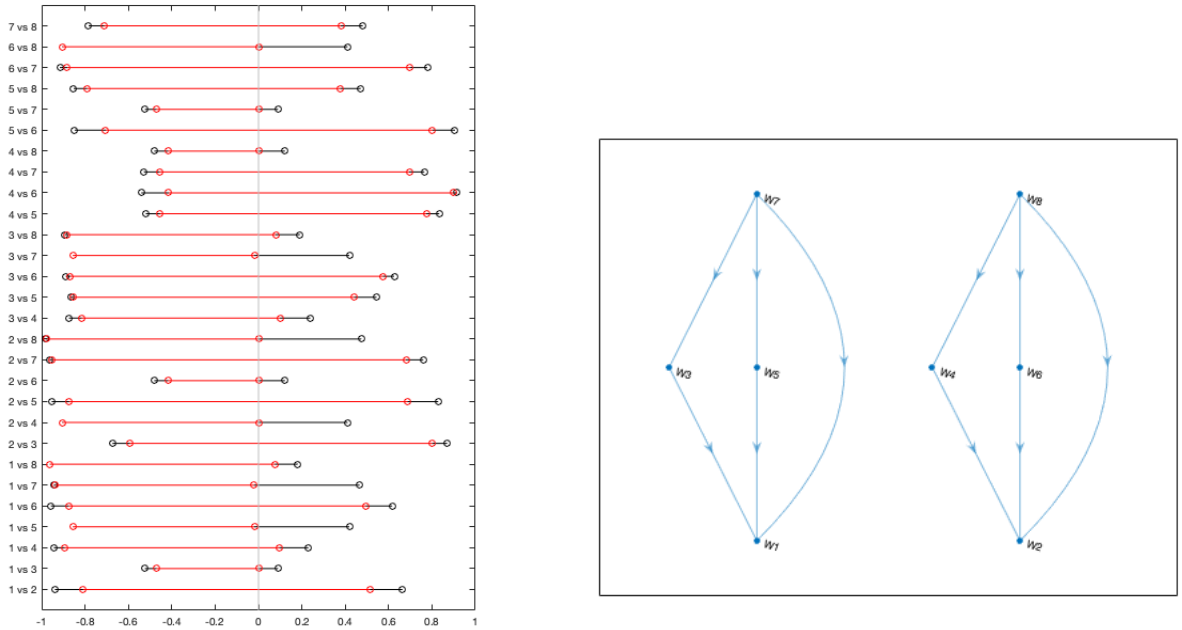

Figure 5: Left: Sharp Bounds on Welfare Gaps under M1 (black) and M2 (red); Right: Sharp DAG under M2 (IV strength: 25% of Figures 1 and 2)

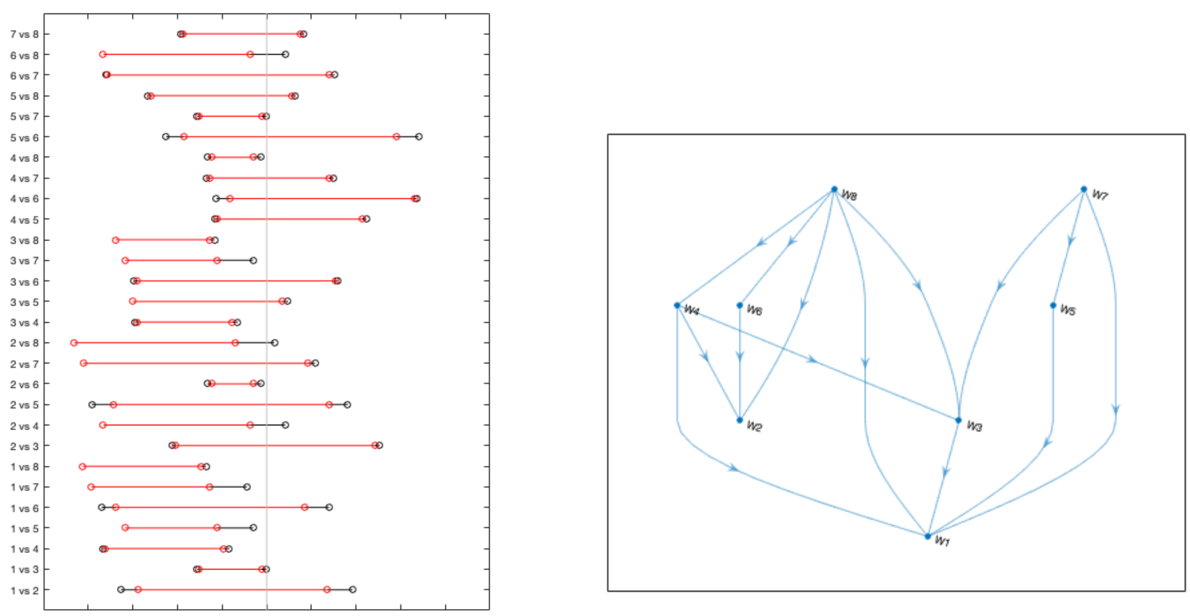

Figure 6: Left: Sharp Bounds on Welfare Gaps under M1 (black) and M2 (red); Right: Sharp DAG under M2 (IV strength: 150% of Figures 1 and 2)

complicated (e.g., with a longer horizon). We also discuss the cases where the set of possible regimes can be reduced. Section F.5 briefly discusses inference.

F.1 Set of the n-th Best Policies

When the partial ordering of welfare is the parameter of interest, the identified set of $\boldsymbol{\delta}^*(\cdot)$ can be viewed as a summary of the partial ordering. This view can be extended to introduce a set of the n-th best regimes, which further summarizes the partial ordering. With slight abuse of notation, we can formalize it as follows.

Recall $\mathcal{K}$ is the set of all regime indices. Motivated from (3.7), let $\mathcal{K}_p^{(1)} \equiv \{k' : \nexists k \in \mathcal{K} \text{ such that } L_{k,k'} > 0 \text{ and } k \neq k' \in \mathcal{K}\}$ be the set of maximal elements of the partial ordering and let $\mathcal{D}_p^{(1)} \equiv \{\boldsymbol{\delta}_{k'}(\cdot) : k' \in \mathcal{K}_p^{(1)}\}$. Theorem 3.1(ii) can be simply stated as $\mathcal{D}_p^* = \mathcal{D}_p^{(1)}$. To define the set of second-best regimes, we first remove all the elements in $\mathcal{K}_p^{(1)}$ from the set of candidate. Accordingly, by defining

$$\mathcal{K}_p^{(2)} \equiv \{k' : \nexists k \in \mathcal{K} \backslash \mathcal{K}_p^{(1)} \text{ such that } L_{k,k'} > 0 \text{ and } k \neq k' \in \mathcal{K} \backslash \mathcal{K}_p^{(1)}\},$$

we can introduce the set of second-best regimes: $\mathcal{D}_p^{(2)} \equiv \{\boldsymbol{\delta}_{k'}(\cdot) : k' \in \mathcal{K}_p^{(2)}\}$. Iteratively, we can define the set of n-th best regimes as $\mathcal{D}_p^{(n)} \equiv \{\boldsymbol{\delta}_{k'}(\cdot) : k' \in \mathcal{K}_p^{(n)}\}$ where

$$\mathcal{K}_p^{(n)} = \left\{ k' : \nexists k \in \mathcal{K} \backslash \bigcup_{j=1}^{n-1} \mathcal{K}_p^{(j)} \text{ such that } L_{k,k'} > 0 \text{ and } k \neq k' \in \mathcal{K} \backslash \bigcup_{j=1}^{n-1} \mathcal{K}_p^{(j)} \right\}.$$

The sets $\mathcal{D}_p^{(1)}, ..., \mathcal{D}_p^{(n)}$ can be recovered from the linear programs (3.6) and are useful policy benchmarks. For instance, the policy maker can conduct a sensitivity analysis for her chosen regime (e.g., from a parametric model) by inspecting in which set the regime is contained.

F.2 Topological Sorts as Observational Equivalence

Another way to summarize the partial ordering is to use topological sorts. A *topological sort* of a DAG is a linear ordering of its vertices that does not violate the order in the partial ordering given by the DAG. That is, for every directed edge $k \to k'$, k comes before k' in this linear ordering. Apparently, there can be multiple topological sorts for a DAG. Let L_G be the number of topological sorts of DAG $G(\mathcal{K}, \mathcal{E}_p)$, and let $k_{l,1} \in \mathcal{K}$ be the initial vertex of the l-th topological sort for $1 \leq l \leq L_G$. For example, given the DAG in Figure 1(a) (of the main text), $(\boldsymbol{\delta}_1, \boldsymbol{\delta}_4, \boldsymbol{\delta}_2, \boldsymbol{\delta}_3)$ is an example of a topological sort (with $k_{l,1} = 1$), but $(\boldsymbol{\delta}_1, \boldsymbol{\delta}_2, \boldsymbol{\delta}_4, \boldsymbol{\delta}_3)$ is not. Topological sorts are routinely reported for a given DAG, and there are well-known algorithms that efficiently find topological sorts, such as Kahn (1962)'s algorithm.

In fact, topological sorts can be viewed as total orderings that are *observationally equivalent* to the true *total* ordering of welfares. That is, each q generates the total ordering of welfares via $W_k = A_k q$, and q's in $\{q : Bq = p\} \cap \mathcal{Q}$ generates observationally equivalent total orderings. This insight enables us to interpret the partial ordering we establish using the more conventional notion of partial identification: the ordering is partially identified in the sense that the set of all topological sorts is not a singleton. This insight yields an alternative way of characterizing the identified set $\mathcal{D}_p^*$ of the optimal regime.

Theorem F.1. *Suppose Assumptions SX and B hold. The identified set $\mathcal{D}_p^*$ defined in (3.5) satisfies*

$$\mathcal{D}_p^* = \{\boldsymbol{\delta}_{k_{l,1}}(\cdot) : 1 \leq l \leq L_G\},$$

where $k_{l,1}$ is the initial vertex of the l-th topological sort of $G(\mathcal{K}, \mathcal{E}_p)$.

Suppose the DAG we recover from the data is not too sparse. By definition, a topological sort provides a ranking of regimes that is *not inconsistent* with the partial welfare ordering. Therefore, not only $\boldsymbol{\delta}_{k_{l,1}}(\cdot) \in \mathcal{D}_p^*$ but also the full sequence of a topological sort

$$\left(\boldsymbol{\delta}_{k_{l,1}}(\cdot), \boldsymbol{\delta}_{k_{l,2}}(\cdot), ..., \boldsymbol{d}_{k_{l,|\mathcal{D}|}}(\cdot)\right) \tag{F.1}$$

can be useful. A policymaker can be equipped with any of such sequences as a policy benchmark.

F.3 Bounds on Sorted Welfares

The set of n-th best regimes and topological sorts provide ordinal information about counterfactual welfares. To gain more comprehensive knowledge about the welfares, they can be accompanied by cardinal information: bounds on the sorted welfares. One might especially be interested in the bounds on "top-tier" welfares that are associated with the identified set or the first few elements in the topological sort. Bounds on gains from adaptivity and regrets can also be computed. These bounds can be calculated by solving linear programs. For instance, the sharp lower and upper bounds on welfare W_k can be calculated via

$$\begin{aligned} U_k &= \max_{q\in\mathcal{Q}} A_k q, \\ L_k &= \min_{q\in\mathcal{Q}} A_k q, \end{aligned} \qquad s.t. \quad Bq = p. \tag{F.2}$$

F.4 Cardinality Reduction

The typical time horizons we consider in this paper are short. For example, a multi-stage experiment called the Fast Track Prevention Program (Conduct Problems Prevention Research Group (1992)) considers $T = 4$. When T is not small, the cardinality of $\mathcal{D}$ may be too large, and we may want to reduce it for computational, institutional, and practical purposes.

One way to reduce the cardinality is to reduce the dimension of the adaptivity. Define a simpler adaptive treatment rule $\delta_t : \{0,1\} \times \{0,1\} \to \{0,1\}$ that maps only the lagged outcome and treatment onto a treatment allocation $d_t \in \{0,1\}$:

$$\delta_t(y_{t-1}, d_{t-1}) = d_t$$

for $t = 2, ..., T$ and $\delta_1(x) = d_1 \in \{0,1\}$. In this case, we have $|\mathcal{D}| = 2^{2(T-1)} \times 2^{|\mathcal{X}|}$ instead of $2^{2^T-2} \times 2^{|\mathcal{X}|}$. An even simpler rule, $\delta_t(y_{t-1})$, appears in Murphy et al. (2001).

Another possibility is to be motivated by institutional or budget constraints. For example, it may be the case that adaptive allocation is available every second period or only later in the horizon due to cost considerations. For example, suppose that the policymaker decides to introduce the adaptive rule at $t = T$ while maintaining static rules for $t \leq T - 1$. Finally, $\mathcal{D}$ can be restricted by budget or policy constraints that, e.g., the treatment is allocated to each individual at most once.

F.5 Inference

Although we do not fully investigate inference in the current paper, we briefly discuss it. For simplicity, we focus on the setting where $p(x)$ is known and thus $\Delta_{k,k'}$ is known. To conduct inference on the optimal regime $\boldsymbol{\delta}^*(\cdot)$, we can construct a confidence set (CS) for $\mathcal{D}_p^*$ with the following procedure. We consider a sequence of hypothesis tests, in which we eliminate regimes that are (statistically) significantly inferior to others. This is a statistical analog of the elimination procedure encoded in (3.7) or (3.8). For each test given $\tilde{\mathcal{K}} \subset \mathcal{K}$, we construct a null hypothesis that W_k and $W_{k'}$ are not comparable for all $k, k' \in \tilde{\mathcal{K}}$. Given (3.6), the incomparability of W_k and $W_{k'}$ is equivalent to $L_{k,k'} \leq 0 \leq U_{k,k'}$. In constructing this null hypothesis, it is helpful to invoke strong duality for the primal programs (3.6) and write the following dual programs:

$$U_{k,k'} = \min_{\lambda} \tilde{p}'\lambda, \qquad s.t. \quad \tilde{B}'\lambda \geq \Delta'_{k,k'} \tag{F.3}$$

$$L_{k,k'} = \max_{\lambda} -\tilde{p}'\lambda, \qquad s.t. \quad \tilde{B}'\lambda \geq -\Delta'_{k,k'} \tag{F.4}$$

where $\tilde{B} \equiv \begin{bmatrix} B \\ \mathbf{1}' \end{bmatrix}$ is a $(d_p+1) \times d_q$ matrix with $\mathbf{1}$ being a $d_q \times 1$ vector of ones and $\tilde{p} \equiv \begin{bmatrix} p \\ 1 \end{bmatrix}$ is a $(d_p + 1) \times 1$ vector. Let $\Lambda^U_{k,k'} \equiv \{\lambda : \tilde{B}'\lambda \geq \Delta'_{k,k'}\}$ and $\Lambda^L_{k,k'} \equiv \{\lambda : \tilde{B}'\lambda \geq -\Delta'_{k,k'}\}$. Then, we have $U_{k,k'} = \min_{\lambda \in \Lambda^U_{k,k'}} \tilde{p}'\lambda$ and $L_{k,k'} = \max_{\lambda \in \Lambda^L_{k,k'}} -\tilde{p}'\lambda$. Therefore, the null hypothesis that $L_{k,k'} \leq 0 \leq U_{k,k'}$ for $k, k' \in \tilde{\mathcal{K}}$ can be written as

$$H_{0,\tilde{\mathcal{K}}} : \tilde{p}'\lambda \geq 0 \text{ for all } \lambda \in \Lambda_{\tilde{\mathcal{K}}}. \tag{F.5}$$

where $\Lambda_{\tilde{\mathcal{K}}} \equiv \bigcup_{k,k' \in \tilde{\mathcal{K}}} \Lambda_{k,k'}$ with $\Lambda_{k,k'} \equiv \Lambda^U_{k,k'} \cup \Lambda^L_{k,k'}$.

Then, the procedure of constructing the CS, denoted as $\widehat{\mathcal{D}}_{CS}$, is as follows: *Step 0. Initially set* $\tilde{\mathcal{K}} = \mathcal{K}$. *Step 1. Test* $H_{0,\tilde{\mathcal{K}}}$ *at level* α *with test function* $\phi_{\tilde{\mathcal{K}}} \in \{0, 1\}$. *Step 2. If* $H_{0,\tilde{\mathcal{K}}}$ *is not rejected, define* $\widehat{\mathcal{D}}_{CS} = \{\boldsymbol{\delta}_k(\cdot) : k \in \tilde{\mathcal{K}}\}$*; otherwise eliminate vertex* $k_{\tilde{\mathcal{K}}}$ *from* $\tilde{\mathcal{K}}$ *and repeat from Step 1.* In Step 1, $T_{\tilde{\mathcal{K}}} \equiv \min_{k,k' \in \tilde{\mathcal{K}}} t_{k,k'}$ can be used as the test statistic for $H_{0,\tilde{\mathcal{K}}}$ where $t_{k,k'} \equiv \min_{\lambda \in \Lambda_{k,k'}} t_\lambda$ and t_λ is a standard t-statistic. The distribution of $T_{\tilde{\mathcal{K}}}$ can be estimated using bootstrap. In Step 2, a candidate for $k_{\tilde{\mathcal{K}}}$ is $k_{\tilde{\mathcal{K}}} \equiv \arg\min_{k \in \tilde{\mathcal{K}}} \min_{k' \in \tilde{\mathcal{K}}} t_{k,k'}$.

The eliminated vertices (i.e., regimes) are statistically suboptimal regimes, which are already policy-relevant outputs of the procedure. Note that the null hypothesis (F.5) consists of multiple inequalities. This incurs the issue of uniformity in that the null distribution depends on binding inequalities, whose identities are unknown. Such a problem has been studied in the literature, as in Hansen (2005), Andrews and Soares (2010), and Chen and Szroeter (2014). Hansen et al. (2011)'s bootstrap approach for constructing the model confidence set builds on Hansen (2005). We apply a similar inference method as in Hansen et al. (2011), but in this novel context and by being conscious about the computational challenge of our problem. In particular, the dual problem (F.3)–(F.4) and the vertex enumeration algorithm are introduced to ease the computational burden in simulating the distribution of $T_{\tilde{\mathcal{K}}}$. That is, the calculation of $\Lambda_{\tilde{\mathcal{K}}}$, the computationally intensive step, occurs only once, and then for each bootstrap sample, it suffices to calculate $\hat{p}$ instead of solving the linear programs (3.6) for all $k, k' \in \tilde{\mathcal{K}}$.

Analogous to Hansen et al. (2011), we can show that the resulting CS has desirable properties. Let $H_{A,\tilde{\mathcal{K}}}$ be the alternative hypothesis.

Assumption CS. *For any* $\tilde{\mathcal{K}}$, *(i)* $\limsup_{n\to\infty} \Pr[\phi_{\tilde{\mathcal{K}}} = 1 | H_{0,\tilde{\mathcal{K}}}] \leq \alpha$, *(ii)* $\lim_{n\to\infty} \Pr[\phi_{\tilde{\mathcal{K}}} = 1 | H_{A,\tilde{\mathcal{K}}}] = 1$, *and (iii)* $\lim_{n\to\infty} \Pr[\boldsymbol{\delta}_{k_{\tilde{\mathcal{K}}}}(\cdot) \in \mathcal{D}_p^* | H_{A,\tilde{\mathcal{K}}}] = 0$.

Proposition F.1. *Under Assumption CS, it satisfies that* $\liminf_{n\to\infty} \Pr[\mathcal{D}_p^* \subset \widehat{\mathcal{D}}_{CS}] \geq 1 - \alpha$ *and* $\lim_{n\to\infty} \Pr[\boldsymbol{\delta}(\cdot) \in \widehat{\mathcal{D}}_{CS}] = 0$ *for all* $\boldsymbol{\delta}(\cdot) \notin \mathcal{D}_p^*$.

The procedure of constructing the CS does not suffer from the problem of multiple testings. This is because the procedure stops as soon as the first hypothesis is not rejected, and asymptotically, maximal elements will not be questioned before all sub-optimal regimes are eliminated. The resulting CS can also be used to conduct a specification test for a less palatable assumption, such as Assumption M2. We can refute the assumption when the CS under that assumption is empty.

To implement the procedure in practice, we need to compute $\Lambda^U_{k,k'}$ and $\Lambda^L_{k,k'}$ for all $k, k' \in \mathcal{K}$. Note that $U_{k,k'} = \min_{\lambda \in \Lambda^U_{k,k'}} \tilde{p}'\lambda = \min_{\lambda \in \tilde{\Lambda}^U_{k,k'}} \tilde{p}'\lambda$ and $L_{k,k'} = \min_{\lambda \in \Lambda^L_{k,k'}} \tilde{p}'\lambda = \min_{\lambda \in \tilde{\Lambda}^L_{k,k'}} \tilde{p}'\lambda$ where $\tilde{\Lambda}^U_{k,k'}$ and $\tilde{\Lambda}^L_{k,k'}$ are sets of vertices in $\Lambda^U_{k,k'}$ and $\Lambda^L_{k,k'}$, respectively. Therefore, implementing the procedure reduces down to enumerating vertices of the polyhedra $\Lambda^U_{k,k'}$ and $\Lambda^L_{k,k'}$ or relevant subsets of them. This can be done by using a version of vertex enumeration algorithm (e.g., Avis and Fukuda (1992)). However, we note that the enumeration may be computationally extremely challenging especially when the dimension of q is large (which happens when we do not impose any additional identifying assumptions). There may be strategies that avoid the full enumeration, but this question is beyond the scope of the paper.

Inference on the welfare bounds in (F.2) may be conducted by using recent results as in Deb et al. (2017), who develop uniformly valid inference for bounds obtained via linear programming. Inference on optimized welfare $W_{\boldsymbol{\delta}^*}$ or $\max_{\boldsymbol{\delta}(\cdot) \in \widehat{\mathcal{D}}_{CS}} W_{\boldsymbol{\delta}}$ can also be an interesting problem. Andrews et al. (2019) consider inference on optimized welfare (evaluated at the estimated policy) in the context of Kitagawa and Tetenov (2018), but with point-identified welfare under the unconfoundedness assumption. Extending the framework to the current setting with partially identified welfare and dynamic regimes under treatment endogeneity would also be interesting future work; e.g., see Han and McCloskey (2022).

References

Andrews, D. W. and G. Soares (2010): "Inference for parameters defined by moment inequalities using generalized moment selection," *Econometrica*, 78, 119–157.

Andrews, I., T. Kitagawa, and A. McCloskey (2019): "Inference on winners," Tech. rep., National Bureau of Economic Research.

Avis, D. and K. Fukuda (1992): "A pivoting algorithm for convex hulls and vertex enumeration of arrangements and polyhedra," *Discrete & Computational Geometry*, 8, 295–313.

Chen, L.-Y. and J. Szroeter (2014): "Testing multiple inequality hypotheses: A smoothed indicator approach," *Journal of Econometrics*, 178, 678–693.

Conduct Problems Prevention Research Group (1992): "A developmental and clinical model for the prevention of conduct disorder: The FAST Track Program," *Development and Psychopathology*, 4, 509–527.

Deb, R., Y. Kitamura, J. K.-H. Quah, and J. Stoye (2017): "Revealed price preference: Theory and stochastic testing," *Cowles Foundation Discussion Paper*.

Han, S. and A. McCloskey (2022): "Inference on welfare for selected treatment rules that are set-identified," *U of Bristol and U of Colorado Boulder*.

Han, S. and S. Yang (2022): "Sharp Bounds on Treatment Effects for Policy Evaluation," *arXiv preprint arXiv:2009.13861*.

Hansen, P. R. (2005): "A test for superior predictive ability," *Journal of Business & Economic Statistics*, 23, 365–380.

Hansen, P. R., A. Lunde, and J. M. Nason (2011): "The model confidence set," *Econometrica*, 79, 453–497.

Kahn, A. B. (1962): "Topological sorting of large networks," *Communications of the ACM*, 5, 558–562.

Kitagawa, T. and A. Tetenov (2018): "Who should be treated? Empirical welfare maximization methods for treatment choice," *Econometrica*, 86, 591–616.

Manski, C. F. (2004): "Statistical treatment rules for heterogeneous populations," *Econometrica*, 72, 1221–1246.

Murphy, S. A. (2003): "Optimal dynamic treatment regimes," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 65, 331–355.

Murphy, S. A., M. J. van der Laan, J. M. Robins, and C. P. P. R. Group (2001): "Marginal mean models for dynamic regimes," *Journal of the American Statistical Association*, 96, 1410–1423.

Shaikh, A. M. and E. J. Vytlacil (2011): "Partial identification in triangular systems of equations with binary dependent variables," *Econometrica*, 79, 949–955.

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