Optimal Dynamic Treatment Regimes and Partial Welfare Ordering

Sukjin Han

University of Bristol

7 April 2022

Harvard & MIT Seminar

Dynamic (i.e., Adaptive) Treatment Regimes

Dynamic treatment regimes are seq's of treatment allocations...

- ...tailored to individual heterogeneity
- each period t, assignment rule $\delta_t(\cdot)$ maps previous outcome (and covariates) onto a current allocation decision

$$\delta_t(y_{t-1}) \in \{0, 1\}$$

Dynamic (i.e., Adaptive) Treatment Regimes

Dynamic treatment regimes are seq's of treatment allocations...

- ...tailored to individual heterogeneity
- each period t, assignment rule $\delta_t(\cdot)$ maps previous outcome (and covariates) onto a current allocation decision

$$\delta_t(y_{t-1}) \in \{0, 1\}$$

Optimal dynamic treatment regime is a dynamic regime that maximizes counterfactual welfare

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

Identification of Optimal Dynamic Treatment Regime

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

This paper investigates the possibility of identification of $\boldsymbol{\delta}^*(\cdot)$ when data are from...

- multi-stage experiments with possible non-compliance,

  or

- more generally, observational studies

Motivating Example: Returns to Schooling & Training

Y_2 employed after program

D_2 receiving job training program

Y_1 employed before program

D_1 receiving high school diploma

Motivating Example: Returns to Schooling & Training

Y_2 employed after program

D_2 receiving job training program

Y_1 employed before program

D_1 receiving high school diploma

Let $Y_1(d_1)$ and $Y_2(d_2)$ be counterfactual employment status

Treatment effects: $E[Y_1(1)] - E[Y_1(0)]$ and $E[Y_2(1)] - E[Y_2(0)]$

Motivating Example: Returns to Schooling & Training

Y_2 employed after program

D_2 receiving job training program

Y_1 employed before program

D_1 receiving high school diploma

Let $Y_1(d_1)$ and $Y_2(d_2)$ be counterfactual employment status

Treatment effects: $E[Y_1(1)] - E[Y_1(0)]$ and $E[Y_2(1)] - E[Y_2(0)]$

May be interested in the effects of sequence of treatments using $Y_2(d_1, d_2)$

Then, e.g., $E[Y_2(1, 0)] - E[Y_2(0, 1)]$ or complementarity:

$$E[Y_2(1, 1)] - E[Y_2(1, 0)] \text{ vs. } E[Y_2(0, 1)] - E[Y_2(0, 0)]$$

Motivating Example: Returns to Schooling & Training

Y_2 employed after program

D_2 receiving job training program

Y_1 employed before program

D_1 receiving high school diploma

Since (d_1, d_2) are *not* simultaneously provided, Y_1 responds to d_1 (as $Y_1(d_1)$)

Motivating Example: Returns to Schooling & Training

Y_2 employed after program

D_2 receiving job training program

Y_1 employed before program

D_1 receiving high school diploma

Since (d_1, d_2) are *not* simultaneously provided, Y_1 responds to d_1 (as $Y_1(d_1)$)

So want to incorporate this knowledge in recommendation of d_2

thus, instead of d_2, consider $\delta_2(Y_1(d_1))$ as hypothetical policy

Motivating Example: Returns to Schooling & Training

Y_2 employed after program

D_2 receiving job training program

Y_1 employed before program

D_1 receiving high school diploma

Since (d_1, d_2) are *not* simultaneously provided, Y_1 responds to d_1 (as $Y_1(d_1)$)

So want to incorporate this knowledge in recommendation of d_2

thus, instead of d_2, consider $\delta_2(Y_1(d_1))$ as hypothetical policy

And, instead of $Y_2(d_1, d_2)$, consider

$$Y_2(d_1, \delta_2(Y_1(d_1))) \equiv Y_2(\boldsymbol{\delta})$$

Motivating Example: Returns to Schooling & Training

Y_2 employed after program

D_2 receiving job training program

Y_1 employed before program

D_1 receiving high school diploma

Optimal dynamic regime: schedule $\boldsymbol{\delta}(\cdot) = (\delta_1, \delta_2(\cdot))$ of allocation *rules* that maximizes $W_{\boldsymbol{\delta}} = E[Y_2(\boldsymbol{\delta})]$ where

$$\delta_1 = d_1, \qquad \delta_2(Y_1(\delta_1)) = d_2$$

Motivating Example: Returns to Schooling & Training

Y_2 employed after program

D_2 receiving job training program

Y_1 employed before program

D_1 receiving high school diploma

Optimal dynamic regime: schedule $\boldsymbol{\delta}(\cdot) = (\delta_1, \delta_2(\cdot))$ of allocation *rules* that maximizes $W_{\boldsymbol{\delta}} = E[Y_2(\boldsymbol{\delta})]$ where

$$\delta_1 = d_1, \qquad \delta_2(Y_1(\delta_1)) = d_2$$

Policy implication of $\boldsymbol{\delta}^*(\cdot)$ s.t. $\delta_2^*(1) = 0$, $\delta_2^*(0) = 1$,...

- more training resources to disadvantaged workers
- with δ_1^* combined, interaction with earlier schooling

Instrument Variables from Sequential Designs

How to learn $W_{\boldsymbol{\delta}}$'s and $\boldsymbol{\delta}^{*}(\cdot)$, esp. when treatments are endogenous?

Instrument Variables from Sequential Designs

How to learn W_δ's and $\delta^*(\cdot)$, esp. when treatments are endogenous?

$\Rightarrow$ we show IVs from sequential (quasi-) experiments are helpful

- e.g., medical trials, field experiments, A/B testings
- e.g., seq of policy shocks, sequential fuzzy RDs

Instrument Variables from Sequential Designs

How to learn W_δ's and $\delta^*(\cdot)$, esp. when treatments are endogenous?

$\Rightarrow$ we show IVs from sequential (quasi-) experiments are helpful

- e.g., medical trials, field experiments, A/B testings
- e.g., seq of policy shocks, sequential fuzzy RDs

In motivating example,

- distance to (or density of) high schools can be Z_1
- random assignment of job training can be Z_2

Instrument Variables from Sequential Designs

How to learn W_δ's and $\delta^*(\cdot)$, esp. when treatments are endogenous?

$\Rightarrow$ we show IVs from sequential (quasi-) experiments are helpful

- e.g., medical trials, field experiments, A/B testings
- e.g., seq of policy shocks, sequential fuzzy RDs

In motivating example,

- distance to (or density of) high schools can be Z_1
- random assignment of job training can be Z_2

Single IV can still be helpful esp. with short horizon

This Paper: Partial ID of Optimal Regime and Welfares

This paper proposes a nonparametric framework where we can (at least partially) learn optimal dynamic regime and related welfares

This Paper: Partial ID of Optimal Regime and Welfares

This paper proposes a nonparametric framework where we can (at least partially) learn optimal dynamic regime and related welfares

1. We establish mapping from data to sharp partial ordering (i.e., ranking) of $W_{\boldsymbol{\delta}}$'s w.r.t. $\boldsymbol{\delta}(\cdot) \in \mathcal{D}$

Sharp Partial Welfare Ordering in Numerical Exercise

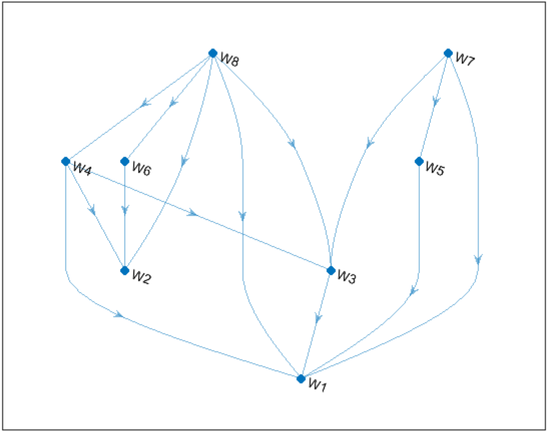


Figure: Partial Ordering as Directed Acyclic Graph

This Paper: Partial ID of Optimal Regime and Welfares

This paper proposes a nonparametric framework where we can (at least partially) learn optimal dynamic regime and related welfares

1. We establish mapping from data to sharp partial ordering (i.e., ranking) of $W_{\boldsymbol{\delta}}$'s w.r.t. $\boldsymbol{\delta}(\cdot) \in \mathcal{D}$

This Paper: Partial ID of Optimal Regime and Welfares

This paper proposes a nonparametric framework where we can (at least partially) learn optimal dynamic regime and related welfares

1. We establish mapping from data to sharp partial ordering (i.e., ranking) of $W_{\boldsymbol{\delta}}$'s w.r.t. $\boldsymbol{\delta}(\cdot) \in \mathcal{D}$

2. Based on partial ordering, we characterize (sharp) identified set for optimal regime $\boldsymbol{\delta}^*(\cdot)$

  - as a set of maximal elements

Sharp Partial Welfare Ordering in Numerical Exercise

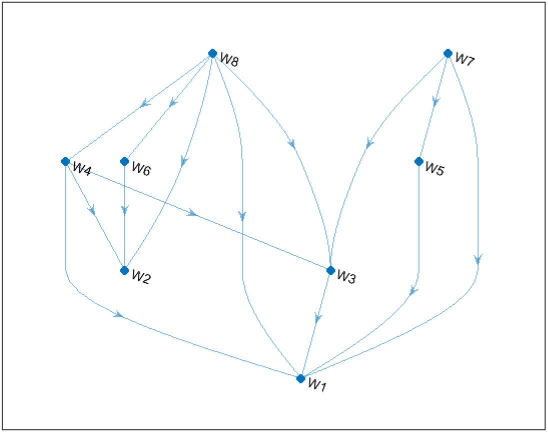


Figure: Partial Ordering as Directed Acyclic Graph

This Paper: Partial ID of Optimal Regime and Welfares

This paper proposes a nonparametric framework where we can (at least partially) learn optimal dynamic regime and related welfares

1. We establish mapping from data to sharp partial ordering (i.e., ranking) of W_{δ}'s w.r.t. $\delta(\cdot) \in \mathcal{D}$

2. Based on partial ordering, we characterize (sharp) identified set for optimal regime $\delta^*(\cdot)$

- as a set of maximal elements

This Paper: Partial ID of Optimal Regime and Welfares

This paper proposes a nonparametric framework where we can (at least partially) learn optimal dynamic regime and related welfares

1. We establish mapping from data to sharp partial ordering (i.e., ranking) of $W_{\boldsymbol{\delta}}$'s w.r.t. $\boldsymbol{\delta}(\cdot) \in \mathcal{D}$

2. Based on partial ordering, we characterize (sharp) identified set for optimal regime $\boldsymbol{\delta}^*(\cdot)$

   - as a set of maximal elements

3. We propose additional assumptions that tighten the ID'ed set

   - e.g., on agent's behavior or dynamics

This Paper: Partial ID of Optimal Regime and Welfares

This paper proposes a nonparametric framework where we can (at least partially) learn optimal dynamic regime and related welfares

1. We establish mapping from data to sharp partial ordering (i.e., ranking) of W_{δ}'s w.r.t. $\delta(\cdot) \in \mathcal{D}$

2. Based on partial ordering, we characterize (sharp) identified set for optimal regime $\delta^*(\cdot)$
   - as a set of maximal elements

3. We propose additional assumptions that tighten the ID'ed set
   - e.g., on agent's behavior or dynamics

4. We apply the method in policy analysis using schooling & post-school training as sequence of treatments

Contribution 1: Treatment Endogeneity

Dynamic treatment regimes:

- Murphy et al. 01, Murphy 03, Robins 04,...
- sequential randomization: "randomize treatment in the current period conditional on past treatments and outcomes"

Statistical treatment rules and policy learning:

- Manski 04, Hirano & Porter 09, Bhattacharya & Dupas 12, Stoye 12, Kitagawa & Tetenov 18, Sakaguchi 19, Athey & Wager 21, Mbakop & Tabord-Meehan 21,...
- versions of unconfoundedness assumption

Contribution 1: Treatment Endogeneity

Dynamic treatment regimes:

- Murphy et al. 01, Murphy 03, Robins 04,...
- sequential randomization: "randomize treatment in the current period conditional on past treatments and outcomes"

Statistical treatment rules and policy learning:

- Manski 04, Hirano & Porter 09, Bhattacharya & Dupas 12, Stoye 12, Kitagawa & Tetenov 18, Sakaguchi 19, Athey & Wager 21, Mbakop & Tabord-Meehan 21,...
- versions of unconfoundedness assumption

Not plausible in experiments with partial compliance and many observational studies

Contribution 1: Treatment Endogeneity

Dynamic treatment regimes:

- Murphy et al. 01, Murphy 03, Robins 04,...
- sequential randomization: "randomize treatment in the current period conditional on past treatments and outcomes"

Statistical treatment rules and policy learning:

- Manski 04, Hirano & Porter 09, Bhattacharya & Dupas 12, Stoye 12, Kitagawa & Tetenov 18, Sakaguchi 19, Athey & Wager 21, Mbakop & Tabord-Meehan 21,...
- versions of unconfoundedness assumption

Not plausible in experiments with partial compliance and many observational studies

This paper: relaxes sequential randomization

Contribution 2: Partial ID in Multi-Period Settings

ID of optimal regime (as fcn of covariates) using IVs:

- Cui & Tchetgen Tchetgen 20, Qiu et al. 20, Han 21; Kasy 16, Pu & Zhang 2021
  - single-period setting
  - rely on independence of compliance type or rank preservation
  - or partial ID
- Han 20
  - dynamic treatment effects and optimal regime in multi-period setting
  - rely on existence of extra exogenous variables

Contribution 2: Partial ID in Multi-Period Settings

ID of optimal regime (as fcn of covariates) using IVs:

- Cui & Tchetgen Tchetgen 20, Qiu et al. 20, Han 21; Kasy 16, Pu & Zhang 2021
  - single-period setting
  - rely on independence of compliance type or rank preservation
  - or partial ID
- Han 20
  - dynamic treatment effects and optimal regime in multi-period setting
  - rely on existence of extra exogenous variables

This paper:

- partial ID of optimal adaptive regime and dynamic treatment effects

Contribution 3: Linear Programming Approach to Partial ID

Calculating bounds using linear programming (LP)

- Balke & Pearl 97, Manski 07, Mogstad et al. 18, Kitamura & Stoye 19, Torgovitsky 19, Machado et al. 19, Kamat 19, Han & Yang 20,...

This paper:

- establish partial ordering via a set of LPs...
- that are governed by the same DGP...
- and characterize bounds on welfare gaps

Simple estimation and inference procedures for optimal regime

Contribution 3: Linear Programming Approach to Partial ID

Calculating bounds using linear programming (LP)

- Balke & Pearl 97, Manski 07, Mogstad et al. 18, Kitamura & Stoye 19, Torgovitsky 19, Machado et al. 19, Kamat 19, Han & Yang 20,...

This paper:

- establish partial ordering via a set of LPs...
- that are governed by the same DGP...
- and characterize bounds on welfare gaps

Simple estimation and inference procedures for optimal regime

Broader applicability:

- rankings across different counterfactual scenarios

Roadmap

I. Dynamic treatment regime and counterfactual welfare

II. Partial ID of optimal dynamic regime

- linear programming
- partial ordering and ID'ed set

III. Additional identifying assumptions

IV. Numerical illustration

V. Empirical application

VI. Inference

I. Dynamic Treatment Regime and Counterfactual Welfare

Dynamic (i.e., Adaptive) Treatment Regimes

Consider two-period case ($T = 2$) only for simplicity

Dynamic regime is defined as

$$\boldsymbol{\delta}(\cdot) \equiv (\delta_1, \delta_2(\cdot)) \in \mathcal{D}$$

where

$$\begin{aligned} \delta_1 &= d_1 \in \{0, 1\} \\ \delta_2(y_1) &= d_2 \in \{0, 1\} \end{aligned}$$

- e.g., y_t symptom, d_t medical treatment
- (stochastic rules in the paper)

Dynamic (i.e., Adaptive) Treatment Regimes

Regime #	δ_1	$\delta_2(1)$	$\delta_2(0)$
1	0	0	0
2	1	0	0
3	0	1	0
4	1	1	0
5	0	0	1
6	1	0	1
7	0	1	1
8	1	1	1

Table: Dynamic Regimes $\boldsymbol{\delta}(\cdot) \equiv (\delta_1, \delta_2(\cdot))$ when $T = 2$

Non-Adaptive Treatment Regimes

Regime #	d_1	d_2
1	0	0
2	1	0
3	0	1
4	1	1

Table: Non-Adaptive Regimes $\boldsymbol{d} \equiv (d_1, d_2)$ when $T = 2$

Counterfactual Outcomes

Define potential outcome as a function of dynamic regime

Counterfactual Outcomes

Define potential outcome as a function of dynamic regime

Potential outcomes with non-adaptive regime $\boldsymbol{d} = (d_1, d_2)$:

$$
\begin{aligned}
Y_1(d_1) \\
Y_2(d_1, d_2)
\end{aligned}
$$

Counterfactual Outcomes

Define potential outcome as a function of dynamic regime

Potential outcomes with non-adaptive regime $\boldsymbol{d} = (d_1, d_2)$:

$$Y_1(d_1)$$
$$Y_2(d_1, d_2)$$

Potential outcomes with dynamic regime $\boldsymbol{\delta}(\cdot) = (\delta_1, \delta_2(\cdot))$:

$$\begin{aligned} Y_1(\delta_1) &= Y_1(d_1) \\ Y_2(\boldsymbol{\delta}) &= Y_2(\delta_1, \delta_2(Y_1(\delta_1))) \end{aligned}$$

Welfare and Optimal Dynamic Regime

Let $\boldsymbol{Y}(\boldsymbol{\delta}) \equiv (Y_1(\delta_1), Y_2(\boldsymbol{\delta}))$

Counterfactual welfare as linear funct'l of $q_{\boldsymbol{\delta}}(\boldsymbol{y}) \equiv \Pr[\boldsymbol{Y}(\boldsymbol{\delta}(\cdot)) = \boldsymbol{y}]$

$$W_{\boldsymbol{\delta}} \equiv f(q_{\boldsymbol{\delta}})$$

- e.g., $E[Y_T(\boldsymbol{\delta}(\cdot))] = \Pr[Y_T(\boldsymbol{\delta}(\cdot)) = 1]$ Details
- e.g., $\sum_{t=1}^{T}\{\omega_t E[Y_t(\boldsymbol{\delta}^t(\cdot))]\}$ (less the cost of treatments)

Welfare and Optimal Dynamic Regime

Let $\boldsymbol{Y}(\boldsymbol{\delta}) \equiv (Y_1(\delta_1), Y_2(\boldsymbol{\delta}))$

Counterfactual welfare as linear funct'l of $q_{\delta}(\boldsymbol{y}) \equiv \Pr[\boldsymbol{Y}(\boldsymbol{\delta}(\cdot)) = \boldsymbol{y}]$

$$W_{\delta} \equiv f(q_{\delta})$$

- e.g., $E[Y_T(\boldsymbol{\delta}(\cdot))] = \Pr[Y_T(\boldsymbol{\delta}(\cdot)) = 1]$ ▸ Details
- e.g., $\sum_{t=1}^{T}\{\omega_t E[Y_t(\boldsymbol{\delta}^t(\cdot))]\}$ (less the cost of treatments)

Optimal dynamic regime as

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

II. Partial ID of Optimal Dynamic Regime

Observed Data

For $t = 1, ..., T$ on a finite horizon,

- $Y_t \in \{0, 1\}$ outcome at t (e.g., symptom indicator)
  - extension: continuous Y_t with discretized rule (later)
- $D_t \in \{0, 1\}$ treatment at t (e.g., medical treatment received)
- $Z_t \in \{0, 1\}$ instrument at t (e.g., medical treatment assigned)

Large N small T panel of $(\boldsymbol{Y}, \boldsymbol{D}, \boldsymbol{Z})$

- (cross-sectional index i suppressed; covariates suppressed)
- more generally, e.g., single IV is allowed

Partial ID of Optimal Dynamic Regime

Let $\boldsymbol{Y}(\boldsymbol{d})$ be vector of $Y_t(\boldsymbol{d}^t)$'s and $\boldsymbol{D}(\boldsymbol{z})$ be vector of $D_t(\boldsymbol{z}^t)$'s.

Assumption SX

$Z_t \perp (\boldsymbol{Y}(\boldsymbol{d}), \boldsymbol{D}(\boldsymbol{z}))|\boldsymbol{Z}^{t-1}$.

- e.g., sequential randomized experiments, sequential fuzzy RDs

Partial ID of Optimal Dynamic Regime

Let $\boldsymbol{Y}(\boldsymbol{d})$ be vector of $Y_t(\boldsymbol{d}^t)$'s and $\boldsymbol{D}(\boldsymbol{z})$ be vector of $D_t(\boldsymbol{z}^t)$'s.

Assumption SX

$Z_t \perp (\boldsymbol{Y}(\boldsymbol{d}), \boldsymbol{D}(\boldsymbol{z})) | \boldsymbol{Z}^{t-1}$.

- e.g., sequential randomized experiments, sequential fuzzy RDs

Goal: to characterize ID'ed set for $\delta^*(\cdot)$ given the distribution of $(\boldsymbol{Y}, \boldsymbol{D}, \boldsymbol{Z})$

Partial ID of Optimal Dynamic Regime

Let $\boldsymbol{Y}(\boldsymbol{d})$ be vector of $Y_t(\boldsymbol{d}^t)$'s and $\boldsymbol{D}(\boldsymbol{z})$ be vector of $D_t(\boldsymbol{z}^t)$'s.

Assumption SX

$Z_t \perp (\boldsymbol{Y}(\boldsymbol{d}), \boldsymbol{D}(\boldsymbol{z})) | \boldsymbol{Z}^{t-1}$.

- e.g., sequential randomized experiments, sequential fuzzy RDs

Goal: to characterize ID'ed set for $\boldsymbol{\delta}^*(\cdot)$ given the distribution of $(\boldsymbol{Y}, \boldsymbol{D}, \boldsymbol{Z})$

ID'ed set as a subset of the discrete set $\mathcal{D}$:

$$\mathcal{D}^* \subset \mathcal{D}$$

Partial ID of Optimal Dynamic Regime

As first step, establish *sharp partial ordering* of welfare W_δ w.r.t. $\delta(\cdot)$ based on $(\boldsymbol{Y}, \boldsymbol{D}, \boldsymbol{Z})$

- cf. total ordering is needed for point ID of $\delta^*(\cdot)$
- can only recover obs'ly equivalent total orderings

Partial ID of Optimal Dynamic Regime

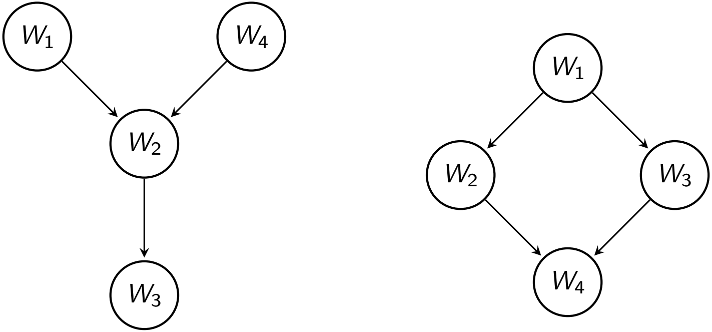
As first step, establish *sharp partial ordering* of welfare W_{δ} w.r.t. $\delta(\cdot)$ based on $(\boldsymbol{Y}, \boldsymbol{D}, \boldsymbol{Z})$

- cf. total ordering is needed for point ID of $\delta^{*}(\cdot)$
- can only recover obs'ly equivalent total orderings

Partial ordering = a *directed acyclic graph* (DAG)

- parameter of independent interest
- *topological sorts* of DAG = obs'ly equivalent total orderings

Partial Ordering of Welfare $W_k \equiv W_{\delta_k}$



(a) (b)

Figure: Partially Ordered Sets as DAGs

Sharp Partial Ordering of Welfare W_δ

We want this partial ordering to be *sharp*

Definition (Sharp Partial Ordering, i.e., Sharp DAG)

In the DAG, no more edges can be created without additional assumptions.

Sharp Partial Ordering of Welfare W_δ

We want this partial ordering to be *sharp*

Definition (Sharp Partial Ordering, i.e., Sharp DAG)

In the DAG, no more edges can be created without additional assumptions.

To guarantee this, characterize sharp lower and upper bounds on

$$W_\delta - W_{\delta'}$$

as optima of linear programming

Linear Programming for Bounds on Welfare Gap

For each $\delta, \delta' \in \mathcal{D}$, welfare gap (i.e., dynamic treatment effect) is

$$W_{\delta} - W_{\delta'} = (A_{\delta} - A_{\delta'})q$$

where $q \in \mathcal{Q}$ is vector of latent distribution

Linear Programming for Bounds on Welfare Gap

For each $\boldsymbol{\delta}, \boldsymbol{\delta}' \in \mathcal{D}$, welfare gap (i.e., dynamic treatment effect) is

$$W_{\boldsymbol{\delta}} - W_{\boldsymbol{\delta}'} = (A_{\boldsymbol{\delta}} - A_{\boldsymbol{\delta}'})q$$

where $q \in \mathcal{Q}$ is vector of latent distribution

Sharp lower and upper bounds via linear programming:

$$\begin{aligned} L_{\boldsymbol{\delta},\boldsymbol{\delta}'} &= \min_{q \in \mathcal{Q}} (A_{\boldsymbol{\delta}} - A_{\boldsymbol{\delta}'})q \\ U_{\boldsymbol{\delta},\boldsymbol{\delta}'} &= \max_{q \in \mathcal{Q}} (A_{\boldsymbol{\delta}} - A_{\boldsymbol{\delta}'})q \end{aligned} \qquad s.t. \quad Bq = p$$

- $A_{\boldsymbol{\delta}}$, $A_{\boldsymbol{\delta}'}$, and B are known to researcher
- p is vector of data distribution for $(\boldsymbol{Y}, \boldsymbol{D}, \boldsymbol{Z})$
- q is unknown decision variable in standard simplex $\mathcal{Q}$

Sharp Partial Ordering and Identified Set

Theorem

Suppose SX holds. (i) DAG is sharp with set of edges

$$\{(W_\delta, W_{\delta'}) : L_{\delta,\delta'} > 0 \text{ for } \delta \neq \delta'\}$$

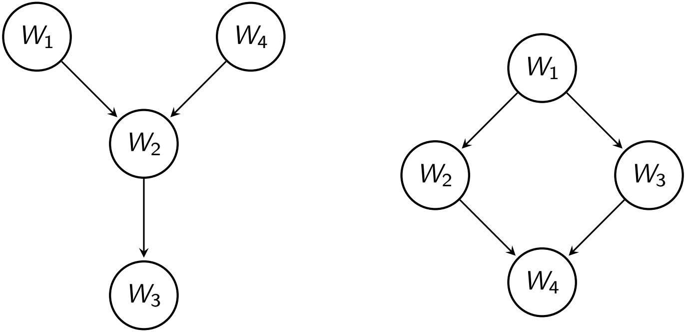
(ii) $\mathcal{D}_p^*$ satisfies

$$\begin{aligned}\mathcal{D}_p^* &= \{\delta' : \nexists \delta \text{ such that } L_{\delta,\delta'} > 0 \text{ for } \delta \neq \delta'\} \\ &= \{\delta' : L_{\delta,\delta'} \leq 0 \text{ for all } \delta \text{ and } \delta \neq \delta'\}\end{aligned}$$

i.e., the rhs set is sharp

- $\mathcal{D}_p^*$ is the set of *maximal elements* associated with the DAG
- key insight: despite separate optimizations, DAG is governed by *common* latent dist q's in $\{q : Bq = p\}$ (i.e., that are obs'ly equivalent)

Partial Ordering of Welfare $W_k \equiv W_{\delta_k}$



(a) $\boldsymbol{\delta}^*(\cdot)$ is partially ID'ed
$\mathcal{D}_p^* = \{\boldsymbol{\delta}_{\#1}, \boldsymbol{\delta}_{\#4}\}$

(b) $\boldsymbol{\delta}^*(\cdot)$ is point ID'ed
$\mathcal{D}_p^* = \{\boldsymbol{\delta}_{\#1}\}$

Figure: Partially Ordered Sets as DAGs

Discussion: Identified Set

Given the minimal structure, the size of $\mathcal{D}_p^*$ may be large

Discussion: Identified Set

Given the minimal structure, the size of $\mathcal{D}_p^*$ may be large

Such $\mathcal{D}_p^*$ still has implications for policy:

(i) it recommends the planner to eliminate sub-optimal regimes from her options

(ii) it warns about the lack of informativeness of data (e.g., even with experimental data)

Discussion: Identified Set

Given the minimal structure, the size of $\mathcal{D}_p^*$ may be large

Such $\mathcal{D}_p^*$ still has implications for policy:

(i) it recommends the planner to eliminate sub-optimal regimes from her options

(ii) it warns about the lack of informativeness of data (e.g., even with experimental data)

The size of $\mathcal{D}_p^*$ is related to...

- the strength of Z_t (i.e., the size of the complier group at t),
- the strength of the dynamic treatment effects

III. Additional Identifying Assumptions

Additional Identifying Assumptions

Researchers are willing to impose more assumptions based on priors about agent's behavior or dynamics

- monotonicity/uniformity ‣ Assumption M1 ‣ Assumption M2
  - Imbens & Angrist 94, Manski & Pepper 00
  - for each t, either $Y_t(1) \geq Y_t(0)$ w.p.1 or $Y_t(1) \leq Y_t(0)$ w.p.1. conditional on $(\boldsymbol{Y}^{t-1}, \boldsymbol{D}^{t-1})$
- agent's learning ‣ Assumption L
- Markovian structure ‣ Assumption K
- positive state dependence, stationarity, etc.
  - Torgovitsky 19

Easy to incorporate within the linear programming

These assumptions tighten the ID'ed set $\mathcal{D}_p^*$ by...

- reducing the dimension of the simplex $\mathcal{Q}$

IV. Numerical Illustration

Numerical Illustration

For $T = 2$, DGP is

$$
\begin{aligned}
D_{i1} &= 1\{\pi_1 Z_{i1} + \alpha_i + v_{i1} \geq 0\} \\
Y_{i1} &= 1\{\mu_1 D_{i1} + \alpha_i + e_{i1} \geq 0\} \\
D_{i2} &= 1\{\pi_{21} Y_{i1} + \pi_{22} D_{i1} + \pi_{23} Z_{i2} + \alpha_i + v_{i2} \geq 0\} \\
Y_{i2} &= 1\{\mu_{21} Y_{i1} + \mu_{22} D_{i2} + \alpha_i + e_{i2} \geq 0\}
\end{aligned}
$$

and

$$W_{\boldsymbol{\delta}} = E[Y_2(\boldsymbol{\delta})]$$

Calculate $[L_{\boldsymbol{\delta}_k,\boldsymbol{\delta}_l}, U_{\boldsymbol{\delta}_k,\boldsymbol{\delta}_l}]$ for $W_{\boldsymbol{\delta}_k} - W_{\boldsymbol{\delta}_l}$ for all pairs $k, l \in \{1, ..., 8\}$

We make $\binom{8}{2} = 28$ comparisons, i.e., 28×2 linear programs

Bounds on Welfare Gaps $W_{\delta_k} - W_{\delta_l}$

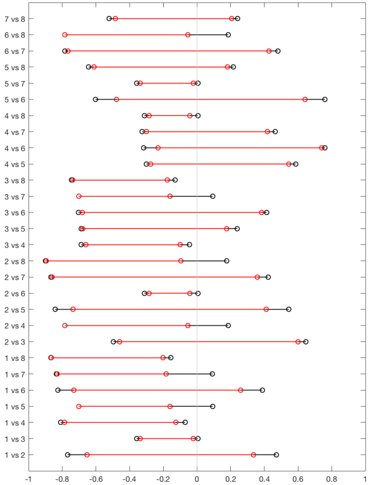


Figure: Sharp Bounds on Welfare Gaps (red: under M2)

Sharp Partial Welfare Ordering

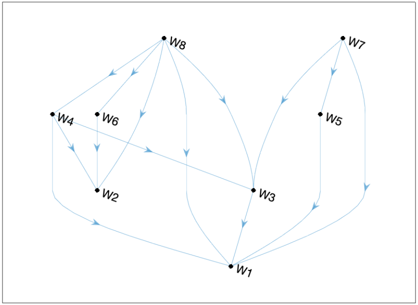


Figure: Partial Ordering as DAG and ID'ed Set for δ^* (under M2)

Sharp Partial Welfare Ordering

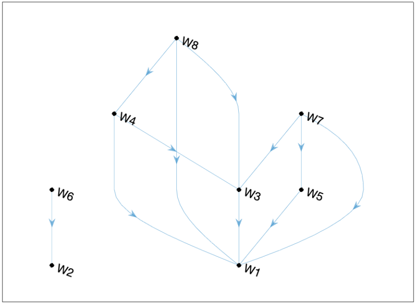


Figure: Partial Ordering as DAG with Only Z_1 (under M2)

V. Empirical Application: Returns to Schooling and Training

Empirical Application: Returns to Schooling and Training

Individuals who face "barriers to employment"

Y_2 above median 30-mo earnings

D_2 receiving job training program

Z_2 random assignment of the program

Y_1 above 80th pctle pre-program earnings

D_1 receiving high school diploma (or GED)

Z_1 number of schools per sq mile (e.g., Neal 97)

Empirical Application: Returns to Schooling and Training

Individuals who face "barriers to employment"

Y_2 above median 30-mo earnings

D_2 receiving job training program

Z_2 random assignment of the program

Y_1 above 80th pctle pre-program earnings

D_1 receiving high school diploma (or GED)

Z_1 number of schools per sq mile (e.g., Neal 97)

Consider $W_{\boldsymbol{\delta}} = E[Y_2(\boldsymbol{\delta})]$ and $= E[Y_1(\delta_1)] + E[Y_2(\boldsymbol{\delta})]$

Empirical Application: Returns to Schooling and Training

Individuals who face "barriers to employment"

Y_2 above median 30-mo earnings

D_2 receiving job training program

Z_2 random assignment of the program

Y_1 above 80th pctle pre-program earnings

D_1 receiving high school diploma (or GED)

Z_1 number of schools per sq mile (e.g., Neal 97)

Consider $W_{\boldsymbol{\delta}} = E[Y_2(\boldsymbol{\delta})]$ and $= E[Y_1(\delta_1)] + E[Y_2(\boldsymbol{\delta})]$

Data: JTPA (e.g., Abadie, Angrist & Imbens 02, Kitagawa & Tetenov 18)
+ NCES + US Census

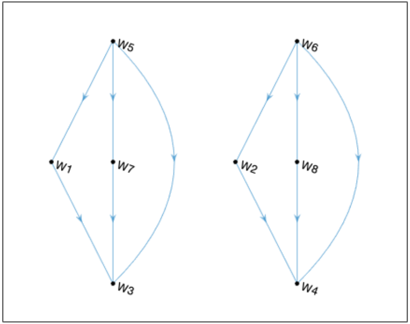
Estimation

Estimation of DAG and $\mathcal{D}_p^*$ is straightforward

- replace data distribution p in LP with sample frequencies $\hat{p}$, a vector of

$$\hat{p}_{\boldsymbol{y},\boldsymbol{d}|\boldsymbol{z}} = \sum_{i=1}^{N} 1\{\boldsymbol{Y}_i = \boldsymbol{y}, \boldsymbol{D}_i = \boldsymbol{d}, \boldsymbol{Z}_i = \boldsymbol{z}\} / \sum_{i=1}^{N} 1\{\boldsymbol{Z}_i = \boldsymbol{z}\}$$

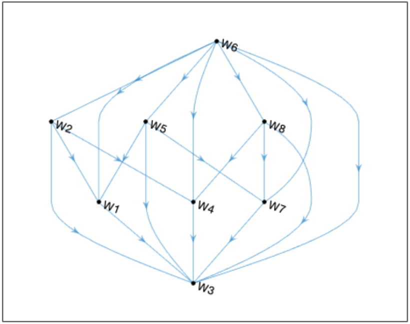
Policy Analysis with Schooling and Training



Regime #	δ_1	$\delta_2(1)$	$\delta_2(0)$
1	0	0	0
2	1	0	0
3	0	1	0
4	1	1	0
5	0	0	1
6	1	0	1
7	0	1	1
8	1	1	1

Figure: DAG of $W_{\boldsymbol{\delta}} = E[Y_2(\boldsymbol{\delta})]$ and Est'ed Set for $\boldsymbol{\delta}^*$ (under M2)

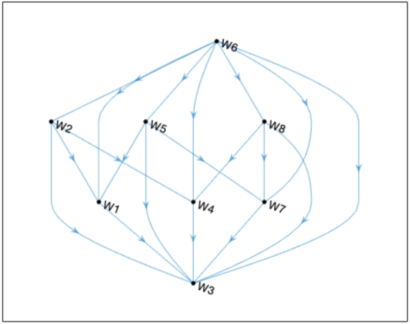
Policy Analysis with Schooling and Training



Regime #	δ_1	$\delta_2(1)$	$\delta_2(0)$
1	0	0	0
2	1	0	0
3	0	1	0
4	1	1	0
5	0	0	1
6	1	0	1
7	0	1	1
8	1	1	1

Figure: DAG of $W_{\boldsymbol{\delta}} = E[Y_1(\delta_1)] + E[Y_2(\boldsymbol{\delta})]$ and Est'ed Set for $\boldsymbol{\delta}^*$ (under M2)

Policy Analysis with Schooling and Training



Regime #	δ_1	$\delta_2(1)$	$\delta_2(0)$
1	0	0	0
2	1	0	0
3	0	1	0
4	1	1	0
5	0	0	1
6	1	0	1
7	0	1	1
8	1	1	1

Figure: DAG of $W_{\boldsymbol{\delta}} = E[Y_1(\delta_1)] + E[Y_2(\boldsymbol{\delta})]$ and Est'ed Set for $\boldsymbol{\delta}^*$ (under M2)

Policy Analysis with Schooling and Training

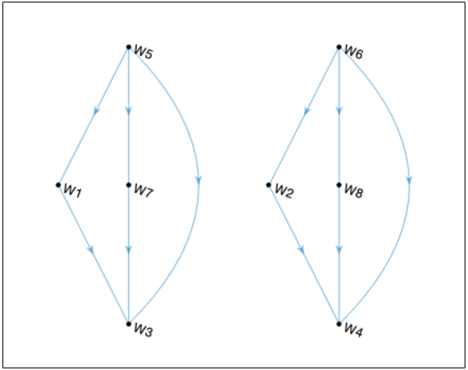


Figure: Partial Ordering with only Z_2 (under M2)

VI. Inference

Inference

For the inference on $\delta^*(\cdot)$, we construct confidence set for $\mathcal{D}_p^*$

- by seq of hypothesis tests (Hansen, Lunde & Nason 11)
  - to eliminate regimes that are significantly inferior to others
  - null hypotheses in terms of multiple ineq's as functions of p
    - e.g., Hansen 05, Andrews & Soares 10,...
  - no need to solve LPs for every bootstrap repetition
    - by using strong duality and vertex enumeration
- also useful for specification tests of (less palatable) identifying assumptions

Inference

Recall $W_\delta - W_{\delta'} = (A_\delta - A_{\delta'})q$ and

$$\begin{array}{l} L_{\delta,\delta'} = \min_{q\in\mathcal{Q}}(A_\delta - A_{\delta'})q \\ U_{\delta,\delta'} = \max_{q\in\mathcal{Q}}(A_\delta - A_{\delta'})q \end{array} \qquad s.t. \quad Bq = p$$

Dual programs with vertex enumeration (e.g., Avis & Fukuda 92):

$$L_{\delta,\delta'} = \max_{\lambda\in\Lambda_{\delta,\delta'}} -\tilde{p}'\lambda$$

$$U_{\delta,\delta'} = \min_{\lambda\in\tilde{\Lambda}_{\delta,\delta'}} \tilde{p}'\lambda$$

Null hypothesis for sequence of tests:

$$H_{0,\tilde{\mathcal{D}}} : L_{\delta,\delta'} \leq 0 \leq U_{\delta,\delta'} \qquad \forall \boldsymbol{\delta}, \boldsymbol{\delta}' \in \tilde{\mathcal{D}}$$

Inference

Recall $W_\delta - W_{\delta'} = (A_\delta - A_{\delta'})q$ and

$$\begin{array}{l} L_{\delta,\delta'} = \min_{q\in\mathcal{Q}}(A_\delta - A_{\delta'})q \\ U_{\delta,\delta'} = \max_{q\in\mathcal{Q}}(A_\delta - A_{\delta'})q \end{array} \qquad s.t. \quad Bq = p$$

Dual programs with vertex enumeration (e.g., Avis & Fukuda 92):

$$L_{\delta,\delta'} = \max_{\lambda\in\Lambda_{\delta,\delta'}} -\tilde{p}'\lambda$$

$$U_{\delta,\delta'} = \min_{\lambda\in\tilde{\Lambda}_{\delta,\delta'}} \tilde{p}'\lambda$$

Null hypothesis for sequence of tests:

$$H_{0,\tilde{\mathcal{D}}} : \tilde{p}'\lambda > 0 \qquad \forall\lambda \in \bigcup_{\delta,\delta'\in\tilde{\mathcal{D}}} (\Lambda_{\delta,\delta'} \cup \tilde{\Lambda}_{\delta,\delta'})$$

Inference

Let $\widehat{\mathcal{D}}_{CS}$ be the confidence set for $\mathcal{D}_p^*$

Algorithm (Constructing $\widehat{\mathcal{D}}_{CS}$)

Step 0. Initially set $\tilde{\mathcal{D}} = \mathcal{D}$.
Step 1. Test $H_{0,\tilde{\mathcal{D}}}$ *at level* α.
Step 2. If $H_{0,\tilde{\mathcal{D}}}$ *is not rejected, define* $\widehat{\mathcal{D}}_{CS} = \tilde{\mathcal{D}}$;
otherwise eliminate a regime $\boldsymbol{\delta}^-$ *from* $\tilde{\mathcal{D}}$ *and repeat from Step 1.*

- in Step 1, $T_{\tilde{\mathcal{D}}} \equiv \min_{\boldsymbol{\delta},\boldsymbol{\delta}' \in \tilde{\mathcal{D}}} t_{\boldsymbol{\delta},\boldsymbol{\delta}'}$ where $t_{\boldsymbol{\delta},\boldsymbol{\delta}'} \equiv \min_{\lambda \in \Lambda_{\boldsymbol{\delta},\boldsymbol{\delta}'} \cup \tilde{\Lambda}_{\boldsymbol{\delta},\boldsymbol{\delta}'}} t_\lambda$ with standard t-statistic t_λ
  - distribution of $T_{\tilde{\mathcal{D}}}$ can be estimated using bootstrap on p
- in Step 2, $\boldsymbol{\delta}^- \equiv \arg\min_{\boldsymbol{\delta} \in \tilde{\mathcal{D}}} \min_{\boldsymbol{\delta}' \in \tilde{\mathcal{D}}} t_{\boldsymbol{\delta},\boldsymbol{\delta}'}$.

Inference

Assumption CS

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

Proposition

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 $\boldsymbol{\delta}(\cdot) \notin \mathcal{D}^*_p$

Extension: Continuous Outcomes

This paper's analysis can be extended to the case of continuous Y_t

But the cost of incremental customization with Y_{t-1} can be high

- thus planner may want to employ a threshold-crossing rule:

$$\delta_t(1\{y_{t-1} \geq \gamma_{t-1}\}) \in \{0,1\}$$

Then a similar analysis can be done for optimal regime $(\boldsymbol{\delta}^*(\cdot), \boldsymbol{\gamma}^*)$

With continuous Y_t, two challenges in LP:

- q is infinite dimensional $\Longrightarrow$ approximate using Bernstein polynomials
- continuum of constraints $\Longrightarrow$ use mean absolute deviation of constraints
- Han & Yang 22

VI. Conclusions

Concluding Remarks

Propose a partial ID framework for optimal dynamic treatment regimes and welfares

- allowing for observational data

Sharp partial welfare ordering and ID'ed set for optimal regime

- via a set of linear programs

Applicability:

- e.g., when establishing rankings across multiple treatments or counterfactual policies

Concluding Remarks

Propose a partial ID framework for optimal dynamic treatment regimes and welfares

- allowing for observational data

Sharp partial welfare ordering and ID'ed set for optimal regime

- via a set of linear programs

Applicability:

- e.g., when establishing rankings across multiple treatments or counterfactual policies

Follow-ups:

1. inference on welfare with selected (set-ID'ed) regime
2. treatment allocation with distributional welfare

Thank You

Distribution of Counterfactual Outcome

With $T = 2$,

$$\begin{aligned} & \Pr[Y_2(\boldsymbol{\delta}) = 1] \\ = & \sum_{y_1 \in \{0,1\}} \Pr[Y_2(\delta_1, \delta_2(Y_1(\delta_1))) = 1 | Y_1(\delta_1) = y_1] \Pr[Y_1(\delta_1) = y_1] \end{aligned}$$

- for example, Regime #4 yields

$$\begin{aligned} \Pr[Y_2(\boldsymbol{\delta}_{\#4}) = 1] = & P[Y_1(1) = 1, Y_2(1,1) = 1] \\ & + P[Y_1(1) = 0, Y_2(1,0) = 1] \end{aligned}$$

Return

Monotonicity/Uniformity in D_t

Assumption M1

Conditional on $(\boldsymbol{Y}^{t-1}, \boldsymbol{D}^{t-1}, \boldsymbol{Z}^{t-1})$, either
$D_t(\boldsymbol{Z}^{t-1}, 1) \geq D_t(\boldsymbol{Z}^{t-1}, 0)$ w.p.1 or
$D_t(\boldsymbol{Z}^{t-1}, 1) \leq D_t(\boldsymbol{Z}^{t-1}, 0)$ w.p.1.

Assumption M1 imposes that there is no defying (complying) behavior in the decision D_t conditional on $(\boldsymbol{Y}^{t-1}, \boldsymbol{D}^{t-1}, \boldsymbol{Z}^{t-1})$

- without conditional on $(\boldsymbol{Y}^{t-1}, \boldsymbol{D}^{t-1}, \boldsymbol{Z}^{t-1})$, general non-uniform pattern of $\boldsymbol{Z}^t$ influencing $\boldsymbol{D}^t$

By extending Vytlacil 02, M1 is implied by

$$D_t = 1\{\pi_t(\boldsymbol{Y}^{t-1}, \boldsymbol{D}^{t-1}, \boldsymbol{Z}^t) \geq \nu_t\}$$

Monotonicity/Uniformity in Y_t

Assumption M2

M1 holds, and conditional on $(\boldsymbol{Y}^{t-1}, \boldsymbol{D}^{t-1}, \boldsymbol{Z}^{t-1})$, either $Y_t(\boldsymbol{D}^{t-1}, 1) \geq Y_t(\boldsymbol{D}^{t-1}, 0)$ w.p.1 or $Y_t(\boldsymbol{D}^{t-1}, 1) \leq Y_t(\boldsymbol{D}^{t-1}, 0)$ w.p.1.

Assumption M2 implicitly imposes rank similarity

- without conditional on $(\boldsymbol{Y}^{t-1}, \boldsymbol{D}^{t-1}, \boldsymbol{Z}^{t-1})$, general non-uniform pattern of $\boldsymbol{D}^t$ influencing $\boldsymbol{Y}^t$

Assumption M2 (and M1) does not assume the direction of monotonicity

M2 is implied by

$$Y_t = 1\{\mu_t(\boldsymbol{Y}^{t-1}, \boldsymbol{D}^t) \geq \varepsilon_t\}$$
$$D_t = 1\{\pi_t(\boldsymbol{Y}^{t-1}, \boldsymbol{D}^{t-1}, \boldsymbol{Z}^t) \geq \nu_t\}$$

Agent's Learning

Assumption L

$D_t(\boldsymbol{y}^{t-1}, \boldsymbol{d}^{t-1}, \boldsymbol{z}^t) \geq D_t(\tilde{\boldsymbol{y}}^{t-1}, \tilde{\boldsymbol{d}}^{t-1}, \boldsymbol{z}^t)$ w.p.1 for $(\boldsymbol{y}^{t-1}, \boldsymbol{d}^{t-1})$ and $(\tilde{\boldsymbol{y}}^{t-1}, \tilde{\boldsymbol{d}}^{t-1})$ s.t. $\left\|\boldsymbol{y}^{t-1} - \boldsymbol{d}^{t-1}\right\| < \left\|\tilde{\boldsymbol{y}}^{t-1} - \tilde{\boldsymbol{d}}^{t-1}\right\|$ (long memory) or $y_{t-1} - d_{t-1} < \tilde{y}_{t-1} - \tilde{d}_{t-1}$ (short memory).

Assumption L assumes agents have the ability to revise his next period's decision based on his memory

- e.g., consider $D_2(y_1, d_1)$
- agent who would switch his decision had he experienced $y_1 = 0$ after $d_1 = 1$, i.e., $D_2(0, 1) = 0$, would remain to take treatment had he experienced $y_1 = 1$, i.e., $D_2(1, 1) = 1$
- more importantly, if $D_2(0, 1) = 1$, it should only because of unobserved preference, *not* because he cannot learn from the past, i.e., $D_2(1, 1) = 0$ cannot happen

Return

Markovian Structure

Assumption K

$Y_t | (\boldsymbol{Y}^{t-1}, \boldsymbol{D}^t) \stackrel{d}{=} Y_t | (Y_{t-1}, D_t)$ and
$D_t | (\boldsymbol{Y}^{t-1}, \boldsymbol{D}^{t-1}, \boldsymbol{Z}^t) \stackrel{d}{=} D_t | (Y_{t-1}, D_{t-1}, Z_t)$.

In terms of the triangular model under M2, Assumption K implies

$$\begin{aligned} Y_t &= 1\{\mu_t(Y_{t-1}, D_t) \geq \varepsilon_t\} \\ D_t &= 1\{\pi_t(Y_{t-1}, D_{t-1}, Z_t) \geq \nu_t\} \end{aligned}$$

- a familiar structure of dynamic discrete choice models in the literature

Return