Policy Choice in Time Series by Empirical Welfare Maximization

Toru Kitagawa (Brown), Weining Wang (Bristol), Mengshan Xu (Mannheim)

Introduction

- **Causal Inference for what!?**
- The treatment choice literature initiated by Manski (04, Ecta) uses the statistical decision theory to link causal inference and planner's policy decision in the microeconometrics context.
- How about in macro/finance? **Statistical treatment (policy) choice with time-series data?**
- E.g.,
  1. FOMC's monetary policy choice?
  2. Central bank's intervention to exchange markets?
  3. Policy choice of relaxing or tightening the Covid containment policy?

Motivating Example and Preview

- The Bank of Japan (BOJ) occasionally intervenes the JPY/USD exchange market to stabilize the exchange rate

- Given the time-series observations of exchange rate, treatment (intervened or not), and other variables, should BOJ intervene today?

- Assuming treatments are randomized conditional on observables in the data, we propose T-EWM: estimate welfare impact of the today's policy choice using the past data and maximize it.

- We formulate the decision problem using potential outcome time-series and study theoretical guarantee of the proposal in terms of welfare convergence

Review: (Static) Statistical Treatment Choice

Assume randomized controlled trial (RCT) or observational data for
$i = 1, \ldots, n$ iid individuals

- $X_i \in \mathcal{X}$ - pre-treatment observed covariates (e.g., education, previous earnings)
- $W_i \in \{0, 1\}$ - randomized treatment (possibly conditional on X_i) (e.g., job training program)
- $Y_i \in \mathbb{R}$ - post-treatment observed outcome (e.g., employment, income)

- An assignment rule $g : \mathcal{X} \to \{0, 1\}$
- Let $(Y(1), Y(0))$ be the potential outcomes. Additive (utilitarian) welfare:

$$\mathcal{W}(g) = E\left[Y(1) \cdot g(X) + Y(0) \cdot (1 - g(X))\right].$$

Manski's Empirical Success Rule

- **Goal**: Use the data (already collected) to learn policy rule $\hat{g}$ that performs well in terms of $\mathcal{W}(\hat{g})$.

- For discrete X case, Manski (04, Ecta) proposes the **conditional empirical success (CES) rule**:

  $$\hat{g}_{CES}(x) \equiv 1\{\hat{\tau}(x) \geq 0\},$$

  where $\hat{\tau}(x)$ is the difference-in-means estimator for CATE,

  $$\hat{\tau}(x) = \widehat{E}[Y|W = 1, X = x] - \widehat{E}[Y|W = 0, X = x] \tag{1}$$

- Manski assesses the performance of the CES rule by bounding the **expected welfare regret**,

  $$E_{P^n}\left[\max_g \mathcal{W}(g) - \mathcal{W}(\hat{g}_{CES})\right] \leq \frac{C}{\sqrt{n}}.$$

Extending to Time-Series?

Questions and challenges:

1. Policy impact can have dynamic causal effects and can be heterogeneous (nonstationary). External validity?
2. Observations are no longer iid.
3. What could be a reasonable welfare objective? How to obtain its empirical analogue?
4. Can we obtain welfare regret convergence? Under what condition? What rate?

Baseline Setting

- Suppose a social planner is at the beginning of period T, wants to perform the policy choice at T, W_T, based on the information available at the beginning of T.

- The planner has access to time-series observations: for $t = 0, 1, 2, \ldots, T - 1$.

- $Y_t \in \mathbb{R}$: outcome, e.g., output, unemployment, stock price, Covid deaths, etc.

- $W_t \in \{0, 1\}$: (manipulatable) treatment, e.g., high/low target rate, lockdown or not, structural policy shock, etc

- $Z_t \in \mathbb{R}$: contextual information, e.g., macroeconomic indices other than Y_t,

- Denote the sample by $X_{0:T-1} = (Y_{0:T-1}, W_{0:T-1}, Z_{0:T-1})$, given initial value X_0.
- A **treatment path** $w_{0:T} \in \{0, 1\}^{T+1}$.
- Potential outcomes indexed by the treatment paths: for each $t = 1, \ldots, T$, $Y_t(w_{0:T})$ denotes the counterfactual outcome at period t when the treatment path were exogenously set to $w_{0:T}$.

Definition: Potential Outcome Time-Series (Bojinov & Shephard 19 JASA, and Rambachan & Shephard 21)

1. **Non-anticipating potential outcomes**: For each $t = 1, \ldots, T$,

   $$Y_t(w_{0:t}, w_{t+1:T}) = Y_t(w_{0:t}, w'_{t+1:T})$$

   holds a.s. for any $w_{0:t}$, and $w_{t+1:T} \neq w'_{t+1:T}$, i.e., the potential outcomes are indexed only by the past and current treatments

2. **Non-anticipating treatment paths (sequential unconfoundedness)**: For every $t = 1, \ldots, T - 1$ and $s = t, t + 1, \ldots, T$, and any $w_{t:T}$,

   $$W_t \perp Y_s(W_{0:t-1}, w_{t:s}) | X_{0:t-1},$$

   In the data, W_t is randomized once conditioned on the history of observables

3. **Observed data**: $Y_t = Y_t(W_{0:t})$.

Policies and Welfare

- **One-period (nonrandomized) policy**: $g : \mathcal{X}_{0:T-1} \to \{0, 1\}$
- The planner wants to maximize **one-period conditional welfare**:

$$\mathcal{W}_T(g|X_{0:T-1})$$
$$\equiv E\left[Y_T(W_{0:T-1}, 1)g(X_{0:T-1}) + Y_T(W_{0:T-1}, 0)(1 - g(X_{0:T-1}))|X_{0:T-1}\right].$$

- cf. in the cross-sectional setting, **we focus on the unconditional welfare**.

Let's simplify (and generalize later)

> **Toy model: one-period Markovian model**
>
> Bivariate time-series, $X_t = (Y_t, W_t) \in \mathbb{R} \times \{0, 1\}$, $t = 0, 1, 2, \ldots$, given X_0.
>
> ① **Markovian Exclusion**: For each $t = 2, \ldots, T$,
>
> $$Y_t(w_{0:t-2}, w_{t-1}, w_t) = Y_t(w'_{0:t-2}, w_{t-1}, w_t) \equiv Y_t(w_{t-1}, w_t)$$
>
> for all $w_{0:t-2} \neq w'_{0:t-2}$, and (w_{t-1}, w_t).
>
> ② **Markovian Exogeneity**: For each $t = 1, \ldots, T - 1$ and $w_{0:t}$,
>
> $$(Y_t(w_{0:t}), W_t) \perp X_{0:t-1} | W_{t-1}, \text{ and } Y_T(w_{T-1}, w_T) \perp X_{0:T-1} | W_{T-1}.$$

- Under this simplification, the planner's objective (one-period social welfare) satisfies

$$
\begin{aligned}
&\mathcal{W}_T(g|X_{1:T-1})\\
&= E\left[Y_T(W_{T-1},1)g(X_{0:T-1}) + Y_T(W_{T-1},0)(1-g(X_{0:T-1}))|X_{0:T-1}\right]\\
&\qquad (\because \text{ Markovian exclusion })\\
&= E[Y_T(W_{T-1},1)|W_{T-1}]g(X_{0:T-1}) + E[Y_T(W_{T-1},0)|W_{T-1}](1-g(X_{0:T-1}))\\
&\qquad (\because \text{ Markovian exogeneity })\\
&= (E[Y_T(W_{T-1},1) - Y_T(W_{T-1},0)|W_{T-1}]) \cdot g(X_{0:T-1}) + E[Y_T(W_{T-1},0)|W_{T-1}]
\end{aligned}
$$

- We can reduce $g(\cdot)$ to a binary map of W_{T-1} without loss of the welfare and focus on maximizing

$$
\mathcal{W}_T(g|W_{T-1}) \equiv E\left[Y_T(W_{T-1},1)g + Y_T(W_{T-1},0)(1-g)|W_{T-1}\right].
$$

Empirical Welfare

- Empirical analogue of $\mathcal{W}_T(g|W_{T-1} = w)$? Try a historical simple average.

Policy rule: Time-Series CES

We obtain $\hat{g}_{CES}(w)$ by maximizing

$$\widehat{\mathcal{W}}(g|W_{t-1} = w) = \frac{1}{T(w)} \sum_{t:W_{t-1}=w} \left[\frac{Y_t W_t}{e_t(W_{t-1})}g + \frac{Y_t(1-W_t)}{1-e_t(W_{t-1})}(1-g) \right]$$

in $g \in \{0, 1\}$, where $T(w) = |\{t : W_{t-1} = w\}|$ and $e_t(W_{t-1}) = \Pr(W_t = 1|W_{t-1})$ assumed to satisfy a strict overlap condition

---

Any concentration of $\widehat{\mathcal{W}}(g|W_{t-1} = w)$?

- Let $\{\mathcal{F}_t = \sigma(X_{1:t-1}) : t = 1, \ldots, T-1\}$ be the filtration and consider decomposing $\widehat{\mathcal{W}}(g|W_{t-1} = w)$ into sums of innovations and conditional expectations w.r.t $\{\mathcal{F}_t\}$ (Bojinov & Shephard 19).

$$\widehat{\mathcal{W}}(g|W_{t-1} = w) = \frac{1}{T(w)} \sum_{t=1}^{T} \xi_{t,w}(g) + \bar{\mathcal{W}}(g|w)$$

where

$$\bar{\mathcal{W}}(g|w) \equiv \frac{1}{T(w)} \sum_{t:W_{t-1}=w} E\left[Y_t(W_{t-1}, 1)g + Y_t(W_{t-1}, 0)(1-g)|\mathcal{F}_{t-1}\right],$$

$$\xi_{t,w}(g) \equiv 1\{W_{t-1} = w\} \cdot \begin{bmatrix} \frac{Y_t W_t}{e_t(W_{t-1})}g + \frac{Y_t(1-W_t)}{1-e_t(W_{t-1})}(1-g) \\ -E\left[Y_t(W_{t-1}, 1)g + Y_t(W_{t-1}, 0)(1-g)|\mathcal{F}_{t-1}\right] \end{bmatrix}$$

- Under Markovian exogeneity and sequential unconfoundedness:

$$\bar{W}(g|w) \equiv \frac{1}{T(w)} \sum_{t: W_{t-1}=w} E\left[Y_t(W_{t-1}, 1)g + Y_t(W_{t-1}, 0)(1-g)|W_{t-1} = w\right],$$

$$E[\xi_{t,w}(g)|\mathcal{F}_{t-1}] \equiv 1\{W_{t-1} = w\} \cdot \begin{bmatrix} E\left[\frac{Y_t W_t}{e_t(W_{t-1})}g + \frac{Y_t(1-W_t)}{1-e_t(W_{t-1})}(1-g)|W_{t-1}\right] \\ -E\left[Y_t(W_{t-1}, 1)g + Y_t(W_{t-1}, 0)(1-g)|W_{t-1}\right] \end{bmatrix}$$
$$= 0$$

- I.e., $\xi_{t,w}(g)$ is a martingale difference sequence (MDS), and the empirical welfare centered at $\bar{W}(g|w)$ is an average of MDS w.r.t $\{\mathcal{F}_t\}$,

$$\widehat{W}(g|w) - \bar{W}(g|w) = \frac{1}{T(w)} \sum_{t=1}^{T} \xi_{t,w}(g),$$

Key assumption: Invariance of welfare ordering
Bridge the period-T welfare regret and the past average welfare regret

Assumption: Invariance of welfare ordering

Let $g^* = \arg\max_g \mathcal{W}_T(g|W_{T-1} = w)$. There exists $c > 0$ such that for any w and $g \in \{0, 1\}$,

$$\mathcal{W}_T(g^*|w) - \mathcal{W}_T(g|w) \le c(\bar{W}(g^*|w) - \bar{W}(g|w))$$

- The welfare ordering of the policies today agrees with the historical average over the periods sharing the conditioning events similar to today's
- Weaker than stationarity of welfare(i.e., $\mathcal{W}_t(g|w)$ independent of t)
- In structural MA(2): $Y_t = \alpha_t + \beta_t W_t + \gamma_t W_{t-1} + \epsilon_t$, this assumption holds if

$$sign(\beta_T) = sign\left(\frac{1}{T(w)} \sum_{t:W_{t-1}=w} \beta_t\right).$$

- Putting everything together,

$$\mathcal{W}_T(g^*|w) - \mathcal{W}_T(\hat{g}_{CES}|w) \le c(\bar{\mathcal{W}}(g^*|w) - \bar{\mathcal{W}}(\hat{g}_{CES}|w))$$

$$\le 2c \sup_g |\widehat{\mathcal{W}}(g|w) - \bar{\mathcal{W}}(g|w)|$$

- Hence, the average conditional welfare regret can be bounded as follows:

$$E_P[\mathcal{W}_T(g^*|w) - \mathcal{W}_T(\hat{g}_{CES}|w)]$$

$$\le 2cE_P \left[\sup_g \left| \frac{1}{T(w)} \sum_{t=1}^{T-1} \xi_{t,w}(g) \right| \right]$$

- Imposing additional regularity (Y_t bounded for all t), we can apply Freedman's large deviation inequality for MDS to obtain a finite-sample regret bound uniform in w and P.

$$E_P[\mathcal{W}_T(g^*|w) - \mathcal{W}_T(\hat{g}_{CES}|w)] \leq \frac{C}{\sqrt{T-1}}.$$

- **Result**: Under the imposed assumptions, Manski's CES-rule can be generalized to time-series and attain $\sqrt{T}$-convergence.

Extension: more covariates?

- Maintaining the one-period Markovian model, consider adding continuous variables to X_{t-1}:

  $$X_{t-1} = (W_{t-1}, Y_{t-1}, Z_{t-1}) \in \{0, 1\} \times \mathbb{R} \times \mathbb{R}^d.$$

- The conditional welfare $\mathcal{W}_T(g|x_{T-1})$ as planner's objective:

  $$\mathcal{W}_T(g|x_{T-1})$$
  $$\equiv E\left[Y_T(W_{T-1}, 1)g + Y_T(W_{T-1}, 0)(1-g)|X_{T-1} = x_{T-1}\right].$$

- Conditioning on the continuous variables, a simple sample analogue is not available

- **Our approach**: analogous to EWM, put a class of policies $\mathcal{G} \equiv \{g\}$, and optimize a sample analogue of the **unconditional** welfare over $\mathcal{G}$.

- Define the **unconditional welfare** at period T:

$$\mathcal{W}_T(g) \equiv E\left[Y_T(W_{T-1}, 1)g(X_{T-1}) + Y_T(W_{T-1}, 0)(1 - g(X_{T-1}))\right]$$

Key assumptions for EWM for time-series:

1. **Correct specification** at $x_{T-1} \in \mathcal{X}_{T-1}$,

$$\arg\sup_{g \in \mathcal{G}} \mathcal{W}_T(g) \subset \arg\sup_{g \in \mathcal{G}} \mathcal{W}_T(g|x_{T-1}),$$

A sufficient condition:
$\{x_{T-1} : E[Y_T(W_{T-1}, 1) - Y_T(W_{T-1}, 0)|x_{T-1}] \geq 0\} \in \mathcal{G}$.

- With additional assumption on the marginal distribution of X_{T-1}, we can bound the conditional welfare regret by the unconditional welfare regret, i.e., $\exists c > 0$ such that $\forall G \in \mathcal{G}$,

$$\sup_{g \in \mathcal{G}} \mathcal{W}_T(g|x_{T-1}) - \mathcal{W}_T(g|x_{T-1}) \leq c(\sup_{g \in \mathcal{G}} \mathcal{W}_T(g) - \mathcal{W}_T(g)).$$

EWM policy and regret bounds

$$\hat{g}^{EWM} \in \arg \max_{g \in \mathcal{G}} \frac{1}{T-1} \sum_{t=1}^{T-1} \left[\frac{Y_t W_t}{e_t(X_{t-1})} g(X_{t-1}) + \frac{Y_t(1-W_t)}{1-e_t(X_{t-1})}(1-g(X_{t-1})) \right]$$

Theorem: Time-series EWM regret bounds

Assume the one-period Markovian model with bounded outcome, correct specification of $\mathcal{G}$, and a complexity restriction of $\mathcal{G}$ (i.e., $\{\xi_{t,x_{t-1}}(G) : G \in \mathcal{G}\}$ has finite VC dimension v), there exists a constant $C > 0$ such that

$$E_P[\sup_{g \in \mathcal{G}} \mathcal{W}_T(g|x_{T-1}) - \mathcal{W}_T(\hat{g}_{EWM}|x_{T-1})] \leq C \sqrt{\frac{v \log T}{T}}$$

holds.

Extension: longer dependence?

- Introducing qth-order Markovian structure, $2 \le q < \infty$, is feasible
- $q = \infty$ is more challenging, though common SVAR modeling implies so, e.g.,

$$Y_t = \sum_{h=0}^{\infty} (\theta_h W_{t-h} + \phi_h \epsilon_{t-h})$$

- Approximate ∞-Markov by finite order m-th Markov and run T-EWM. The regret bounds depend also on the approximation bias,

$$\mathcal{W}_T(g^* | \mathcal{F}_{T-1}) - \mathcal{W}_T(\hat{g}_{CES} | \mathcal{F}_{T-1})$$
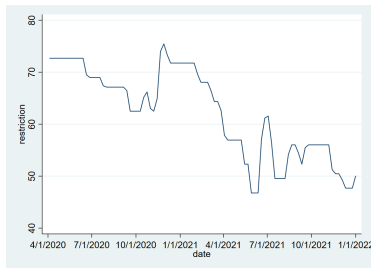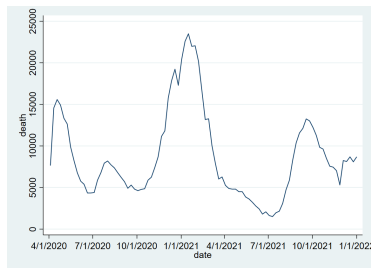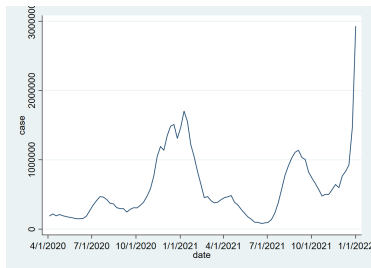$$\le 2c \sup_g |\widehat{\mathcal{W}}(g | w_{T-m:T-1}) - \bar{\mathcal{W}}(g | \mathcal{F}_{T-1})| + bias(m),$$

where $bias(m) = 2 \sup_g |\mathcal{W}_T(g | w_{-\infty:T-1}) - \mathcal{W}_T(g | w_{T-m:T-1})|$.

Empirical Example (preliminary)

- Binary treatment: relax the covid restriction level ($W = 0$) or not ($W = 1$)
- Outcome (welfare): $-1\times$ two-week ahead deaths in US:
- Contextual information for the model of propensity score: $X_t = (\text{cases}_t, \text{deaths}_t, \text{change of cases}_t, \text{change of deaths}_t, \text{restriction level}_t, \text{vaccine coverage}_t, \text{economic condition}_t)$
- We maximize the expected welfare, $E(-1 \cdot \text{deaths}_{T+1})$, over the set of quadrant (threshold) policies based on (change of deaths$_{T-1}$, restriction level$_{T-1}$)
- Data source:
    - Cases, deaths, and vaccinations: The CDC website
    - Restriction level: The Oxford Stringency Index
    - Economic condition: The Lewis-Mertens-Stock business conditions index
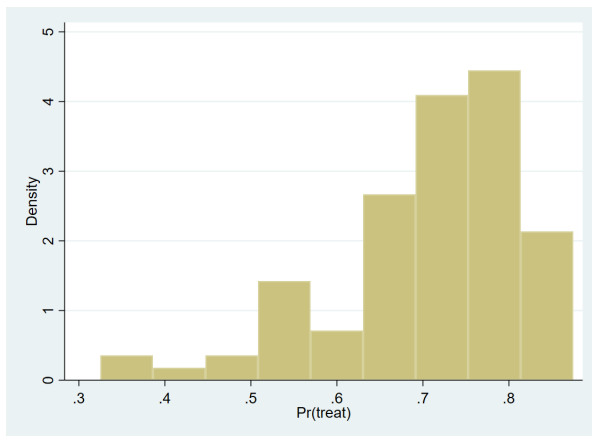
Empirical Example

Data overview (from the top left to the bottom right: cases, deaths, restriction level, and economic condition)
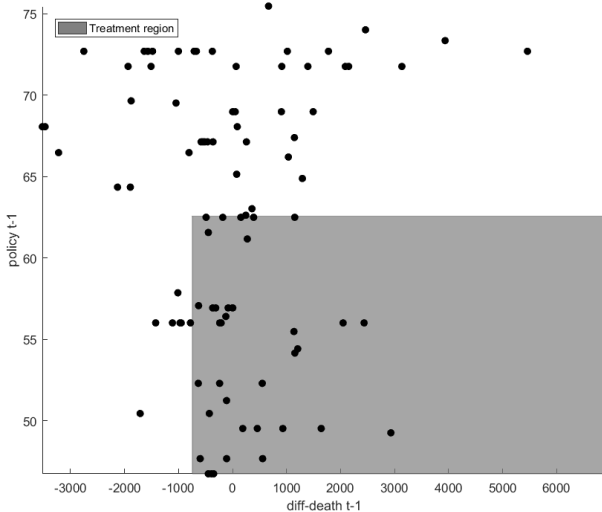
Empirical Example (preliminary)

The propensity score model and the estimated propensity score

$$\log\left(\frac{\Pr(\text{keeping or increasing restriction at week } t)}{\Pr(\text{decreasing restriction at week } t)}\right) = \alpha + \beta X_{t-1}$$

Empirical Example (preliminary)

Policy choice

Literature

- Potential outcome time-series: Angrist, Jorda, Kuersteiner (18 JBES), Bojinov & Shephard (19 JASA), Rambachan & Shephard (21)

- Treatment choice: Manski(04 Ecta), Dehejia (05 JoE), Hirano & Porter (09 Ecta), Stoye (09, 12 JoE), Tetenov (12 JoE), Chamberlain (11 Handbook Chap), Kitagawa & Tetenov (18 Ecta), Mbakop & Tabord-Meehan (21 Ecta), Athey & Wager (21 Ecta), and more

- (Individualized) Dynamic treatment regimes (Large *N* - short *T* panel): Murphy (03 JRSSB), Zhao, Zeng, Laber, Kosorok (15 JASA), Sakaguchi (21), among many others.

- Empirical risk minimization in time-series: Jiang & Tanner (10 ET), Brownlees & Guðmundsson (21), Brownlees & Llorens-Terrazas (21)

Concluding Remarks

- Propose a framework for data-driven policy choice in the time-series setting

- Non-trivial challenges distinguishing the time-series policy choice from the static one

- Under sequential unconfoundedness, finite-order Markovian, and invariance of welfare ordering, Manski's CES rule can be extended to time-series

- With the correct specification assumption added, EWM can be extended to time-series as well

- Our approach relies only on the causal and conditional independence structure among the variables, and is free from functional form, stationarity (in the outcome generating process), or prior distribution thereof.