Semiparametric Models for Dynamic Treatment Effects and Mediation Analyses with Observational Data

Sukjin Han & Sungwon Lee

U of Bristol & Sogang University

11 March 2023

Warwick-Turing Economics Data Science Workshop

Dynamics of Treatments Affecting Outcomes

A sequence of interventions
interacting with a sequence of outcomes over periods

Dynamics of Treatments Affecting Outcomes

A sequence of interventions
interacting with a sequence of outcomes over periods

To design informed policies,
important to understand the dynamic causal mechanism

Dynamics of Treatments Affecting Outcomes

Example: Schooling (D_1) and post-school training (D_2) on employment status (Y_1 and Y_2)

Dynamics of Treatments Affecting Outcomes

Example: Schooling (D_1) and post-school training (D_2) on employment status (Y_1 and Y_2)

Post-training employment status (Y_2) is affected by...

(i) the job training (D_2),

(ii) directly by HS degree (D_1),

(iii) indirectly by HS degree through previous status (Y_1) via...

- state dependence and
- time-invariant heterogeneity
- i.e., Y_1 is a mediator

Dynamics of Treatments Affecting Outcomes

Example: Schooling (D_1) and post-school training (D_2) on employment status (Y_1 and Y_2)

Post-training employment status (Y_2) is affected by...

(i) the job training (D_2),

(ii) directly by HS degree (D_1),

(iii) indirectly by HS degree through previous status (Y_1) via...

- state dependence and
- time-invariant heterogeneity
- i.e., Y_1 is a mediator

Want to understand various channels of causal effects

Challenge: Dynamically Endogenous Selection

In observational settings, individuals make dynamically endogenous decisions whether to select into treatments

Challenge: Dynamically Endogenous Selection

In observational settings, individuals make dynamically endogenous decisions whether to select into treatments

Example (conti'ed): Workers decide (D_2) to participate in the training based on...

- their schooling decision (D_1)
- previous labor market outcome (Y_1)
- prospect of labor market outcome (counterfactual Y_2)

Challenge: Dynamically Endogenous Selection

In observational settings, individuals make dynamically endogenous decisions whether to select into treatments

Example (conti'ed): Workers decide (D_2) to participate in the training based on...

- their schooling decision (D_1)
- previous labor market outcome (Y_1)
- prospect of labor market outcome (counterfactual Y_2)

Even in experimental settings, especially if D_1 and D_2 are the same kind, non-compliance due to learning

Our Approach: Instrumental Variables

Want to address this challenge, while remaining flexible in...

- modeling dynamics and
- treatment heterogeneity

$\Longrightarrow$ We use sequence of IVs

Our Approach: Instrumental Variables

Want to address this challenge, while remaining flexible in...

- modeling dynamics and
- treatment heterogeneity

$\Longrightarrow$ We use sequence of IVs

We consider the most challenging case of binary IVs

Candidates:

- multi-period/stage experiments
- sequence of natural experiments
- sequential fuzzy regression continuity

Our Approach: Instrumental Variables

Want to address this challenge, while remaining flexible in...

- modeling dynamics and
- treatment heterogeneity

$\Longrightarrow$ We use sequence of IVs

We consider the most challenging case of binary IVs

Candidates:

- multi-period/stage experiments
- sequence of natural experiments
- sequential fuzzy regression continuity

Our theory immediately extends with IVs of richer variation

Our Approach: Instrumental Variables

Example (conti'ed): For HS diploma (D_1) and job training (D_2)

- Z_1: close/far from HS
- Z_2: randomization from JTPA
- one of our empirical applications

Sequential fuzzy RD:

- Z_t: eligibility from running variables

Sequential experiments:

- A/B testing, multi-stage field experiments

This Paper: Semiparametric Models for Dynamic Effects

Consider a class of semiparametric models for dynamic treatment and mediation effects

- nonparametric outcome and selection equations
- nonparametric marginal distribution of unobservables
- multivariate parametric copula for dependence of unobservables

This Paper: Semiparametric Models for Dynamic Effects

Why semiparametric?

1. Point ID with alternative assumptions while being flexible
   - existing results:
     - irreversible treatments (Heckman & Navarro 07, Heckman, Humphries & Veramendi 16),
     - rely on IVs with large support or extra exogenous variables (above refs, Han 21)
     - resort to partial identification (Han 23)
2. Avoid the curse of dim with nonparametric joint dist of unobs
   - selection endogeneity parameters as byproduct
3. Avoid misspecification of marginals of unobs
   - treatment effects are direct functions of marginals

Other Related Work

Dynamic discrete choice models, dynamic sample selection models
(Honoré & Kyriazidou 00, Kyriazidou 01, Honoré & Tamer 06, Torgovitsky 19, Honoré & Weidner 21)

- main focus is state dependence
- linear index; dependence of unobs via fixed effects
- no need of IVs
- but requirement on T

Policy evaluation with multiple treatments
(Heckman & Pinto 18, Lee & Salanié 18, Balat & Han 23)

Dynamic treatment regimes
(Murphy et al. 01, Cui & Tchetgen Tchetgen 21, Qiu et al. 21, Han 23)

I. Model and Assumptions

Two-Period Semiparametric Model

Consider $T = 2$ and binary Y_t (for simplicity), and write

$$
\begin{aligned}
Y_2 &= 1[\mu_2(Y_1, D) \geq U_2(Y_1, D)] \\
D_2 &= 1[\pi_2(Y_1, D_1, Z_2) \geq V_2] \\
Y_1 &= 1[\mu_1(D_1) \geq U_1(D_1)] \\
D_1 &= 1[\pi_1(Z_1) \geq V_1]
\end{aligned}
$$

where $D \equiv (D_1, D_2)$ and unobs are normalized as $U[0, 1]$

- (general T, continuous Y_t and covariate X in the paper)

Two-Period Semiparametric Model

Consider $T = 2$ and binary Y_t (for simplicity), and write

$$\begin{aligned} Y_2 &= 1[\mu_2(Y_1, D) \geq U_2(Y_1, D)] \\ D_2 &= 1[\pi_2(Y_1, D_1, Z_2) \geq V_2] \\ Y_1 &= 1[\mu_1(D_1) \geq U_1(D_1)] \\ D_1 &= 1[\pi_1(Z_1) \geq V_1] \end{aligned}$$

where $D \equiv (D_1, D_2)$ and unobs are normalized as $U[0,1]$

- (general T, continuous Y_t and covariate X in the paper)

$U_1(d_1)$ and $U_2(y_1, d)$ for flexible heterogeneity

- let $Y_1(d_1)$ and $Y_2(y_1, d)$ be the counterfactual outcomes
  - e.g., $Y_1(1) = 1[\mu_1(1) \geq U_1(1)]$ and $Y_1(0) = 1[\mu_1(0) \geq U_1(0)]$
  - compare to: $Y_1(d_1) = 1[\mu_1(d_1) \geq U_1]$ (rank invariance)
- even though $U_1(d_1)$ and $U_2(y_1, d)$ are normalized as $U[0,1]$, allow distinct selection patterns

Two-Period Semiparametric Model

$$
\begin{aligned}
Y_2 &= 1[\mu_2(Y_1, D) \geq U_2(Y_1, D)] \\
D_2 &= 1[\pi_2(Y_1, D_1, Z_2) \geq V_2] \\
Y_1 &= 1[\mu_1(D_1) \geq U_1(D_1)] \\
D_1 &= 1[\pi_1(Z_1) \geq V_1]
\end{aligned}
$$

- consistency:

$$
\begin{aligned}
Y_1 &= D_1 Y_1(1) + (1 - D_1) Y_1(0) \\
Y_2 &= \sum_{y_1, d \in \{1,0\}^3} 1[Y_1 = y_1, D = d] Y_2(y_1, d)
\end{aligned}
$$

- therefore Y_2 is a function of a full vector of $(U_2(1,1,1), U_2(1,1,0), U_2(1,0,1), U_2(1,0,0), U_2(0,1,1), U_2(0,1,0), U_2(0,0,1), U_2(0,0,0))$

Two-Period Semiparametric Model

$$
\begin{aligned}
Y_2 &= 1[\mu_2(Y_1, D) \geq U_2(Y_1, D)] \\
D_2 &= 1[\pi_2(Y_1, D_1, Z_2) \geq V_2] \\
Y_1 &= 1[\mu_1(D_1) \geq U_1(D_1)] \\
D_1 &= 1[\pi_1(Z_1) \geq V_1]
\end{aligned}
$$

Special case (dynamic Roy model):

$$
\begin{aligned}
\pi_2(Y_1, D_1, Z_2) - V_2 &\equiv \delta_2(Z_2) + \mu_2(Y_1, D_1, 1) - \mu_2(Y_1, D_1, 0) \\
&\quad - (U_2(Y_1, D_1, 1) - U_2(Y_1, D_1, 0)) \\
\pi_1(Z_1) - V_1 &\equiv \delta_1(Z_1) + \mu_1(1) - \mu_1(0) - (U_1(1) - U_1(0))
\end{aligned}
$$

- not possible with scalar unobservable

Two-Period Semiparametric Model

Assumption C

$(V_1, V_2, U_1(d_1), U_2(y_1, d)) \sim C\,(v_1, v_2, u_1, u_2; \Sigma(y_1, d))$.

- C makes the model semiparametric
  - nonparametric marginals are subsumed in $\mu_t(\cdot)$ and $\pi_t(\cdot)$
- notable elements in $\Sigma(y_1, d, x)$ are:
  - $\rho_{V_1, U_1(d_1)}$ and $\rho_{V_t, U_2(y_1, d)}$ (for $t = 1, 2$)
  - which capture the (sign/degree of) treatment-state-specific selection

II. Dynamic Treatment and Mediation Effects

Causal Objects of Interest

If we identify $\mu_1(d_1)$, $\mu_2(y_1, d)$, $\Sigma(y_1, d)$, we can identify various dynamic treatment and mediation effects

Relevant counterfactual outcomes:

$$Y_2(y_1, d),\ Y_2(d),\ Y_2(d_2),\ Y_2(d_1),\ Y_1(d_1)$$

- note that $Y_2(Y_1, D_1, d_2) = Y_2(d_2)$,
- but $Y_2(Y_1(d_1), d) = Y_2(d)$ and $Y_2(d_1, D_2(d_1)) = Y_2(d_1)$ where $D_2(d_1)$ is counterfactual treatment

Causal Objects of Interest

If we identify $\mu_1(d_1)$, $\mu_2(y_1, d)$, $\Sigma(y_1, d)$, we can identify various dynamic treatment and mediation effects

Relevant counterfactual outcomes:

$$Y_2(y_1, d),\ Y_2(d),\ Y_2(d_2),\ Y_2(d_1),\ Y_1(d_1)$$

- note that $Y_2(Y_1, D_1, d_2) = Y_2(d_2)$,
- but $Y_2(Y_1(d_1), d) = Y_2(d)$ and $Y_2(d_1, D_2(d_1)) = Y_2(d_1)$ where $D_2(d_1)$ is counterfactual treatment

Building blocks:

$$E[Y_2(y_1, d)] = \Pr[U_2(y_1, d) \leq \mu_2(y_1, d)] = \mu_2(y_1, d)$$
$$E[Y_1(d_1)] = \Pr[U_1(d_1) \leq \mu_1(d_1)] = \mu_1(d_1)$$

Causal Objects of Interest

Also, for example,

$$
\begin{aligned}
E[Y_2(d)] &= \sum_{y_1 \in \{0,1\}} \Pr[Y_1(d_1) = y_1, Y_2(y_1, d) = 1] \\
&= C(\mu_1(d_1), \mu_2(1, d); \rho_{U_1(d_1), U_2(1,d)}) \\
&+ \mu_2(0, d) - C(\mu_1(d_1), \mu_2(0, d); \rho_{U_1(d_1), U_2(0,d)})
\end{aligned}
$$

$$
\begin{aligned}
E[Y_2(d_1)] &= \sum_{y_1, d_2 \in \{0,1\}^2} \Pr[Y_1(d_1) = y_1, D_2(d_1) = d_2, Y_2(y_1, d) = 1] \\
&= \cdots
\end{aligned}
$$

Dynamic Treatment and Mediation Effects

Dynamic treatment effects (type 1):

$$E[Y_2(\tilde{y}_1, \tilde{d}) - Y_2(y_1, d)] \quad \text{and} \quad E[Y_1(\tilde{d}_1) - Y_1(d_1)]$$

- dynamic complementarity: $\tilde{y}_1 = y_1$ and comparing $\tilde{d} = (0, 1)$ and $d = (0, 0)$ with $\tilde{d} = (1, 1)$ and $d = (1, 0)$:

$$E[Y_2(y_1, 0, 1) - Y_2(y_1, 0, 0)] \quad \text{vs} \quad E[Y_2(y_1, 1, 1) - Y_2(y_1, 1, 0)]$$

- state dependence: $\tilde{d} = d$ and $\tilde{y}_1 = 1$ and $y_1 = 0$:

$$E[Y_2(1, d) - Y_2(0, d)]$$

Dynamic Treatment and Mediation Effects

Dynamic treatment effects (type 1):

$$E[Y_2(\tilde{y}_1, \tilde{d}) - Y_2(y_1, d)] \quad \text{and} \quad E[Y_1(\tilde{d}_1) - Y_1(d_1)]$$

- dynamic complementarity: $\tilde{y}_1 = y_1$ and comparing $\tilde{d} = (0, 1)$ and $d = (0, 0)$ with $\tilde{d} = (1, 1)$ and $d = (1, 0)$:

$$E[Y_2(y_1, 0, 1) - Y_2(y_1, 0, 0)] \quad \text{vs} \quad E[Y_2(y_1, 1, 1) - Y_2(y_1, 1, 0)]$$

- state dependence: $\tilde{d} = d$ and $\tilde{y}_1 = 1$ and $y_1 = 0$:

$$E[Y_2(1, d) - Y_2(0, d)]$$

Dynamic treatment effects (type 2):

$$E[Y_2(\tilde{d}_1, d_2) - Y_2(d_1, d_2)]$$
$$E[Y_2(\tilde{d}_1) - Y_2(d_1)]$$

Dynamic Treatment and Mediation Effects

Direct effect and indirect effect mediated by Y_1:

$$
\begin{aligned}
& E[Y_2(1, d_2) - Y_2(0, d_2)] \\
& = E[Y_2(Y_1(0), 1, d_2) - Y_2(Y_1(0), 0, d_2)] \\
& + E[Y_2(Y_1(1), 1, d_2) - Y_2(Y_1(0), 1, d_2)] \\
& = E[Y_2(Y_1(1), 1, d_2) - Y_2(Y_1(1), 0, d_2)] \\
& + E[Y_2(Y_1(1), 0, d_2) - Y_2(Y_1(0), 0, d_2)]
\end{aligned}
$$

Dynamic Treatment and Mediation Effects

Direct effect and indirect effect mediated by Y_1:

$$
\begin{aligned}
& E[Y_2(1, d_2) - Y_2(0, d_2)] \\
& = E[Y_2(Y_1(0), 1, d_2) - Y_2(Y_1(0), 0, d_2)] \\
& + E[Y_2(Y_1(1), 1, d_2) - Y_2(Y_1(0), 1, d_2)] \\
& = E[Y_2(Y_1(1), 1, d_2) - Y_2(Y_1(1), 0, d_2)] \\
& + E[Y_2(Y_1(1), 0, d_2) - Y_2(Y_1(0), 0, d_2)]
\end{aligned}
$$

Again, for all above, need to identify $\mu_1(d_1)$, $\mu_2(y_1, d)$, $\Sigma(y_1, d)$

Dynamic Treatment and Mediation Effects

Direct effect and indirect effect mediated by Y_1:

$$
\begin{aligned}
& E[Y_2(1, d_2) - Y_2(0, d_2)] \\
& = E[Y_2(Y_1(0), 1, d_2) - Y_2(Y_1(0), 0, d_2)] \\
& + E[Y_2(Y_1(1), 1, d_2) - Y_2(Y_1(0), 1, d_2)] \\
& = E[Y_2(Y_1(1), 1, d_2) - Y_2(Y_1(1), 0, d_2)] \\
& + E[Y_2(Y_1(1), 0, d_2) - Y_2(Y_1(0), 0, d_2)]
\end{aligned}
$$

Again, for all above, need to identify $\mu_1(d_1)$, $\mu_2(y_1, d)$, $\Sigma(y_1, d)$

In addition, $\pi_2(y_1, d_1, z_2)$ captures habit and learning

$$E[D_2(\tilde{y}_1, \tilde{d}_1) - D_2(y_1, d_1)]$$

III. Identification Analysis

Identifying Assumptions

Assumption Z

(i) $(Z_1, Z_2) \perp (V_1, V_2, U_1(d_1), U_2(y_1, d))$ and (ii) π_1 and π_2 are non-trivial functions of Z_1 and Z_2.

Assumption S

(i) $C(v_1, u_1; \rho_{v_1 u_1}) \prec_S C(v_1, u_1; \tilde{\rho}_{v_1 u_1})$ for any $\rho_{v_1 u_1} < \tilde{\rho}_{v_1 u_1}$;
(ii) $C(v_1, v_2, u_1; \rho_{v_1 v_2}) \prec_S C(v_1, v_2, u_1; \tilde{\rho}_{v_1 v_2})$ for any $\rho_{v_1 v_2} < \tilde{\rho}_{v_1 v_2}$;
(iii) $C(v_1, v_2, u_1, u_2; \rho_{v_2 u_2}) \prec_S C(v_1, v_2, u_1, u_2; \tilde{\rho}_{v_2 u_2})$ for any $\rho_{v_2 u_2} < \tilde{\rho}_{v_2 u_2}$.

- the ordering "$\prec_S$" denote "more stochastic increasing" (cf. "more positive regression dependent")
- Han & Vytlacil 17 use S(i) for *static* triangular model
  - extension to S(ii)–(iii) is not straightforward (next slide)
  - we consider much flexible models (even if it were static) and allow for continuous outcome

Sufficient Condition for Assumption S

Assumption S*

(i) Same as S(i); (ii) The copulas are generated by

$$
\begin{aligned}
&C(v_1, v_2, u_1; \rho_{v_1 v_2}, \rho_{v_1 u_1}, \rho_{v_2 u_1}) \\
&= \int^{v_1} C\left(C(v_2|\tilde{v}_1), C(u_1|\tilde{v}_1); \rho(\rho_{v_1 v_2}, \rho_{v_1 u_1}, \rho_{v_2 u_1})\right) d\tilde{v}_1 \\
&C(v_1, v_2, u_1, u_2; \Sigma) \\
&= \int^{v_1, v_2} C\left(C(u_1|\tilde{v}_1, \tilde{v}_2), C(u_2|\tilde{v}_1, \tilde{v}_2); \rho(\Sigma)\right) dC(\tilde{v}_1, \tilde{v}_2)
\end{aligned}
$$

where the outer $C(\cdot,\cdot;\rho)$ satisfies $C(\cdot,\cdot;\rho) \prec_S C(\cdot,\cdot;\tilde{\rho})$ for $\rho < \tilde{\rho}$; (iii) $\rho(\rho_{v_1 v_2}, \rho_{v_1 u_1}, \rho_{v_2 u_1})$ and $\rho(\Sigma)$ are strictly increasing in $\rho_{v_2 u_1}$ and $\rho_{u_1 u_2}$, respectively.

- S*(ii) exploits vine copula structure
- S* holds for multivariate Gaussian copula!

Identification Analysis: Step 1

First, consider

$$Y_1 = 1[\mu_1(D_1) \geq U_1(D_1)]$$
$$D_1 = 1[\pi_1(Z_1) \geq V_1]$$

Note $\pi_1(z_1)$ is trivially ID'ed as $\pi_1(z_1) = \Pr[D_1 = 1|Z_1 = z_1]$ by normalization

Consider $\Pr[D_1 = d, Y_1 = y|Z_1 = z]$ for $(d, y, z) \in \{0, 1\}^3$:

By Assumptions Z and C,

$$\begin{aligned}
\Pr[D_1 = 1, Y_1 = 1|Z_1 = 0] &= \Pr[V_1 \leq \pi_1(0), U_1(1) \leq \mu_1(1)] \\
&= C(\pi_1(0), \mu_1(1); \rho_{V_1,U_1(1)}) \\
\Pr[D_1 = 1, Y_1 = 1|Z_1 = 1] &= \Pr[V_1 \leq \pi_1(1), U_1(1) \leq \mu_1(1)] \\
&= C(\pi_1(1), \mu_1(1); \rho_{V_1,U_1(1)})
\end{aligned}$$

Identification Analysis: Step 1

The system of nonlinear equations:

$$\Pr[D_1 = 1, Y_1 = 1 | Z_1 = 0] = C(\pi_1(0), \mu_1(1); \rho_{V_1,U_1(1)})$$
$$\Pr[D_1 = 1, Y_1 = 1 | Z_1 = 1] = C(\pi_1(1), \mu_1(1); \rho_{V_1,U_1(1)})$$

- unique solution for $(\mu_1(1), \rho_{V_1,U_1(1)})$ if its Jacobian is P-matrix (Gale & Nikaido 65)
- this is true if and only if

$$\frac{C_2(\pi_1(0), \mu_1(1))}{C_{\rho_{V_1,U_1(1)}}(\pi_1(0), \mu_1(1))} \neq \frac{C_2(\pi_1(1), \mu_1(1))}{C_{\rho_{V_1,U_1(1)}}(\pi_1(1), \mu_1(1))}$$

- which is guaranteed by Assumptions Z(ii) and S(i)

Identification Analysis: Step 1

The system of nonlinear equations:

$$\Pr[D_1 = 1, Y_1 = 1 | Z_1 = 0] = C(\pi_1(0), \mu_1(1); \rho_{V_1,U_1(1)})$$
$$\Pr[D_1 = 1, Y_1 = 1 | Z_1 = 1] = C(\pi_1(1), \mu_1(1); \rho_{V_1,U_1(1)})$$

- unique solution for $(\mu_1(1), \rho_{V_1,U_1(1)})$ if its Jacobian is P-matrix (Gale & Nikaido 65)
- this is true if and only if

$$\frac{C_2(\pi_1(0), \mu_1(1))}{C_{\rho_{V_1,U_1(1)}}(\pi_1(0), \mu_1(1))} \neq \frac{C_2(\pi_1(1), \mu_1(1))}{C_{\rho_{V_1,U_1(1)}}(\pi_1(1), \mu_1(1))}$$

- which is guaranteed by Assumptions Z(ii) and S(i)

Similarly for $(\mu_1(0), \rho_{V_1,U_1(0)})$ with

$$\Pr[D_1 = 0, Y_1 = 1 | Z_1 = 0] = \mu_1(0) - C(\pi_1(0), \mu_1(0); \rho_{V_1,U_1(0)})$$
$$\Pr[D_1 = 0, Y_1 = 1 | Z_1 = 1] = \mu_1(0) - C(\pi_1(1), \mu_1(0); \rho_{V_1,U_1(0)})$$

Identification Analysis: Step 2

ID'ed in Step 1: $\pi_1(z_1), \mu_1(d_1), \rho_{V_1,U_1(d_1)}$

Next, consider

$$\begin{aligned} D_2 &= 1[\pi_2(Y_1, D_1, Z_2) \geq V_2] \\ Y_1 &= 1[\mu_1(D_1) \geq U_1(D_1)] \\ D_1 &= 1[\pi_1(Z_1) \geq V_1] \end{aligned}$$

Fix z_2, and for each $z_1 \in \{0, 1\}$,

$$\begin{aligned} &\Pr[D_1 = 1, D_2 = 1, Y_1 = 1 | Z_1 = z_1, Z_2 = z_2] \\ =&\Pr[V_1 \leq \pi_1(z_1), V_2 \leq \pi_2(1, 1, z_2), U_1(1) \leq \mu_1(1)] \\ =&C(\pi_1(z_1), \pi_2(1, 1, z_2), \mu_1(1); \rho_{V_1,V_2}, \rho_{V_1,U_1(1)}) \end{aligned}$$

Relevant Jacobian is P-matrix by Assumptions Z(ii) and S(ii) as

$$\frac{C_2(\pi_1(0), \pi_2(1, 1, z_2), \mu_1(1))}{C_{\rho_{V_1,V_2}}(\pi_1(0), \pi_2(1, 1, z_2), \mu_1(1))} \neq \frac{C_2(\pi_1(1), \pi_2(1, 1, z_2), \mu_1(1))}{C_{\rho_{V_1,V_2}}(\pi_1(1), \pi_2(1, 1, z_2), \mu_1(1))}$$

Identification Analysis: Step 3

ID'ed in Steps 1, 2: $\pi_1(z_1), \mu_1(d_1), \rho_{V_1,U_1(d_1)}, \pi_2(y_1, d_1, z_2), \rho_{V_1,V_2}$

Finally, consider

$$\begin{aligned} Y_2 &= 1[\mu_2(Y_1, D) \geq U_2(Y_1, D)] \\ D_2 &= 1[\pi_2(Y_1, D_1, Z_2) \geq V_2] \\ Y_1 &= 1[\mu_1(D_1) \geq U_1(D_1)] \\ D_1 &= 1[\pi_1(Z_1) \geq V_1] \end{aligned}$$

Fix z_1, and for each $z_2 \in \{0, 1\}$,

$$\begin{aligned} &\Pr[D_1 = 1, D_2 = 1, Y_1 = 1, Y_2 = 1 | Z_1 = z_1, Z_2 = z_2] \\ =&\Pr[V_1 \leq \pi_1(z_1), V_2 \leq \pi_2(1, 1, z_2), \ldots \\ &\qquad U_1(1) \leq \mu_1(1), U_2(1, 1, 1) \leq \mu_2(1, 1, 1)] \\ =&C(\pi_1(z_1), \pi_2(1, 1, z_2), \mu_1(1), \mu_2(1, 1, 1); \rho_{V_1,V_2}, \rho_{V_1,U_1(1)}, \rho_{V_2,U_2(1,1,1)}) \end{aligned}$$

Relevant Jacobian is P-matrix by Assumptions Z(ii) and S(iii)

Identification Analysis: Summary and Extensions

Regardless of the length of T and complexity with $U_t(y^{t-1}, d^t)$, each step involves...

- pairs of probabilities that produces...
- 2×2 Jacobian matrices and
- relevant stochastic ordering of mutlivariate copula,
- which is guaranteed by the vine copula assumptions

Identification Analysis: Summary and Extensions

Regardless of the length of T and complexity with $U_t(y^{t-1}, d^t)$, each step involves...

- pairs of probabilities that produces...
- 2×2 Jacobian matrices and
- relevant stochastic ordering of mutlivariate copula,
- which is guaranteed by the vine copula assumptions

With continuous Y_t, we directly ID the distribution of counterfactual outcomes in place of $\mu_t(y^{t-1}, d^t)$

- then we also identify dynamic *quantile* treatment and mediation effects (of all kinds defined earlier)

IV. Discussions and Conclusions

Estimation

With continuous X, we propose sieve ML to jointly estimate

$$\pi_1(z_1, x), \mu_1(d_1, x), \rho_{V_1, U_1(d_1)}(x),$$
$$\pi_2(y_1, d_1, z_2, x), \rho_{V_1, V_2}(x),$$
$$\mu_2(y_1, d, x), \rho_{V_2, U_2(y_1, d)}(x)$$

as functions of $x \in \mathcal{X}$

- show consistency of sieve MLE and asymptotic normality of its functionals
- for inference, provide theory for sieve LR test

Estimation

With continuous X, we propose sieve ML to jointly estimate

$$\pi_1(z_1, x), \mu_1(d_1, x), \rho_{V_1, U_1(d_1)}(x),$$
$$\pi_2(y_1, d_1, z_2, x), \rho_{V_1, V_2}(x),$$
$$\mu_2(y_1, d, x), \rho_{V_2, U_2(y_1, d)}(x)$$

as functions of $x \in \mathcal{X}$

- show consistency of sieve MLE and asymptotic normality of its functionals
- for inference, provide theory for sieve LR test

With discrete X or without X, we have parametric ML!

- a semiparametric but saturated model for treatment effects
- not true if the focus is underlying parameters with linear index (Han & Lee 19)

Two Related Works

1. Local Gaussian representation in static model (Chernozhukov, Fernandez-Val, Han & Wüthrich 23)

- copula as a representation (instead of restriction)
- dependence parameter as an implicit function
- no longer a representation with multivariate copula

$\longrightarrow$ not applicable to dynamic settings

Two Related Works

1. Local Gaussian representation in static model (Chernozhukov, Fernandez-Val, Han & Wüthrich 23)

- copula as a representation (instead of restriction)
- dependence parameter as an implicit function
- no longer a representation with multivariate copula

  $\longrightarrow$ not applicable to dynamic settings

2. Empirical study of pre- and post-natal maternal smoking on child development (Cattan, Conti, Han & Salvati 23)

- specific structure of dynamics
- mainly multivariate probit models with linear index
- besides treatment effects, we also study habits and change in endogenous selection

Concluding Remarks

Propose a practically useful framework to study dynamic treatment and mediation effects

- a class of semiparametric model
- point ID of a large array of parameters
- sieve estimation and inference procedures
- possibility of restoring parametric ML

Concluding Remarks

Propose a practically useful framework to study dynamic treatment and mediation effects

- a class of semiparametric model
- point ID of a large array of parameters
- sieve estimation and inference procedures
- possibility of restoring parametric ML

Follow-up questions:

- endogenous drop-out?
- formal theory for extrapolation of (dynamic) fuzzy RD?

Thank You! ☺