Identification and Estimation of Nonstationary Dynamic Discrete Choice Models

Cheng Chou
Independent

Geert Ridder
USC

Ruoyao Shi
UC Riverside

UC Berkeley
May 1, 2025

Motivation

- Structural DDC models can be very useful:
  - Inter-temporal preference of forward-looking agents.
  - Counterfactual analysis.
  - Examples: labor force participation, demand for durable goods, etc.
- Estimation:
  - Full solution estimators (MLE).
  - CCP-based estimators (simulated GMM, pseudo-MLE).
  - Finite dependence (GMM).

  All require estimation of state transition distribution as input, some even need simulation (HMSS, 1994) or iteration (Aguirregabiria & Mira, 2002), to evaluate the objective function.
- **Practical motivation**: a novel & simple estimator of flow utility parameters that completely bypasses state transition.

Preview of Results

- We articulate a **Markovian** property for observed state variables under coomon assumptions in DDC literature.
- It **eliminates** the need of dealing with the state transitions.
- Represent the optimal decision rule as a simple **linear** system under mild additional assumptions.
- **Identification** conditions for flow utility parameters can be easily discussed.
- Develop a CCP-based semiparametric **estimator** that is much simpler and faster than alternatives ($\approx$ 1000 times faster than HMSS in MC).

Relation to Literature

- **Estimation**: Rust (1987, NFP), Hotz & Miller (1993), Hotz, Miller, Sanders & Smith (1994, HMSS), Aguirregabiria & Mira (2002, NPL), Pesendorfer & Schmidt-Dengler (2008, GMM), Aguirregabiria & Magesan (2013, Euler), Arcidiacono & Miller (2019, FDP), Kalouptsidi, Scott & Souza-Rodrigues (2021), Chiong, Galichon & Shum (2016, duality), Srisuma & Linton (2012, FIE), Buchholz, Shum & Xu (2021, FIE); Adusumilli & Eckardt (2019).
- **Identification**: Rust (1987, 1994), Magnac & Thesmar (2002), Hotz & Miller (1993), Arcidiacono & Miller (2011), Bajari, Chu, Nekipelov & Park (2016), Arcidiacono & Miller (2020), Abbring & Daljord (2020).
- **Counterfactual**: Aguirregabiria (2010), Arcidiacono & Miller (2020), Kalouptsidi et al. (2021); **unobserved heterogeneity**: Kasahara & Shimotsu (2009), Arcidiacono & Miller (2011), Hu & Shum (2012), Higgins & Jochmans (2023); **partial identification**: Norets & Tang (2014), Berry & Compiani (2023), Kalouptsidi, Kitamura et al.

Nonstationary Dynamic Binary Choice Model

- Decision horizon: $\mathcal{T} \equiv \{T_{start}, \ldots, T_{end}\}$.
- In each period $t \in \mathcal{T}$, flow utility is $u(a_t, s_t; \delta_t) + \varepsilon_{a_t t}$, where
  - $a_t \in \{0, 1\}$: a binary choice;
  - s_t: observed state variables;
  - $\varepsilon_t \equiv (\varepsilon_{1t}, \varepsilon_{0t})'$ utility shocks; let $\Omega_t = (s_t', \varepsilon_t')'$.
- Each agent chooses a_t to maximize expected lifetime payoff:

$$\mathbb{E}\left(\sum_{j=0}^{T_{end}-t} \beta^j \left(u(a_{t+j}, x_{t+j}; \delta_{t+j}) + \varepsilon_{a_{t+j}t+j}\right) \Big| s_t, \varepsilon_t, a_t \right).$$

- Model primitives:
  - flow utility parameters: δ_t for all $t \in \mathcal{T}$;
  - state transition distributions: $f_{t+1}(s_{t+1}, \varepsilon_{t+1} | s_t, \varepsilon_t, a_t)$;
  - shock distributions: $F_t(\varepsilon_t)$.
- Nonstationarity: δ_t or f_t time-varying; T_{end} finite.

We Maintain These Common Assumptions

Assumption 1 (Controlled Markov process)

Assume $(s_{t+1}, \varepsilon_{t+1}) \perp\!\!\!\perp (s_{t-j}, \varepsilon_{t-j}, a_{t-j}) \mid (s_t, \varepsilon_t, a_t)$ *for* $j \in \mathbb{N}^+$.

Assumption 2 (Flow utility shocks)

Assume (i) $\varepsilon_t \perp\!\!\!\perp s_t$; *(ii)* $\varepsilon_t \perp\!\!\!\perp s_{t-1}$; *(iii)* $\varepsilon_t \perp\!\!\!\perp \varepsilon_{t'}$ *for* $t \neq t'$; *(iv)* $\varepsilon_{0t} \perp\!\!\!\perp \varepsilon_{1t}$.

Assumption 3 (Conditional independence)

Assume $s_{t+1} \perp\!\!\!\perp \varepsilon_t \mid (s_t, a_t)$.

- A subset of AM (2010) assumptions; sufficient for Markovian below.
- Common "conditional independence" assumption in literature:

$$\begin{aligned} f(s_{t+1}, \varepsilon_{t+1} | s_t, \varepsilon_t, a_t) &= f(\varepsilon_{t+1} | s_{t+1}, s_t, \varepsilon_t, a_t) \cdot f(s_{t+1} | s_t, \varepsilon_t, a_t) \\ &= f(\varepsilon_{t+1}) \cdot f(s_{t+1} | s_t, a_t) \end{aligned}$$

Some Standard Preliminary Results (HM 1993, AM 2011)

- Conditional choice probability (CCP):

$$p_t(s) \equiv Pr.(a_t = 1 | s_t = s) = \int a_t^o(s_t, \varepsilon_t) dF_t(\varepsilon_t).$$

- Choice-specific conditional value function $v_{at}(s_t)$ and integrated value function $\bar{V}_t(\cdot)$ have the relation ($a \in \{0, 1\}$):

$$v_{at}(s_t) \equiv u_t(a, s_t) + \beta \mathbb{E}(\bar{V}_{t+1}(s_{t+1}) | s_t, a_t = a).$$

- For each $a \in \{0, 1\}$, there exists a function ψ_a of CCP only, whose functional form is known given $F_t(\varepsilon_t)$, such that

$$\psi_a(p_t) = \bar{V}_t(s_t) - v_{at}(s_t);$$

intuitively, it adjusts for the fact that choice a might not be optimal.

Some Standard Preliminary Results (cont'd)

- Our approach starts with an implication of these results:

$$\psi_0(p_t)-\psi_1(p_t) = \underbrace{u(1,s_t;\delta_t)-u(0,s_t;\delta_t)}_{\text{static logit}} + \underbrace{\beta\Delta\mathbb{E}\left(\bar{V}_{t+1}(s_{t+1}) \mid s_t\right)}_{\text{main difficulty in dynamic models}}.$$

  where $\Delta\mathbb{E}(h_\tau|s_t) \equiv \mathbb{E}(h_\tau|s_t, a_t = 1) - \mathbb{E}(h_\tau|s_t, a_t = 0)$ for $\tau > t$.

- Repeatedly plug in the Bellman's equation (U_t^o is optimal flow utility):

$$\bar{V}_t(s_t) = U_t^o(s_t) + \beta\mathbb{E}(\bar{V}_{t+1}(s_{t+1})|s_t).$$

- Stop in some period $T^* \leq T_{end}$, we get

$$\begin{aligned} &\psi_0(p_t) - \psi_1(p_t) = u(1,x_t;\delta_t) - u(0,x_t;\delta_t) \\ &+ \sum_{\tau=t+1}^{T^*-1} \beta^{\tau-t}\Delta\mathbb{E}(\mathbb{E}(\cdots\mathbb{E}(U_\tau^o(s_\tau)|s_{\tau-1})\cdots|s_{t+1})|s_t) \\ &+ \beta^{T^*-t}\Delta\mathbb{E}(\mathbb{E}(\cdots\mathbb{E}(\mathbb{E}(\bar{V}_{T^*}(s_{T^*})|s_{T^*-1})|s_{T^*-2})\cdots|s_{t+1}|s_t). \quad (1) \end{aligned}$$

Markovian Property

Theorem 1 (Markovian s_t)

Under Assumptions 1 & 2(i)-(iii), s_t is a first order Markov process; that is, $s_{t+1} \perp\!\!\!\perp s_{t-j} \mid s_t$ for $j \in \mathbb{N}^+$.

- Proof only uses elementary probability theory. proof
- Not implied by Assumption 1 alone, or vice versa. assumptions
- This is the key to our approach, not exploited in literature.

Lemma 2 (Conditional independence)

Under Assumptions 1 & 2(i)-(iii), for $j \in \mathbb{N}^+$ and measurable function $g(\cdot)$,

$$\mathbb{E}(\mathbb{E}(g(s_{t+j}) \mid s_{t+1}) \mid s_t, a_t) = \mathbb{E}(g(s_{t+j}) \mid s_t, a_t).$$

Markovian Property (cont'd)

Theorem 2 (Telescoping)

Under Assumptions 1-3, eq. (1) simplifies to

$$\psi_0(p_t) - \psi_1(p_t) = u(1, x_t; \delta_t) - u(0, x_t; \delta_t)$$
$$+ \sum_{\tau=t+1}^{T^*-1} \beta^{\tau-t} \Delta\mathbb{E}(\mathbb{E}(\cdots \mathbb{E}(\mathbb{E}(U_\tau^o(s_\tau) \mid s_{\tau-1}) \mid s_{\tau-2}) \cdots \mid s_{t+1}) \mid s_t)$$
$$+ \beta^{T^*-t} \Delta\mathbb{E}(\mathbb{E}(\cdots \mathbb{E}(\mathbb{E}(\bar{V}_{T^*}(s_{T^*}) \mid s_{T^*-1}) \mid s_{T^*-2}) \cdots \mid s_{t+1}) \mid s_t)$$
$$= u(1, x_t; \delta_t) - u(0, x_t; \delta_t)$$
$$+ \sum_{\tau=t+1}^{T^*-1} \beta^{\tau-t} \Delta\mathbb{E}(U_\tau^o(s_\tau) \mid s_t) + \beta^{T^*-t} \Delta\mathbb{E}(\bar{V}_{T^*}(s_{T^*}) \mid s_t). \quad (2)$$

Mild Common Assumptions

Assumption 4 (Logit)

ε_{0t} and ε_{1t} both follow a type I extreme value distribution.

- Implies: $\psi_0(p_t) - \psi_1(p_t) = \ln(p_t/(1-p_t))$.

Assumption 5 (Linear flow utility)

For each $t \in \mathcal{T}$, suppose $u(0, x_t; \delta_t) = x_t' \delta_{0,t}$ and $u(1, x_t; \delta_t) = x_t' \delta_{1,t}$ for some $\delta_{0,t}$ and $\delta_{1,t}$. We normalize $\delta_{0,1} = c_{d_x \times 1}$.

- Weaker than what's common in DDC literature.
- $s_t = (x_t', z_t')'$, where z_t (can be empty) is "excluded variables".
- Implies:

$$U_t^o(s_t) = p_t x_t' \delta_{1,t} + (1-p_t) x_t' \delta_{0,t} + \gamma - p_t \ln(p_t) - (1-p_t) \ln(1-p_t).$$

Mild Common Assumptions (cont'd)

- Under Assumptions 1-5, eq. (2) becomes a partially linear system:

$$y_{T^*-1} = x'_{T^*-1}\Delta_{T^*-1} + \beta\Delta\mathbb{E}(\bar{V}_{T^*}(x_{T^*}, z_{T^*})|s_{T^*-1}), \text{ and} \tag{3a}$$

$$y_t = x'_t\Delta_t + \sum_{\tau=t+1}^{T^*-1}\beta^{\tau-t}\Delta\bar{x}_t^{\tau\prime}\delta_{0,\tau} + \sum_{\tau=t+1}^{T^*-1}\beta^{\tau-t}\Delta\bar{x}_{1,t}^{\tau\prime}\Delta_\tau + \beta^{T^*-t}\Delta\mathbb{E}(\bar{V}_{T^*}(s_{T^*})|s_t), \text{ for } t = 1, \ldots, T-2, \tag{3b}$$

where

- We let $\Delta_t \equiv \delta_{1,t} - \delta_{0,t}$.
- $\Delta\bar{x}_{1,t}^{\tau} \equiv \Delta\mathbb{E}(p_\tau x_\tau|x_t, z_t)$ and $\Delta\bar{x}_t^{\tau} \equiv \Delta\mathbb{E}(x_\tau|x_t, z_t)$;
- $y_{T-1} \equiv \ln\left(\frac{p_{T-1}}{1-p_{T-1}}\right)$ and $y_t \equiv \ln\left(\frac{p_t}{1-p_t}\right) + \sum_{\tau=t+1}^{T-1}\beta^{\tau-t}\Delta\bar{\eta}_t^{\tau}$;
- $\Delta\bar{\eta}_t^{\tau} \equiv \Delta\mathbb{E}(\eta_\tau|x_t, z_t)$, $\eta_\tau \equiv p_\tau\ln(p_\tau) + (1-p_\tau)\ln(1-p_\tau)$ for $\tau > t$.

- **Key**: $\Delta\bar{x}_{1,t}^{\tau}$, $\Delta\bar{x}_t^{\tau}$ and $\Delta\bar{\eta}_t^{\tau}$ are conditional mean differences of h_τ involving future CCP ($\tau > t$).

A Mild New Assumption

Assumption 6 (Sample-terminal-period integrated value function)

Suppose there exist are K known functions of (x, z), denoted by $q^K(x, z) \equiv (q^{K,1}(x, z), \dots, q^{K,K}(x, z))'$, and a $K \times 1$ unknown vector of parameters γ^K, such that $\bar{V}_T(x_T, z_T) = q^K(x_T, z_T)'\gamma^K$.

- If $T = T_{end}$ and known, then $q^K(x_T, z_T) = (x_T', p_T x_T' \Delta_T - \eta_T)'$.
- Let $T^* = T$, eq. (3) further becomes a linear system:

$$y_{T-1} = x_{T-1}' \Delta_{T-1} + \beta \Delta \bar{q}_{T-1}^K \gamma^K, \text{ and} \tag{4a}$$

$$y_t = x_t' \Delta_t + \sum_{\tau=t+1}^{T-1} \beta^{\tau-t} \Delta \bar{x}_t^{\tau\prime} \delta_{0,\tau} + \sum_{\tau=t+1}^{T-1} \beta^{\tau-t} \Delta \bar{x}_{1,t}^{\tau\prime} \Delta_\tau + \beta^{T-t} \Delta \bar{q}_t^K \gamma^K, \text{ for } t = 1, \dots, T-2, \tag{4b}$$

where $\Delta \bar{q}_t^K \equiv \Delta \mathbb{E}(q^K(x_T, z_T) | x_t, z_t)$. rank condition

3-Step CCP-Based Semiparametric Estimator

1. **CCP**: use data $\{a_{i,t}, s_{i,t}\}_{i=1}^N$ to obtain CCP estimates $\hat{p}_{i,t}$.
2. **Conditional mean differences**: obtain $\hat{h}_{i,\tau}$ by substituting unknown $p_{i,\tau}$ with $\hat{p}_{i,\tau}$, then use data $\{\hat{h}_{i,\tau}, a_{i,t}, s_{i,t}\}_{i=1}^N$ to obtain the conditional mean difference estimates of $\Delta\bar{\eta}$, $\Delta\bar{x}$ and $\Delta\bar{q}^K$.
3. **Parameter of interest**: let $\bar{m}_N(\delta,\gamma^K) \equiv \frac{1}{N}\sum_{i=1}^N m(a_i, s_i, \delta, \gamma^K, \hat{p}, \widehat{\Delta\bar{\eta}}, \widehat{\Delta\bar{x}}, \widehat{\Delta\bar{q}^K})$ capture the distance between LHS & RHS of eq. (4), and obtain **closed-form solution** to:

$$(\hat{\delta}', \hat{\gamma}^{K\prime})' \equiv \underset{\delta\in\mathbb{R}^{(2T-3)d_x},\gamma_K\in\mathbb{R}^K}{\arg\min} \bar{m}_N(\delta,\gamma^K)' W_N \bar{m}_N(\delta,\gamma^K).$$

asymptotic distribution

Simulation Setup

- Fix a context: a car dealership chooses whether to begin offering EV over a three-year horizon.
  - x_{1t} & x_{2t}: dealership's readiness (service equipment, training, etc.);
  - w_{1t} & w_{2t}: dealership's latent internal investment, determine x_{1t} & x_{2t};
  - z_t: public sentiment.
- $T = T_{end} = 3$, but researcher may or may not know this fact.
- Researcher observes a_t and $s_t = (x_{1t}, x_{2t}, z_t)'$.
- (w_{1t}, w_{2t}, z_t) follows choice-specific VAR(1). VAR(1) details
- Flow utility function is

$$u_t(a, x_t) = \delta_{a,t,0} + \delta_{a,t,1} x_{t,1} + \delta_{a,t,2} x_{t,2}.$$

- Time-invariant (unknown to researcher) **parameters of interest**:
  - $(\delta_{0,t,0}, \delta_{0,t,1}, \delta_{0,t,2}) = (\underline{0}, \underline{-0.5}, \underline{-0.3})$;
  - $(\delta_{1,t,0}, \delta_{1,t,1}, \delta_{1,t,2}) = (-0.5, 0.1, 0.1)$.

Results: $T = T_{end} = 3$ Known ($N = 250$, $R = 1000$)

	$\delta_{a,t,k}$	Truth	CRS estimator			HMSS estimator (NP)			HMSS estimator (P)		
			Abs Bias	Std Dev	MSE	Abs Bias	Std Dev	MSE	Abs Bias	Std Dev	MSE
$t = 1$	$\delta_{1,1,0}$	−0.5	0.081	0.261	0.075	0.047	0.269	0.075	0.116	0.255	0.078
	$\delta_{1,1,1}$	0.1	0.003	0.045	0.002	0.028	0.044	0.003	0.029	0.047	0.003
	$\delta_{1,1,2}$	0.1	0.004	0.038	0.001	0.043	0.042	0.004	0.064	0.043	0.006
$t = 2$	$\delta_{0,2,1}$	−0.5	0.042	0.287	0.084	0.261	0.174	0.098	0.128	0.191	0.053
	$\delta_{0,2,2}$	−0.3	0.030	0.142	0.021	0.141	0.120	0.034	0.287	0.142	0.103
	$\delta_{1,2,0}$	−0.5	0.037	0.165	0.029	0.029	0.199	0.040	0.065	0.193	0.041
	$\delta_{1,2,1}$	0.1	0.005	0.042	0.002	0.012	0.042	0.002	0.009	0.047	0.002
	$\delta_{1,2,2}$	0.1	0.013	0.040	0.002	0.018	0.039	0.002	0.020	0.045	0.002
$t = 3$	$\delta_{0,3,2}$	−0.5	0.071	0.213	0.050	0.303	0.115	0.105	0.111	0.175	0.043
	$\delta_{0,3,3}$	−0.3	0.076	0.125	0.021	0.201	0.088	0.048	0.171	0.117	0.043
	$\delta_{1,3,1}$	−0.5	0.009	0.125	0.016	0.009	0.191	0.037	0.011	0.186	0.035
	$\delta_{1,3,2}$	0.1	0.001	0.034	0.001	0.000	0.045	0.002	0.004	0.045	0.002
	$\delta_{1,3,3}$	0.1	0.009	0.035	0.001	0.004	0.046	0.002	0.005	0.048	0.002

Results: $T = T_{end}$ Unknown ($N = 250$, $R = 1000$)

	$\delta_{a,t,k}$	Truth	CRS estimator			HMSS estimator (NP)			HMSS estimator (P)		
			Abs Bias	Std Dev	MSE	Abs Bias	Std Dev	MSE	Abs Bias	Std Dev	MSE
$t=1$	$\delta_{1,1,0}$	−0.5	0.031	0.311	0.098	0.049	0.273	0.077	0.112	0.274	0.088
	$\delta_{1,1,1}$	0.1	0.004	0.053	0.003	0.028	0.045	0.003	0.030	0.051	0.004
	$\delta_{1,1,2}$	0.1	0.008	0.048	0.002	0.045	0.042	0.004	0.074	0.045	0.008
$t=2$	$\delta_{0,2,1}$	−0.5	0.090	0.301	0.099	0.196	0.168	0.067	0.050	0.162	0.029
	$\delta_{0,2,2}$	−0.3	0.044	0.167	0.030	0.120	0.117	0.028	0.267	0.144	0.092
	$\delta_{1,2,0}$	−0.5	0.012	0.215	0.046	0.062	0.185	0.038	0.068	0.179	0.037
	$\delta_{1,2,1}$	0.1	0.010	0.053	0.003	0.019	0.037	0.002	0.013	0.035	0.001
	$\delta_{1,2,2}$	0.1	0.002	0.053	0.003	0.022	0.035	0.002	0.024	0.035	0.002
$t=3$	$\delta_{0,3,2}$	−0.5	–	–	–	0.438	0.033	0.193	0.429	0.032	0.185
	$\delta_{0,3,3}$	−0.3	–	–	–	0.245	0.037	0.061	0.249	0.031	0.063
	$\delta_{1,3,1}$	−0.5	–	–	–	0.053	0.184	0.037	0.034	0.181	0.034
	$\delta_{1,3,2}$	0.1	–	–	–	0.008	0.044	0.002	0.010	0.043	0.002
	$\delta_{1,3,3}$	0.1	–	–	–	0.000	0.045	0.002	0.001	0.046	0.002

Rank Condition

- Identification of (δ, γ^K): rank condition of the linear system.
- A sufficient condition: $(x_t', \Delta\bar{x}_t^{t+1\prime})'$ not perfectly linearly correlated (recall $\Delta\bar{x}_t^{t+1} \equiv \Delta\mathbb{E}(x_{t+1}|x_t, z_t)$).
- More discussion in paper:
    - **Excluded variable** z_t is auxiliary: not required, but useful in breaking perfect linear correlation between x_t and $\Delta\bar{x}_t^{t+1}$. details
    - Triangularity helps deal with **time-invariant** variables in x_t. details
    - Unknown **discount factor** easily accommodated.
    - **Stationary model** is a special case.
    - When interpreting Assumption 6 as an approximation, bias can be quantified. details
    - Over-identification reduces reliance on Assumption 6. details

Conclusion & Future Research

- More simulations & counterfactuals.
- Unobserved heterogeneity (Chou, Liu & Shi, 2025).
- Identification of δ in partially linear system ($K \to \infty$).
- Complement other estimators with Markovian property.

Thank you!

Proof of Theorem 1

1. $\Omega_{t+1} \perp\!\!\!\perp \Omega_{t-j} | \Omega_t$ for $\Omega_t = (s_t', \varepsilon_t')'$, because by Assumption 1,

$$\begin{aligned} f(\Omega_{t+1}|\Omega_t, \Omega_{t-j}) &= \sum_{a_t=0,1} f(\Omega_{t+1}|a_t, \Omega_t, \Omega_{t-j}) Pr.(a_t|\Omega_t, \Omega_{t-j}) \\ &= \sum_{a_t=0,1} f(\Omega_{t+1}|a_t, \Omega_t) Pr.(a_t|\Omega_t) \\ &= f(\Omega_{t+1}|\Omega_t), \end{aligned}$$

   so $f(s_{t+1}, s_{t-j}|s_t, \varepsilon_t) = f(s_{t+1}|s_t, \varepsilon_t) f(s_{t-j}|s_t, \varepsilon_t)$.

2. Assumptions 1 & 2(i)-(iii) imply $\varepsilon_t \perp\!\!\!\perp (s_{t-j}, a_{t-j})|s_t$ implying

$$f(s_{t+1}, s_{t-j}|s_t, \varepsilon_t) = f(s_{t+1}|s_t, \varepsilon_t) f(s_{t-j}|s_t).$$

3. Integrate both sides w.r.t. $F(\varepsilon_t|s_t)$,

$$f(s_{t+1}, s_{t-j}|s_t) = f(s_{t+1}|s_t) f(s_{t-j}|s_t).$$

Back to Theorem 1

Asymptotic Distribution

Proposition 1 (Asymptotic distribution of $\hat{\delta}$)

$$\sqrt{N}\left(\hat{\delta} - \delta\right) \xrightarrow{d.} \mathcal{N}(0, V),$$

where $V \equiv \mathbb{E}[\psi_\delta(a_i, s_i)\psi'_\delta(a_i, x_i)]$ *and* $\psi_\delta(a, s)$ *is the influence function of* $\hat{\delta}$*, given in the paper.*

Proposition 2 (Consistent estimator of V)

$\widehat{V} \xrightarrow{p.} V$ *for the* $\widehat{V}$ *given in the paper.*

back to estimator

Simulation Setup Details

- $(w_{1t}, w_{2t}, z_t)'$ follows time-invariant choice-specific VAR(1):

$$(w_{t,1}, w_{t,2}, z_t)' = c + A(w_{t-1,1}, w_{t-1,2}, z_{t-1})' + \iota_t,$$

where

$$c' = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} 0.7 & 0.2 & 0.5 a_t z_t \\ 0.2 & 0.6 & 0.5 a_t z_t^2 \\ 0 & 0 & 0.5 \end{bmatrix}.$$

- Latent investment & public sentiment determine readiness:

$$x_{1t} = w_{1t} + z_t,$$
$$x_{2t} = w_{2t} + \left(z_t^2 - \mathbb{E}(z_t^2)\right).$$

back to simulation setup

Function of the Excluded Variable(s) z_t

- **Takeaway 1**: rank condition holds if: (a) x_t affects mean of x_{t+1} in different ways $a_t = 1$ vs. $a_t = 0$; and (b) this difference is nonlinear in x_t (recall $\Delta\mathbb{E}(x_{t+1}|x_t) \equiv \mathbb{E}(x_{t+1}|x_t, a_t = 1) - \mathbb{E}(x_{t+1}|x_t, a_t = 0)$).
- **Takeaway 2**: rank condition holds if: (a) z_t affects mean of x_{t+1} in different ways $a_t = 1$ vs. $a_t = 0$; and (b) this difference is nonlinear in x_t. (d_z not important.)
  - Least favorable case: $\Delta\mathbb{E}(x_{t+1}|x_t, z_t)$ contains an additive component that is linear in x_t (denoted as $\rho_1 x_t$).
  - If there exist $\ell(\cdot) \equiv (\ell_1(\cdot), \ldots, \ell_{d_x}(\cdot))$ and ρ_2 such that

$$\underbrace{\begin{bmatrix} x_t \\ \Delta\mathbb{E}(x_{t+1}|x_t, z_t) \end{bmatrix}}_{2d_x \times 1} = \underbrace{\begin{bmatrix} I_{d_x} & 0_{d_x \times d_z} \\ \rho_1 & \rho_2 \end{bmatrix}}_{2d_x \times 2d_x} \underbrace{\begin{bmatrix} x_t \\ \ell(z_t) \end{bmatrix}}_{2d_x \times 1}.$$

Rank condition holds if: (a) $(x_t', \ell(z_t)')'$ has invertible second moment matrix; and (b) ρ_2 has full rank (d_x). back to rank condition

Identification with Time-invariant Variables in x_t

- **Takeaway**: corresponding coordinate in $\delta_{0,t}$ is unidentified, but Δ_t is, and the counterfactual* is unaffected.
- Suppose Δ_{T-1} and γ^K are identified.
- Suppose $x_{t,1}$ is time-invariant, so $\Delta\bar{x}^{\tau}_{1,t} = 0$ for $\tau > t$.
- Consider eq. (4b) for $t = T-2$ and $t = T-3$:

$$y_{T-2} - \beta\Delta\bar{x}^{T-1\prime}_{1,T-2}\Delta_{T-1} - \beta^2\Delta\bar{q}^{K\prime}_{T-2}\gamma^K$$

$$= x'_{T-2}\Delta_{T-2} + \beta\underbrace{\Delta\bar{x}^{T-1\prime}_{T-2}}_{=0}\delta_{0,T-1},$$

$$y_{T-3} - \beta\Delta\bar{x}^{T-2\prime}_{1,T-3}\Delta_{T-2} - \beta^2\Delta\bar{x}^{T-1\prime}_{1,T-3}\Delta_{T-1} - \beta^2\underbrace{\Delta\bar{x}^{T-1\prime}_{T-3}}_{=0}\delta_{0,T-1} - \beta^3\Delta$$

$$= x'_{T-3}\Delta_{T-3} + \beta\underbrace{\Delta\bar{x}^{T-2\prime}_{T-3}}_{=0}\delta_{0,T-2}.$$

- Continually & intermittently time-varying x_t both work.

back to rank condition

Bias Induced by Assumption 6

- Assumption 6 can be interpreted as using a series basis functions $q^K(x_T, z_T)$ to **approximate** the expected value function $\bar{V}_T(x_T, z_T)$.
- Define the approximation error

$$r^K(x_T, z_T) \equiv \bar{V}_T(x_T, z_T) - q^K(x_T, z_T)'\gamma^K,$$

  and let $\Delta\bar{r}_t^K \equiv \Delta\mathbb{E}(r^K(x_T, z_T)|x_t, z_t)$ for $t = 1, \ldots, T-1$.
- Assume $\bar{V}_T(x_T, z_T)$ is m times continuously differentiable, then the approximation error using **power series** has the order $\mathbb{E}[(\Delta\bar{r}_t^K)^2] = O\left(K^{-\frac{2m}{d_s}}\right)$. This leads to:

Theorem 3 (Asymptotic bias bound of $\hat{\delta}$, power series)

$$\|\delta^K_{pseudo} - \delta\| = O\left(K^{-\frac{m}{d_s}}\right).$$

back to rank condition

Over-identification May Reduce Reliance on Assumption 6

- If T is large, recall eq. (3b) for $t = 1$:

$$y_1 = x_1' \Delta_1 + \sum_{\tau=2}^{T-1} \beta^{\tau-1} \Delta \bar{x}_1^{\tau\prime} \delta_{0,\tau} + \sum_{\tau=2}^{T-1} \beta^{\tau-1} \Delta \bar{x}_{1,1}^{\tau\prime} \Delta_\tau + \beta^{T-1} \Delta \mathbb{E}(\bar{V}_T(x_T, z_T) | s_1),$$

  which contains **all** parameters of interest.
- Also true for eq. (3b) for $t = 2$, which contains **most** parameters of interest.
- So, using eq. (3b) for the first few periods reduces reliance on Assumption 6.
- Similar to the truncation employed by the HMSS estimators.

back to rank condition