Dynamic Random Utility

Mira Frick (Yale)
Ryota Iijima (Yale)
Tomasz Strzalecki (Harvard)

Static Random Utility

- Agent is maximizing utility subject to private information
  - randomness ("utility shocks") at individual level
  - population heterogeneity

- Analyst observes agent: choices appear stochastic because analyst does not have access to this private information
  - for each menu of options, the analyst observes a probability distribution of choices (a stochastic choice rule ρ)

Choice probability:

$$\rho(x, A) = \mathbb{P}\Big(U(x) = \max_{y \in A} U(y)\Big)$$

Dynamic Random Utility (DRU)

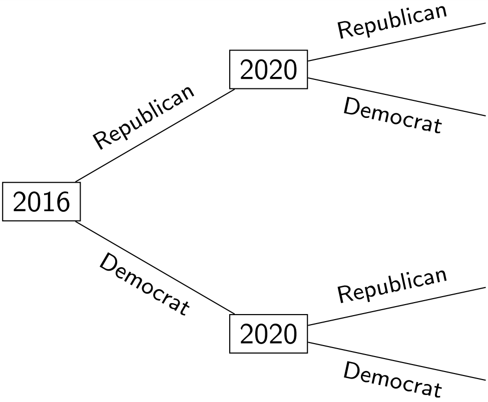
Conditional choice probability:

$$\rho_t(x_t, A_t|h^t) = \mathbb{P}\Big[U_t(x) = \max_{y_t \in A_t} U_t(y_t)\Big| h^t\Big]$$

Two main dynamic effects that connect ρ_t and ρ_{t+1}

- **Backward Looking**: (if U_t and U_{t+1} are correlated)
  - History-Dependence, Choice-Persistence
- **Forward Looking**: (if U_t satisfies the Bellman Equation)
  - Agent is forward-looking and Bayesian-rational

History Dependence and Selection on Unobservables



If political preferences persistent over time, expect history dependence:

$$\rho(R_{2020}|R_{2016}) > \rho(R_{2020}|D_{2016})$$

History independence only if preferences completely independent over time.

History Dependence is a result of informational asymmetry

Types of History Dependence (Heckman, 1981)

1. **Choice-Dependence**: A consequence of the informational asymmetry between the analyst and the agent
   - Dynamic selection on unobservables
   - Utility is serially correlated (past choices partially reveal it)

2. **Consumption-Dependence**: Past consumption changes the state of the agent
   - Habit formation or preference for variety (preferences change)
   - Experimentation (beliefs change)

For today, assume 2 away, focus on 1:

- Frick, Iijima, and Strzalecki (2017) has an extension to 2
- Main question here: how much history-dependence can there be?
- What are the axioms that link ρ_t and ρ_{t+1}?

Dynamic Decisions

Decision Trees: $x_t = (z_t, A_{t+1})$

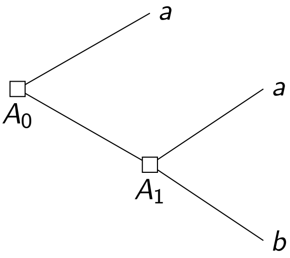
– Choice today leads to an immediate payoff and a menu for tomorrow

– Stopping problems, e.g., patent renewal, Pakes (1986)

Example: Stopping problem

You can either:

- buy the legacy iPhone a in $t = 0$ (and nothing in period 1)
- defer purchase till $t = 1$ and choose between a or new iPhone b



- Formally, $A_0 := \{a, A_1\}$ and $A_1 := \{a, b\}$
- Buying now is a, waiting is A_1

Bellman Equation

$$U_t(z_t, A_{t+1}) = u_t(z_t) + \mathbb{E}\left[\max_{y_{t+1} \in A_{t+1}} U_{t+1}(y_{t+1}) \middle| \mathcal{F}_t \right]$$

Bayesian Rationality:

- Preference for Flexibility (like bigger menus)
- Rational Expectations (dynamic consistency)
- Preference for late decisions (value of information)

Dynamic Discrete Choice (DDC) models in Econometrics often assume

$$U_t(z_t, A_{t+1}) = v_t(z_t) + \mathbb{E}\left[\max_{y_{t+1} \in A_{t+1}} U_{t+1}(y_{t+1}) \middle| \mathcal{F}_t \right] + \epsilon_t^{(z_t, A_{t+1})}$$

If ϵ is i.i.d., this can lead to

- violations of Bayesian Rationality
- biased estimates in optimal stopping problems
- this generalizes beyond i.i.d. ϵ

This paper

Analyzes fully general/nonparametric model of dynamic random utility:

1. Axiomatically characterize implied dynamic stochastic choice behavior
   - Backward-looking axioms
   - Forward-looking axioms

2. Axiomatic analysis and comparative statics of persistence

3. Relationship with the DDC—modeling tradeoffs

Dynamic Random Utility

Decision Trees

Time: $t = 0, 1$

Per-period outcomes: Z

Decision Nodes: $\mathcal{A}_t$ defined recursively:

- period 1: menu A_1 is a subset of $X_1 := Z$
- period 0: menu A_0 is a subset of $X_0 := Z \times \mathcal{A}_1$

  pairs $x_0 = (z_0, A_1)$ of current outcome and continuation menu

Comment: Everything extends to finite horizon by backward induction

Conditional Choice Probabilities

ρ is a sequence of **history-dependent** choice distributions:

period 0: for each menu A_0, observe choice distribution

$$\rho_0(\cdot, A_0) \in \Delta(A_0)$$

period 1: for each menu A_1 and history h^0 that leads to menu A_1, observe choice distribution conditional on h^0

$$\rho_1(\cdot, A_1 | h^0) \in \Delta(A_1)$$

Conditional Choice Probabilities

ρ is a sequence of **history-dependent** choice distributions:

period 0: for each menu A_0, observe choice distribution

$$\rho_0(\cdot, A_0) \in \Delta(A_0)$$

period 1: for each menu A_1 and history h^0 that leads to menu A_1, observe choice distribution conditional on h^0

$$\rho_1(\cdot, A_1 | h^0) \in \Delta(A_1)$$

$\mathcal{H}_0 \cdots\cdots\cdots\cdots$ period-0 histories

$$\mathcal{H}_0 := \{h^0 = (A_0, x_0) : \rho_0(x_0, A_0) > 0\}$$

$\mathcal{H}_0(A_1) \cdots\cdots\cdots$ is set of histories that lead to menu A_1

$$\mathcal{H}_0(A_1) := \{h^0 = (A_0, x_0) \in \mathcal{H}_0 : x_0 = (z_0, A_1) \text{ for some } z_0 \in Z\}$$

Dynamic Random Utility

Definition: A *DRU* representation of ρ consists of

- a probability space $(\Omega, \mathcal{F}, \mathbb{P})$
- a stochastic process of utilities $U_t : \Omega \to \mathbb{R}^{X_t}$

such that for all $x_0 \in A_0$

$$\rho_0(x_0, A_0) = \mathbb{P}\left[U_0(x_0) = \max_{y_0 \in A_0} U_0(y_0)\right]$$

and for all $x_1 \in A_1$ and histories $(A_0, x_0) \in \mathcal{H}_0(A_1)$,

$$\rho_1(x_1, A_1 | A_0, x_0) = \mathbb{P}\left[U_1(x_1) = \max_{y_1 \in A_1} U_1(y_1) \middle| U_0(x_0) = \max_{y_0 \in A_0} U_0(y_0)\right]$$

Ties

- For technical reasons allow for ties and use tie-breaking
- I will say that $\rho(x, A) > 0$ *modulo ties* if $\rho(x^n, A^n) > 0$ for $x^n \to x$ and $A^n \to A$ or something roughly like that
- Formalized by Ahn and Sarver (2013), we use similar notions
- I will gloss over this here and focus on conceptual points

History Independence

General idea:

- Agent's choice history $h^0 = (A_0, x_0)$ reveals something about his period-0 private information, so expect $\rho_1(\cdot|h^0)$ to depend on h^0

- But dependence cannot be arbitrary: some histories are *equivalent* as far as the private information they reveal

- Our axioms:
  - Identify two types of equivalence classes of histories
  - Impose history *independence* of ρ_1 within these classes

Contraction History Independence

Axiom (Contraction History Independence): If

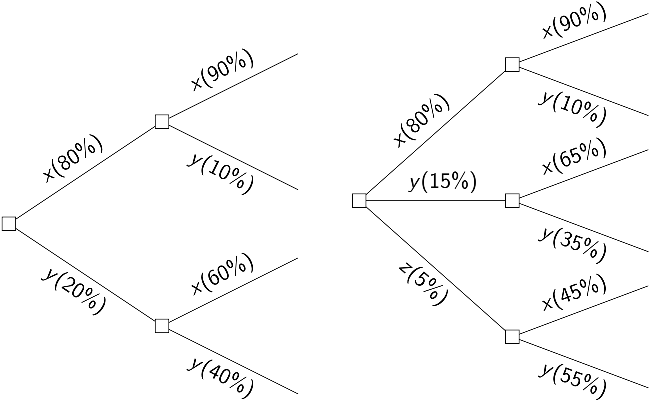
(i) $A_0 \subseteq B_0$

(ii) $\rho_0(x_0, A_0) = \rho_0(x_0, B_0)$,

then

$$\rho_1(\cdot, \cdot | A_0, x_0) = \rho_1(\cdot, \cdot | B_0, x_0)$$

Example



- z does not steal any customers from x in period $t = 0$
- so what people do in $t = 1$ after choosing x should be the same
- (note that z steals from y, so we have a mixture)

Necessity of CHI

Define the event $C(x_t, A_t)$ iff $U_t(x_t) = \max_{y_t \in A_t} U_t(y_t)$

Then

$$\rho_0(x_0, A_0) = \mathbb{P}\left[C(x_0, A_0)\right]$$

and for all $x_1 \in A_1$ and histories $(A_0, x_0) \in \mathcal{H}_0(A_1)$,

$$\rho_1(x_1, A_1 | A_0, x_0) = \mathbb{P}\left[C(x_1, A_1) \middle| C(x_0, A_0)\right]$$

Part (i) of CHI says $A_0 \subseteq B_0$ so $C(x_0, B_0) \subseteq C(x_0, A_0)$

Part (ii) of CHI says $\rho_0(x_0, A_0) = \rho_0(x_0, B_0)$, so the two events are identical almost surely

So conditioning on them should lead to the same prediction going forward

Adding Lotteries

Add lotteries: $X_t = \Delta(Z \times \mathcal{A}_{t+1})$, assume each utility function is vNM

- Denote lotteries by $p_t \in X_t$
- Helps formulate the second kind of history-independence
- Makes it easy to build on the REU axiomatization
- Helps overcome the limited observability problem
  - not all menus observed after a given history; how to impose axioms?
- Helps distinguish choice-dependence from consumption-dependence

$$h^0 = (A_0, x_0) \text{ vs } h^0 = (A_0, p_0, z_0)$$

Consumption History Independence

Assume away consumption dependence and allow only for choice dependence

Axiom (Consumption Independence): For any $p_0 \in A_0$ with $p_0(z_0), p_0(z_0') > 0$

$$\rho_1(\cdot|A_0, p_0, z_0) = \rho_1(\cdot|A_0, p_0, z_0')$$

Weak Linear History Independence

Idea: Under EU-maximization, choosing p_0 from A_0 reveals the same information as choosing option $\lambda p_0 + (1-\lambda) q_0$ from menu $\lambda A_0 + (1-\lambda)\{q_0\}$.

Axiom (Weak Linear History Independence)

$$\rho_1(\cdot,\cdot|A_0,p_0) = \rho_1\big(\cdot,\cdot|\lambda A_0 + (1-\lambda)q_0, \lambda p_0 + (1-\lambda)q_0\big).$$

Necessity of WLHI

Note we have

$$C\left(\frac{1}{2}p_0 + \frac{1}{2}q_0, \frac{1}{2}A_0 + \frac{1}{2}\{q_0\}\right) = C(p_0, A_0)$$

This is true because of Expected Utility:

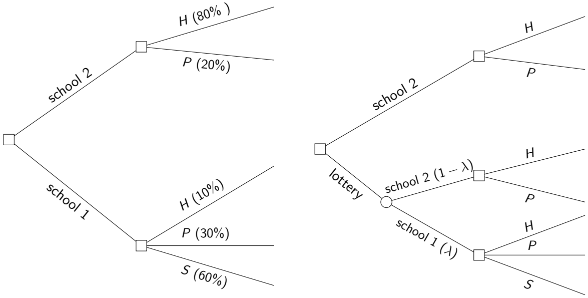
$$U_0\left(\frac{1}{2}p_0 + \frac{1}{2}q_0\right) \geq U_0\left(\frac{1}{2}r_0 + \frac{1}{2}q_0\right) \text{ for all } r_0 \in A_0$$

$$\Updownarrow$$

$$U_0(p_0) \geq U_0(r_0) \text{ for all } r_0 \in A_0$$

So conditioning on either of these events leads to the same prediction

Example



- school 2 offers two after-school programs, school 1 offers three
- different partents self-select to different schools
- how would school-1 parents choose between $\{H, P\}$?
- lottery to get in to the school
- Axiom says choice between $\{H, P\}$ independent of λ

Linear History Independence

Axiom (Weak Linear History Independence)

$$\rho_1(\cdot,\cdot|A_0,p_0) = \rho_1\big(\cdot,\cdot|\lambda A_0 + (1-\lambda)q_0, \lambda p_0 + (1-\lambda)q_0\big).$$

Idea was to mix-in a lottery q_0. Next we mix-in a set of lotteries B_0

Axiom (Linear History Independence)

$$\rho_1(\cdot,\cdot|A_0,p_0)\rho_0(p_0,A_0)$$
$$=\sum_{q_0\in B_0}\rho_1\Big(\cdot,\cdot|\lambda A_0+(1-\lambda)B_0,\lambda p_0+(1-\lambda)q_0\Big)\cdot\rho_0\Big(\lambda p_0+(1-\lambda)q_0,\lambda A_0+(1-\lambda)B_0\Big)$$

Necessity of LHI

Note that by Expected Utility we have

$$C\left(\frac{1}{2}p_0 + \frac{1}{2}q_0, \frac{1}{2}A_0 + \frac{1}{2}B_0\right)$$

$$\Updownarrow$$

$C(p_0, A_0)$ **and** $C(q_0, B_0)$

Necessity of LHI

Axiom (Linear History Independence)

$$\rho_1(\cdot,\cdot|A_0,p_0)\rho_0(p_0,A_0)$$

$$=\sum_{q_0\in B_0}\rho_1\Big(\cdot,\cdot|\lambda A_0+(1-\lambda)B_0,\lambda p_0+(1-\lambda)q_0\Big)\cdot\rho_0\Big(\lambda p_0+(1-\lambda)q_0,\lambda A_0+(1-\lambda)B_0\Big)$$

Under the representation, this is equivalent to:

$$\mathbb{P}(E|C(p_0,A_0))\mathbb{P}(C(p_0,A_0))$$
$$=\sum_{q_0\in B_0}\mathbb{P}\Big(E|C(p_0,A_0)\cap C(q_0,B_0)\Big)\cdot\mathbb{P}\Big(C(p_0,A_0)\cap C(q_0,B_0)\Big)$$

This is equivalent to

$$\mathbb{P}(E\cap C(p_0,A_0))=\sum_{q_0\in B_0}\mathbb{P}\Big(E\cap C(p_0,A_0)\cap C(q_0,B_0)\Big)$$

This is the Law of Total Probability

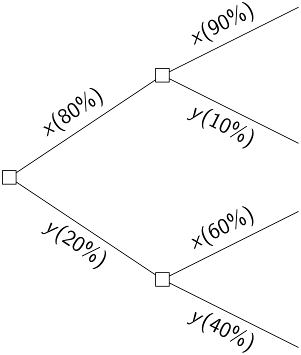
Dynamic Random Expected Utility

Theorem 1: ρ has a DREU representation if and only it satisfies

- Contraction History Independence
- Consumption History Independence
- Linear History Independence
- REU axioms in each period†
- History-Continuity†

Remark: For REU axioms we use the approach of Gul and Pesendorfer (2006); Ahn and Sarver (2013). We need to extend their result to infinite spaces because X_1 is infinite (our Theorem 0).

Consumption Persistence



- $\rho_1(x|x) > \rho_1(x|y)$
- again, there is no habit here
- but serially correlated utility
- widely studied in marketing literature
- comparative statics?

Consumption Persistence

Decision trees in which $t = 0$ choice does not influence $t = 1$ menus.

Let $C \subseteq \Delta(Z)$ denote a typical **consumption menu**

Primitive consists of:

- period 0 consumption choice: $\rho_0(c_0, C_0)$
- period 1 consumption choice: $\rho_1(c_1, C_1 | C_0, c_0)$

Axiom: ρ features **consumption persistence** if for all consumption menus $C_1 \subseteq C_0$ without ties, and $c, c' \in C_0$,

$$\rho_1(c, C_1 | C_0, c) \geq \rho_1(c, C_1 | C_0, c')$$

Example

Suppose that U_t follows an irreducible Markov chain

- $\mathcal{U} := \{u^1, ..., u^m\}$, transition matrix M
- assumptions:
  - no collinearity: $u^i \notin [\mathrm{co}\{u^j, u^k, u^\ell\}]$ for all i, j, k, ℓ
  - uniformly-ranked pair: $\exists \quad \overline{c}, \underline{c} \in \Delta(Z)$ s.t. $u^i(\overline{c}) > u^i(\underline{c})$ for all i
  - initial distribution has full support (but need not be the stationary distribution)

Example

Corollary: In the Markov chain example, TFAE:

1. ρ features consumption persistence

2. $(\mathcal{U}, M)$ is a **renewal process**, i.e., $\exists \alpha \in [0, 1)$ and $\nu \in \Delta(\mathcal{U})$ such that $M_{ii} = \alpha + (1 - \alpha)\nu(u^i)$ and $M_{ij} = (1 - \alpha)\nu(u^j)$

So either you stay put, or switch randomly according to the stationary distribution.

In the paper:

- Comparative statics: definition in terms of ρ' and $\rho \iff \alpha' > \alpha$
- General characterization (outside of Markov)
- Axioms for Markov (trivial for two periods, but not in general)

Dynamic Optimality

How to incorporate Dynamic Optimality?

- In the definition above, no structure on the family (U_t)
- But typically U_t satisfies the Bellman equation

Definition: ρ has an *Bayesian Evolving Utility* (BEU) representation if it has a DREU representation where the process (U_t) satisfies

$$U_t(z_t, A_{t+1}) = u_t(z_t) + \delta \mathbb{E}\left[\max_{p_{t+1} \in A_{t+1}} U_{t+1}(p_{t+1}) | \mathcal{F}_t\right]$$

for $\delta > 0$ and a $\mathcal{F}_t$-adapted process of vNM utilities $u_t : \Omega \to \mathbb{R}^Z$

Question: What are the additional assumptions?

Answer:

- Option value calculation (Preference for Flexibility)
- Rational Expectations (Sophistication)

Preference for Flexibility

We develop the stochastic version of axioms of Kreps (1979); Dekel, Lipman, and Rustichini (2001)

Axiom (Preference for Flexibility): For any A_1, B_1 such that $A_1 \subseteq B_1$

$$\rho_0\big((z_0, B_1), \{(z_0, A_1), (z_0, B_1)\}\big) = 1$$

modulo ties.[†]

Axiom (Stochastic DLR) Preference for Flexibility + Technical conditions[†]

Result Stochastic DLR + Separability[†] implies

$$U_t(z_t, A_{t+1}) = u_t(z_t) + \delta \hat{\mathbb{E}} \left[\max_{p_{t+1} \in A_{t+1}} U_{t+1}(p_{t+1}) | \mathcal{F}_t \right]$$

for some expectation operator $\hat{\mathbb{E}}$, possibly different than the true DGP

Rational Expectations (following trivial history)

- Need an axiom that ensures that $\hat{\mathbb{E}} = \mathbb{E}$ i.e., beliefs=DGP
- Fix a trivial history $h_0 = (\{p_0\}, p_0)$ and menus $B_1 \supset A_1$

  Agent sometimes chooses an option in $B_1 \setminus A_1$ following h_0

  $$\Updownarrow$$

  In some states of the world she must value B_1 strictly more than A_1

- Like Ahn and Sarver (2013) but they have deterministic $t = 0$ choice

Axiom (Sophistication): For any $h_0 = (\{p_0\}, p_0)$ and $B_1 \supset A_1$ the following are equivalent modulo ties†

1. $\rho_1(p_1, B_1 | h^0) > 0$ for some $p_1 \in B_1 \setminus A_1$
2. $\rho_0\Big((z, B_1), \big\{(z, B_1), (z, A_1)\big\}\Big) = 1$

Rational Expectations (following any history)

- Now fix *any* history $h_0 = (A_0, p_0)$ and menus $B_1 \supset A_1$

- Agent sometimes chooses an option in $B_1 \setminus A_1$ following h_0

$$\Updownarrow$$

- In some states of the world in which she chooses p_0 from A_0, she must value B_1 strictly more than A_1

Axiom (Conditional Sophistication): For any $h_0 = (A_0, p_0)$ and $B_1 \supset A_1$ the following are equivalent modulo ties†

1. $\rho_1(p_1, B_1 | A_0, p_0) > 0$ for some $p_1 \in B_1 \setminus A_1$
2. $\rho_0\Big(\frac{1}{2}p_0 + \frac{1}{2}(z, B_1), \frac{1}{2}A_0 + \frac{1}{2}\big\{(z, B_1), (z, A_1)\big\}\Big) > 0$

Analogues in econometrics

- Analogue of Sophistication is the Williams-Daly-Zachary theorem
  - ρ_1 is the gradient of U_0 (in the space of utilities)
  - see, e.g., Chiong, Galichon, and Shum (2016)
  - It is an envelope-theorem result, like the Hotelling lemma
- Hotz and Miller (1993) and the literature that follows exploits this relationship
- Our axiom is in a sense a “test” of this property

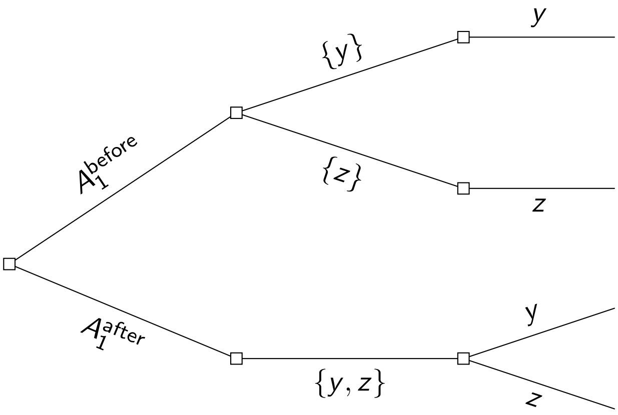
Characterization of BEU

Theorem 2: Suppose that ρ admits a DREU representation.

ρ has a BEU representation iff ρ satisfies Separability, Stochastic DLR, and Conditional Sophistication.

Preference for making choices late

- Suppose you got admitted to PhD programs at Harvard and MIT
- Do you make your decision before the visit days or after?



Preference for making choices late

Proposition 2: If ρ has a BEU representation, then absent ties[†]

$$\rho_0(A_1^{\text{after}}, \{A_1^{\text{before}}, A_1^{\text{after}}\}) = 1$$

Comment:

- BEU has positive value of information: desire to delay the choice as late as possible to capitalize on incoming information (unless there is a cost)

Learning

- Bayesian Evolving Utility: randomness in choice comes from changing tastes
- Bayesian Evolving Beliefs: randomness in choice comes from random signals
  - tastes are time-invariant, but unknown $u_t = \mathbb{E}[\tilde{u}|\mathcal{G}_t]$ for some time-invariant vNM utility $\tilde{u} : \Omega \to \mathbb{R}^Z$
- To characterize BEB, need to add a "martingale" axiom (Theorem 3) or a "consumption-inertia" axiom (Propositon 6)

Identification

- Uniqueness of the utility process, discount factor, and information (Proposition I.1)
- There is a vast DDC literature on identification (Manski, 1993; Rust, 1994; Magnac and Thesmar, 2002; Norets and Tang, 2013)
  - δ not identified unless make assumptions about "observable attributes"
  - How does this compare to the "menu variation" approach

Dynamic Discrete Choice

DDC model

Definition: The *DDC model* is a restriction of DREU to deterministic decision trees that additionally satisfies the Bellman equation

$$U_t(z_t, A_{t+1}) = v_t(z_t) + \delta\mathbb{E}\left[\max_{y_{t+1}\in A_{t+1}} U_{t+1}(y_{t+1})|\mathcal{F}_t\right] + \epsilon_t^{(z_t, A_{t+1})},$$

with deterministic utility functions $v_t : \Omega \to \mathbb{R}^Z$; discount factor $\delta \in (0,1)$; and $\mathcal{F}_t$-adapted zero-mean *payoff shocks* $\tilde{\epsilon}_t : \Omega \to \mathbb{R}^{Y_t}$.

Special cases of DDC

- **BEU** is a special case, which can be written by setting $\epsilon_t^{(z_t, A_{t+1})} = \epsilon_t^{(z_t, B_{t+1})}$
  - **shocks to consumption**
- **i.i.d. DDC** where $\epsilon_t^{(z_t, A_{t+1})}$ and $\epsilon_\tau^{(y_t, B_{t+1})}$ are i.i.d.
  - **shocks to actions**

Other special cases of DDC

- **permanent unobserved heterogeneity**: $\varepsilon_t^{(z_t, A_{t+1})} = \pi_t^{z_t} + \theta_t^{(z_t, A_{t+1})}$, where
  - $\pi_t^{z_t}$ is a "permanent" shock that is measurable with respect to $\mathcal{F}_0$
  - $\theta_t^{(z_t, A_{t+1})}$ is a "transitory" shock, i.i.d. conditional on $\mathcal{F}_0$
- **transitory but correlated shocks to actions**: $\varepsilon_t^{(z_t, A_{t+1})}$ and $\varepsilon_\tau^{(x_\tau, B_{\tau+1})}$ are i.i.d. whenever $t \neq \tau$, but might be correlated within any fixed period $t = \tau$

Dynamic logit

- A special case of i.i.d. DDC where ϵ_t are distributed extreme value
- Dynamic logit is a workhorse for estimation
  - e.g., Miller (1984), Rust (1989), Hendel and Nevo (2006), Gowrisankaran and Rysman (2012)
- Very tractable due to the “log-sum” expression for “consumer surplus”

$$V_t(A_{t+1}) = \log \left(\sum_{x_{t+1} \in A_{t+1}} e^{v_{t+1}(x_{t+1})} \right)$$

(This formula is also the reason why nested logit is so tractable)

- Axiomatization (Fudenberg and Strzalecki, 2015)

Understanding the role of i.i.d. ϵ

Key Assumption: Shocks to actions, $\epsilon_t^{(z_t, A_{t+1})}$ and $\epsilon_t^{(z_t, B_{t+1})}$ are i.i.d. regardless of the nature of the menus A_{t+1} and B_{t+1}

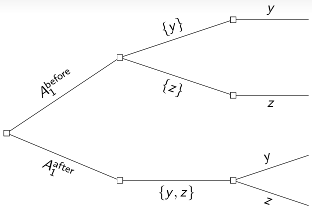
Let $A_0 := \{(z_0, A_1^{\text{small}}), (z_0, A_1^{\text{big}})\}$ where $A_1^{\text{small}} = \{z_1\}$ and $A_1^{\text{big}} = \{z_1, z_1'\}$.

Proposition 1: If ρ has a i.i.d. DDC representation, then

$$0 < \rho_0 \left((z_0, A_1^{\text{small}}), A_0 \right) < 0.5.$$

Moreover, if the ε shocks are scaled by $\lambda > 0$, then this probability is strictly increasing in λ whenever $v_1(z_1') > v_1(z_1)$.

Understanding the role of i.i.d. ϵ



Proposition 2: If ρ has a i.i.d. DDC representation with $\delta < 1$, then

$$0.5 < \rho_0 \left((x, A_1^{\text{early}}), A_0 \right) < 1.$$

Moreover, if ε is scaled by $\lambda > 0$, then $\rho_0((x, A_1^{\text{early}}), A_0)$ is strictly increasing in λ (modulo ties).

Intuition:

- The agent gets the ϵ not at the time of consumption but at the time of decision (even if the decision has only delayed consequences)
- So making decisions early allows him to get the $\max \epsilon$ earlier

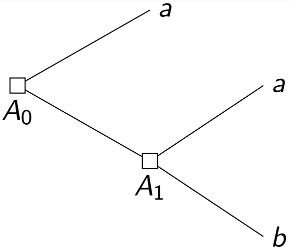
Beyond i.i.d. DDC

- This result extends in a straightforward way to DDC with permanent unobserved heterogeneity
  - this is just a mixture of i.i.d DDC models, so inherits this property
- Also to DDC with transitory but correlated shocks to actions
- Final model: mixture of i.i.d. DDC with BEU
  - horse race between the two effects

Other Decision Problems

- So far, looked at pure manifestations of option value
  - direct choice between nested menus
  - costless option to defer choice
- DDC models typically not applied to those
- But these forces exist in “nearby” choice problems
- So specification of shocks matters more generally

Biased Parameter Estimates



Parameters: $v_0(a) = v_1(a) = w$ and $v_1(b) = 0$, discount factor δ

Proposition 3: Suppose that the data generating process ρ is compatible with both BEU and i.i.d. DDC. If the distribution of ϵ has a symmetric and unimodal density, then the MLE estimators almost surely satisfy:

1. $\lim_n \hat{w}_n^{\mathrm{DDC}} = \lim_n \hat{w}_n^{\mathrm{BEU}}$

2a. $\lim_n \hat{\delta}_n^{\mathrm{DDC}} < \lim_n \hat{\delta}_n^{\mathrm{BEU}}$ if $\rho_0(a; A_0) > 0.5$

2b. $\lim_n \hat{\delta}_n^{\mathrm{DDC}} > \lim_n \hat{\delta}_n^{\mathrm{BEU}}$ if $\rho_0(a; A_0) < 0.5$.

Modeling Choices

- BEU: so far few convenient parametrization (Pakes, 1986) but
  - bigger menus w/prob. 1
  - late decisions w/prob. 1
- i.i.d. DDC: widely used because of statistical tractability, but
  - smaller menus w/prob. $\in (0, \frac{1}{2})$
  - early decisions w/prob. $\in (\frac{1}{2}, 1)$

Comments:

- i.i.d. DDC violates a key feature of Bayesian rationality: positive option value
- Model Misspecification
  - Maybe a fine model of (behavioral) consumers
  - But what about profit maximizing firms?

Thank you!

References I

Ahn, D. S., and T. Sarver (2013): "Preference for flexibility and random choice," *Econometrica*, 81(1), 341–361.

Chiong, K. X., A. Galichon, and M. Shum (2016): "Duality in dynamic discrete-choice models," *Quantitative Economics*, 7(1), 83–115.

Dekel, E., B. Lipman, and A. Rustichini (2001): "Representing preferences with a unique subjective state space," *Econometrica*, 69(4), 891–934.

Frick, M., R. Iijima, and T. Strzalecki (2017): "Dynamic Random Utility," *mimeo*.

Fudenberg, D., and T. Strzalecki (2015): "Dynamic logit with choice aversion," *Econometrica*, 83(2), 651–691.

Gowrisankaran, G., and M. Rysman (2012): "Dynamics of Consumer Demand for New Durable Goods," *mimeo*.

Gul, F., and W. Pesendorfer (2006): "Random expected utility," *Econometrica*, 74(1), 121–146.

Heckman, J. J. (1981): "Heterogeneity and state dependence," in *Studies in labor markets*, pp. 91–140. University of Chicago Press.

Hendel, I., and A. Nevo (2006): "Measuring the implications of sales and consumer inventory behavior," *Econometrica*, 74(6), 1637–1673.

References II

HOTZ, V. J., AND R. A. MILLER (1993): "Conditional choice probabilities and the estimation of dynamic models," *The Review of Economic Studies*, 60(3), 497–529.

KREPS, D. (1979): "A representation theorem for" preference for flexibility"," *Econometrica*, pp. 565–577.

MAGNAC, T., AND D. THESMAR (2002): "Identifying dynamic discrete decision processes," *Econometrica*, 70(2), 801–816.

MANSKI, C. F. (1993): "Dynamic choice in social settings: Learning from the experiences of others," *Journal of Econometrics*, 58(1-2), 121–136.

MILLER, R. (1984): "Job matching and occupational choice," *The Journal of Political Economy*, 92, 1086–1120.

NORETS, A., AND X. TANG (2013): "Semiparametric Inference in dynamic binary choice models," *The Review of Economic Studies*, p. rdt050.

PAKES, A. (1986): "Patents as options: Some estimates of the value of holding European patent stocks," *Econometrica*, 54, 755–784.

RUST, J. (1989): "A Dynamic Programming Model of Retirement Behavior," in *The Economics of Aging*, ed. by D. Wise, pp. 359–398. University of Chicago Press: Chicago.

——— (1994): "Structural estimation of Markov decision processes," *Handbook of econometrics*, 4, 3081–3143.