Choice and Attention Across Time

Xi Zhi "RC" Lim

Presenting at:

Short, Online | 11/12/2022

This paper

Building on:

Choices are affected by attention.

Today:

Past choices affect attention ,

which in turn affects future choices.

- Example: Operating systems, vacation destinations, car brands, coffee rosters, who to date, recruitment, research topics...

This paper

- Cycle 1: makes more decisions ⇒ **considers more alternatives** in budget set ⇒ affects future choices

- Cycle 2: makes more decisions ⇒ **considers different alternatives** in budget set ⇒ affects future choices

- These cycles allow us to:
  1. exploit **connection** between past and future choices
  2. past choices help **predict** future behavior (attention)
  3. future choices help **explain** past behavior (attention)
  4. even when attention is an **unobservable**

In one-shot decisions:

$$\{\underline{\text{macOS}}, \text{Win}\}, \{\text{macOS}, \underline{\text{Win}}, \text{Linux}\}$$

In one-shot decisions:

$$\{\underline{\text{macOS}}, \text{Win}\}, \{\text{macOS}, \underline{\text{Win}}, \text{Linux}\}$$

$$\{\underline{\text{macOS}}, \text{Win}\}, \{\underline{\text{macOS}}, \underline{\text{Win}}, \text{Linux}\}$$

In one-shot decisions:

$$\{\underline{\text{macOS}}, \text{Win}\}, \{\text{macOS}, \underline{\text{Win}}, \text{Linux}\}$$

$$\{\underline{\text{macOS}}, \text{Win}\}, \{\text{macOS}, \underline{\text{Win}}, \text{Linux}\}$$

$\Longrightarrow$ Pins down neither attention

In one-shot decisions:

$$\{\underline{\text{macOS}}, \text{Win}\}, \{\text{macOS}, \underline{\text{Win}}, \text{Linux}\}$$

$$\ldots\ u(\text{macOS}) > u(\text{Win})$$

$$\{\underline{\text{macOS}}, \text{Win}\}, \{\text{macOS}, \underline{\text{Win}}, \text{Linux}\}$$

$$\ldots\ u(\text{Win}) > u(\text{macOS})$$

$\Longrightarrow$ Pins down neither attention nor preferences.

Attention affected by past choices

$$\{\text{macOS}, \underline{\text{Win}}, \text{Linux}\}$$

Attention affected by past choices

$$\{\underline{\text{macOS}}, \text{Win}\} \longrightarrow \{\text{macOS}, \underline{\text{Win}}, \text{Linux}\}$$

Attention affected by past choices

$$\{\underline{\text{macOS}}, \text{Win}\} \longrightarrow \{\text{macOS}, \underline{\text{Win}}, \text{Linux}\}$$

$$\Longrightarrow u(\text{Win}) > u(\text{macOS})$$

Literature

Decision theory and attention

- **Caplin, Dean** (2011) - Search
- **Masatlioglu, Nakajima, Ozbay** (2012) - Revealed Attention
- **Manzini, Mariotti** (2014) - Stochastic Choice
- **Caplin, Dean, Leahy** (2019) - Rational Inattention
- **Gossner**, **Steiner**, **Steward** (2021) - Manipulated Attention

All study one-shot decisions. This paper: choice sequences.

Primitive

Standard:

- X: countable set of alternatives
- $\mathscr{A}$: finite choice sets of at least size 2

Innovation:

- $X^{\mathbf{N}}$: infinite sequences of alternatives
- $\mathscr{A}^{\mathbf{N}}$: infinite sequences of choice sets
- (A_n): a generic sequence of choice sets

The primitive:

$$c : \mathscr{A}^{\mathbf{N}} \to X^{\mathbf{N}} \text{ where } c((A_n))_i \in (A_n)_i$$

Primitive

Assume **future-independence**:

$$(A_n), (B_n) \text{ such that } (A_n)_i = (B_n)_i \text{ for all } i \leq K$$
$$\Rightarrow c((A_n))_i = c((B_n))_i \text{ for all } i \leq K$$

Then, c is characterized by a set of ***cross-sectional*** choice functions:

$$\tilde{c}((A_1, ..., A_K)) : \mathscr{A} \to X$$

A1: Weak Stability

Axiom

Suppose $h < i < j$.

$c((A_n))_i = y$, $x \in A_i$ *and* $c((A_n))_h = x \Rightarrow c((A_n))_j \neq x$ *if* $y \in A_j$

(Blue parts: WARP)

Switching is allowed, but **not back and fourth**.

A **weaker** form of WARP imposed **within sequence**.

A2: Past Dependence

Axiom

$$\tilde{c}((A_1, ..., A_k))(B) \in \tilde{c}((A_1, ..., A_{k-1}))(B) \cup \tilde{c}((A_1, ..., A_{k-1}))(A_k).$$

(Blue parts: Past **In**dependence)

Past choices **may affect** future choices,

in a limited/tractable way.

Axiom 3: Default Attention

Every choice set contains alternatives that

"**always** receive attention."

- Notations:
  - Choice without history: $c_0(A) := \tilde{c}(\emptyset)(A)$
  - $c((A_n))_j = x$, $y \in A_j$, and $c((A_n))_i = y$ with $i < j \Rightarrow xPy$
  - $c((A_n))_j = x$ and $c_0(A_j) = y \Rightarrow xPy$

Axiom

If $c_0(A)\,Py$, then y is never chosen from A.

Attention Across Time (AAT) Representation

- Denote: $\boldsymbol{c}((A_1,...,A_k)) = \{\tilde{c}((A_1,...,A_{i-1}))(A_i) : i = 1,...,k\}$

Theorem

c satisfies A1-A3 if and only if there exist

1. $u : X \to \mathbb{R}$ (*utility function*)
2. $\Gamma : \mathscr{A} \to 2^X \setminus \{\emptyset\}$ (*default attention*)

such that

$$\tilde{c}(h)(A) = \underset{x \in \tilde{\Gamma}(h)(A)}{\arg\max}\, u(x)$$

$$\text{where } \tilde{\Gamma}(h)(A) = (\Gamma(A) \cup \boldsymbol{c}(h)) \cap A$$

The rest of the talk

1. **Counterfactual WARP**
2. **Necessary departures**
3. **Uniqueness of preferences**
4. **Attention across time vs at a time**

Counterfactual WARP Violations

Example 1:

$$u(4) > u(3) > u(2) > u(1)$$

$\Gamma(A)$:

$\{1,2,3,4\}$	$\{1,2,4\}$	$\{2,3,4\}$	$\{1,3\}$	$\{2,3\}$	$\{3,4\}$
$\{1,2,3\}$	$\{1,3,4\}$	$\{1,2\}$	$\{1,4\}$	$\{2,4\}$	

- c_0 **satisfies WARP** $2 \succ_0 4 \succ_0 3 \succ_0 1$, but **WARP violation** after some history:

$$\tilde{c}(\{1,3\})(\{1,2,3,4\}) = 3, \qquad \tilde{c}(\{1,3\})(\{4,3\}) = 4$$

Example 2:

$$u(3) > u(2) > u(1)$$

$$\Gamma(A): \quad \begin{matrix} \{1,2,3\} & \{1,3\} \\ \{1,2\} & \{2,3\} \end{matrix}$$

- **Never** counterfactual WARP violations, but **affected by attention**:

$$c_0(\{1,2\}) = 1, \qquad \tilde{c}(\{2,3\})(\{1,2\}) = 2$$

Last two examples:

Lack of counterfactual WARP violations

does not rule out limited attention.

Proposition

There exists h such that $\hat{c}(h)$ violates WARP if and only if there exist x, y, z such that $z \succ_0 x \succ_0 y$ and $u(x) > u(y) > u(z)$.

Counterfactual WARP violations $\Leftrightarrow$

sufficiently large "jump" between initial choice and true preference.

Necessary Departures

Standard u-rep

Definition

We say c admits a standard u-rep if there exists $u : X \to \mathbb{R}$ such that

$$\tilde{c}(h)(A) = \arg\max_{x \in (A)} u(x)$$

Necessary departures

Strengthen axioms:

A1*: WARP within sequence A2*: Past-Independence

Proposition

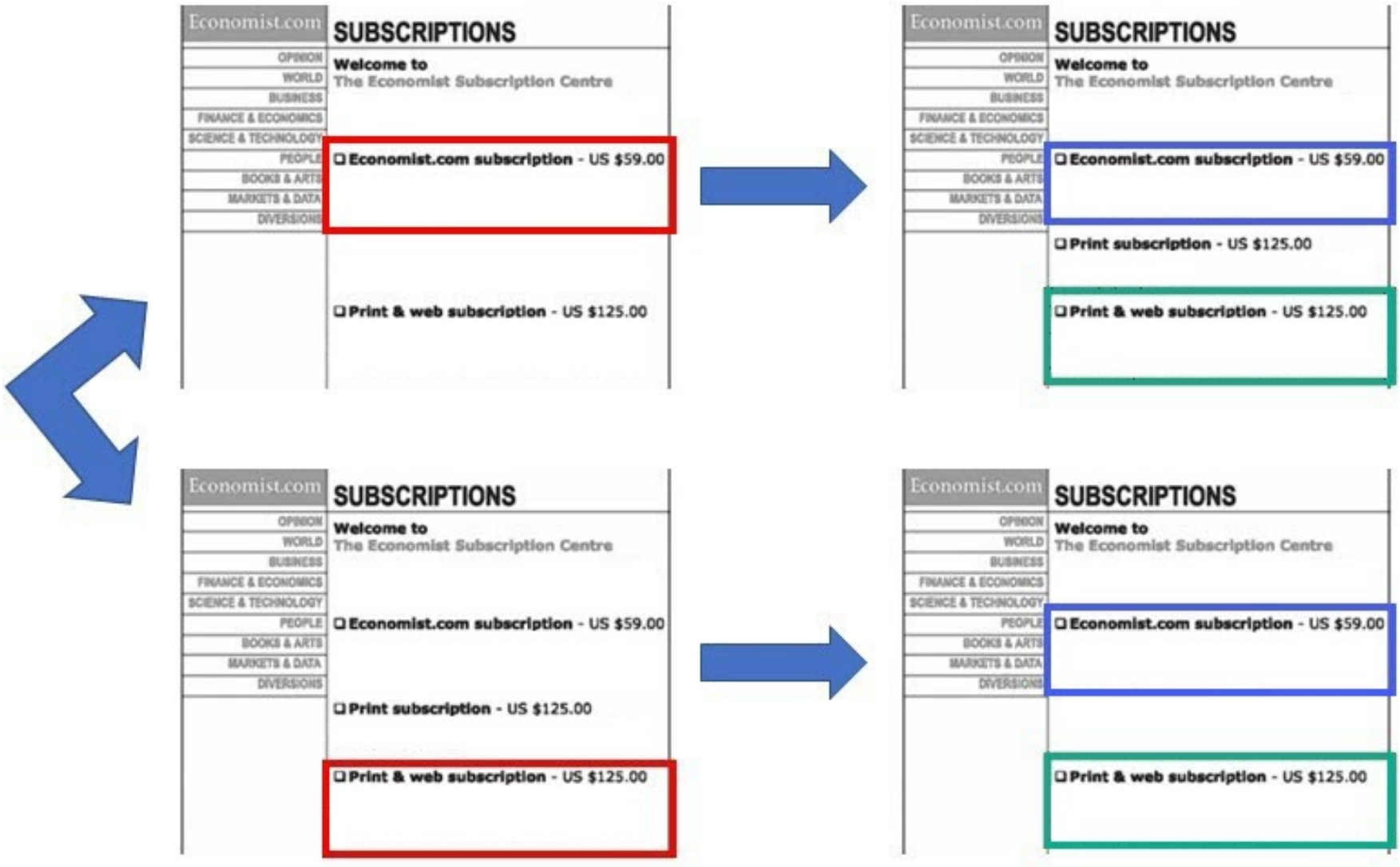
1. $A1 + A2 + A3 \Leftrightarrow$ *AAT*
2. $A1^* + A2 + A3 \Leftrightarrow$ *standard u–rep*
3. $A1 + A2^* + A3 \Leftrightarrow$ *standard u–rep*

In a non-trivial model of attention over time, it is **necessary**

that WARP violations occur AND past choices affect future choices.

Uniqueness of Preferences

AAT ⇒ Convergence



Must converge to **blue** or **green** ⇒ **reveals preference!**

Preferences pinned down

$\hat{X} :=$ Alternatives that are sometimes chosen.

Proposition

Fixed c, $u_1(x) > u_1(y) \iff u_2(x) > u_2(y)$ for all $x, y \in \hat{X}$.

Connection between past and future choices

pins down preferences.

Blessed if and only if cursed

In between two extremes:

No WARP violations	**Lots of WARP violations**
Cannot "confirm" preferences, but can predict as though there is no limited attention.	Can "confirm" preferences, and make predictions even though there is limited attention.

Attention across time vs at a time

Attention across time vs at a time

Choice with Limited Attention (CLA):

$$\hat{c}(A) = \underset{x \in \Gamma(A)}{\arg\max}\, u(x)$$

where Γ *is an attention filter, i.e.,* $z \notin \Gamma(A) \Rightarrow \Gamma(A \setminus \{z\}) = \Gamma(A)$.

...Masatlioglu Nakajima Ozbay (2012)

Proposition

Under AAT, if default attention Γ *is an attention filter, then*

1. $\forall h, \tilde{\Gamma}(h) : \mathscr{A} \to \mathscr{A}$ *is also an attention filter*
2. $\forall h, \tilde{c}(h) : \mathscr{A} \to X$ *is also a CLA*

Attention across time vs at a time

Rational Shortlist Methods (RSM):

$$\hat{c}(A) = \underset{x \in \Gamma(A)}{\arg\max}\, u(x)$$

where Γ *is an shortlist, i.e.,* $\Gamma(A) = \max(A, P)$ *for some binary relation* P*.*

...Manzini Mariotti (2007)

Proposition

Under AAT, if default attention Γ *is a shortlist, then*

1. $\forall h, \tilde{\Gamma}(h) : \mathscr{A} \to \mathscr{A}$ *is also an shortlist*
2. $\forall h, \tilde{c}(h) : \mathscr{A} \to X$ *is also a RSM*

Full Characterization

- $\mathbb{C}_{All}$: the universe choice functions (i.e., model-less)
- $\mathbb{C}$: a generic subset (i.e., a model)

Definitions:

1. g is a κ-*cousin* of f if $\exists X^* \subseteq f(\mathscr{A}) : g(A) = \kappa\left(\{f(A)\} \cup (A \cap X^*)\right) \; \forall A \in \mathscr{A}$.
2. $\mathbb{C}$ is *WARP-convex* if $\forall f \in \mathbb{C}$, $\exists \kappa$ sat. WARP : every κ-cousin of f is also in $\mathbb{C}$.

WARP-convex: "Correction of mistakes are permissible"

Theorem

$\mathbb{C}$ is fully compatible with AAT $\Leftrightarrow$ $\mathbb{C}$ is WARP-convex.

Conclude

Past and future choices are probably connected

This connection may contain information useful to us

This paper:

Models this connection in a framework of attention

What could we study?

Identification of preferences	Counterfactual WARP violations
Necessary departures	Attention across time vs at a time

Thanks for your attention (now & ever after)!