BEHAVIORAL ECONOMICS · LECTURE 6

Bounded rationality II: consideration sets

Niels Boissonnet – Bielefeld University

Master Economics – Winter 2026/27

Before we start: choose the best Scotch



Before we start: choose the best Scotch



Your task

- Which one did you pick? **How many bottles did you actually look at?**
- Why those ones: the price tag, a name you know, the label that stands out?

Learning objectives

By the end of this lecture you should be able to:

- Define a **consideration set** and explain why a model with arbitrary consideration sets has no testable content.
- State the **rational shortlist method** and its characterization (Expansion, Weak WARP).
- State **choice with limited attention**, derive **revealed preference** from WARP violations, and check whether a data set is consistent with it.
- Compute choice probabilities in the **random consideration set** model and say what the **random attention model** adds.
- Solve a simple **market game** in which firms compete for attention, and interpret its equilibrium.

Roadmap

Consideration sets

Shortlist methods

Revealed attention

Stochastic consideration sets

Application: consideration sets and marketing

Wrapping up

Where we are in the course

The assumptions of the Robinson model

0. **Consistent preferences:** Robinson's choices can be described by a utility function u
1. **Costless optimization:** he always finds the option that maximizes u
2. **Correct beliefs & expected utility:** he knows p and maximizes expected utility
3. **Absolute, self-regarding consumption:** only his own (b, y) enters u
4. **Stable tastes:** the same u in both periods

Still **Assumption 1**. Last week: the DM inspects options one at a time and stops when one is good enough. Today: the DM chooses the best option **among those she considers**, and the question is which options she considers.

Part 2: the cost of paying attention

Cost of considering options

- Satisficing (Lecture 5)
- Consideration sets (**today**, Lecture 6)

Cost of considering aspects of a decision

- Salience (Lecture 8)

Cost of acquiring information

- Bayesian updating: the toolkit (Lecture 7)
- Rational inattention (Lecture 9)

These approaches overlap conceptually.

1

Consideration sets

The decision maker chooses the best option among those she considers. Without structure on what she considers, the model can explain any choice.

Basic idea

Idea

Maybe you only considered a **subset** of the bottles.

- Decision makers focus their attention on a subset of the available alternatives.
- Marketers call this subset the **consideration set**.

Basic idea

Idea

Maybe you only considered a **subset** of the bottles.

- Decision makers focus their attention on a subset of the available alternatives.
- Marketers call this subset the **consideration set**.

How is this set determined?

- Cheapest brands?
- Brands you have heard of?
- Brands you have had before?
- Brands that stand out?
- At random?
- And **how many** alternatives?

A naive model

Setting

A finite set X of alternatives; menus $A \subseteq X$, $A \neq \emptyset$; a utility function $u : X \to \mathbb{R}$.

- A **consideration set mapping** E assigns to each menu A a nonempty subset $E(A) \subseteq A$: the alternatives the DM considers.
- **Choice:** the best *considered* alternative,

$$C(A) = \arg \max_{x \in E(A)} u(x).$$

A naive model

Setting

A finite set X of alternatives; menus $A \subseteq X$, $A \neq \emptyset$; a utility function $u : X \to \mathbb{R}$.

- A **consideration set mapping** E assigns to each menu A a nonempty subset $E(A) \subseteq A$: the alternatives the DM considers.
- **Choice:** the best *considered* alternative,

$$C(A) = \arg \max_{x \in E(A)} u(x).$$

Testable implications?

What are the testable implications of this model?

A naive model

Setting

A finite set X of alternatives; menus $A \subseteq X$, $A \neq \emptyset$; a utility function $u : X \to \mathbb{R}$.

- A **consideration set mapping** E assigns to each menu A a nonempty subset $E(A) \subseteq A$: the alternatives the DM considers.
- **Choice:** the best *considered* alternative,

$$C(A) = \arg \max_{x \in E(A)} u(x).$$

Testable implications?

What are the testable implications of this model?

- **None!** Any choice function C can be rationalized: set $E(A) = \{C(A)\}$ for every A.

A naive model

Setting

A finite set X of alternatives; menus $A \subseteq X$, $A \neq \emptyset$; a utility function $u : X \to \mathbb{R}$.

- A **consideration set mapping** E assigns to each menu A a nonempty subset $E(A) \subseteq A$: the alternatives the DM considers.
- **Choice:** the best *considered* alternative,

$$C(A) = \arg \max_{x \in E(A)} u(x).$$

Testable implications?

What are the testable implications of this model?

- **None!** Any choice function C can be rationalized: set $E(A) = \{C(A)\}$ for every A.

A testable model needs **structure on how consideration sets are formed**. Each model today is one such structure.

Reminder: WARP

Weak Axiom of Revealed Preference (WARP)

If $x, y \in A \cap B$, $x \neq y$, and $C(A) = x$, then $C(B) \neq y$.

If you choose x when y is available, you never choose y when x is available.

Reminder: WARP

Weak Axiom of Revealed Preference (WARP)

If $x, y \in A \cap B$, $x \neq y$, and $C(A) = x$, then $C(B) \neq y$.

If you choose x when y is available, you never choose y when x is available.

From Part 1

- A choice function maximizes a strict preference $\succ$ **iff** it satisfies WARP.
- So "choose the best in $E(A)$" can only be told apart from maximization through **violations of WARP**.

Reminder: WARP

Weak Axiom of Revealed Preference (WARP)

If $x, y \in A \cap B$, $x \neq y$, and $C(A) = x$, then $C(B) \neq y$.

If you choose x when y is available, you never choose y when x is available.

From Part 1

- A choice function maximizes a strict preference $\succ$ **iff** it satisfies WARP.
- So "choose the best in $E(A)$" can only be told apart from maximization through **violations of WARP**.

Example (Lecture 1, the decoy)

W = web only, P = print + web, D = print only (the decoy):

$$C(\{W, P\}) = W, \qquad C(\{W, P, D\}) = P.$$

P is chosen when W is available, and W is chosen when P is available: **WARP fails**.

2

Shortlist methods

Eliminate with one criterion, then choose with another: this explains some violations of WARP, but not the decoy effect.

Shortlist decision making

Idea

In many situations people decide in **two stages**:

1. **Shortlist:** eliminate the options that fail some intuitive criterion.
2. **Choice within the shortlist:** choose among the remaining options after a more detailed evaluation.

Examples

- An investor first eliminates portfolios that are too risky for her tolerance, then ranks the survivors by expected return.
- A recruiter first excludes candidates who lack the required skills, then selects the best remaining one on overall merit.

Shortlist decision making

Idea

In many situations people decide in **two stages**:

1. **Shortlist:** eliminate the options that fail some intuitive criterion.
2. **Choice within the shortlist:** choose among the remaining options after a more detailed evaluation.

Examples

- An investor first eliminates portfolios that are too risky for her tolerance, then ranks the survivors by expected return.
- A recruiter first excludes candidates who lack the required skills, then selects the best remaining one on overall merit.

Possibly connected to the fast “System 1” and slow “System 2” of Kahneman (2011).

Rational shortlist methods (Manzini and Mariotti, 2007)

Undominated elements

For an asymmetric relation $\succ$ on X, let the **undominated** elements of A be

$$\max(A; \succ) = \{x \in A : \text{ there is no } y \in A \text{ with } y \succ x\}.$$

Rational shortlist methods (Manzini and Mariotti, 2007)

Undominated elements

For an asymmetric relation $\succ$ on X, let the **undominated** elements of A be

$$\max(A;\succ) = \{x \in A : \text{ there is no } y \in A \text{ with } y \succ x\}.$$

The procedure

Given two asymmetric relations ("rationales") $\succ_1$ and $\succ_2$:

1. **Shortlist:** keep the options not eliminated by the first criterion, $E(A) = \max(A;\succ_1)$.
2. **Choose:** apply the second criterion to the shortlist, $C(A) = \max\left(E(A);\succ_2\right)$.

Rational shortlist methods (Manzini and Mariotti, 2007)

Undominated elements

For an asymmetric relation $\succ$ on X, let the **undominated** elements of A be

$$\max(A;\succ) = \{x \in A : \text{ there is no } y \in A \text{ with } y \succ x\}.$$

The procedure

Given two asymmetric relations ("rationales") $\succ_1$ and $\succ_2$:

1. **Shortlist:** keep the options not eliminated by the first criterion, $E(A) = \max(A;\succ_1)$.
2. **Choose:** apply the second criterion to the shortlist, $C(A) = \max\left(E(A);\succ_2\right)$.

Definition

A choice function C is a **rational shortlist method (RSM)** if there are asymmetric relations $\succ_1, \succ_2$ such that, for every menu A, $\max\left(\max(A;\succ_1);\succ_2\right)$ is a single alternative, namely $C(A)$.

An RSM can violate WARP

Two rationales

Let $X = \{x, y, z\}$ with

- first rationale: $x \succ_1 y$ (and nothing else);
- second rationale: $y \succ_2 z \succ_2 x$ (and $y \succ_2 x$).

An RSM can violate WARP

Two rationales

Let $X = \{x, y, z\}$ with

- first rationale: $x \succ_1 y$ (and nothing else);
- second rationale: $y \succ_2 z \succ_2 x$ (and $y \succ_2 x$).

Two menus

- Menu $\{x, y, z\}$: stage 1 eliminates y, so $E = \{x, z\}$; stage 2 gives $C(\{x, y, z\}) = z$.

An RSM can violate WARP

Two rationales

Let $X = \{x, y, z\}$ with

- first rationale: $x \succ_1 y$ (and nothing else);
- second rationale: $y \succ_2 z \succ_2 x$ (and $y \succ_2 x$).

Two menus

- Menu $\{x, y, z\}$: stage 1 eliminates y, so $E = \{x, z\}$; stage 2 gives $C(\{x, y, z\}) = z$.
- Menu $\{y, z\}$: nothing is eliminated, so $E = \{y, z\}$; stage 2 gives $C(\{y, z\}) = y$.

An RSM can violate WARP

Two rationales

Let $X = \{x, y, z\}$ with

- first rationale: $x \succ_1 y$ (and nothing else);
- second rationale: $y \succ_2 z \succ_2 x$ (and $y \succ_2 x$).

Two menus

- Menu $\{x, y, z\}$: stage 1 eliminates y, so $E = \{x, z\}$; stage 2 gives $C(\{x, y, z\}) = z$.
- Menu $\{y, z\}$: nothing is eliminated, so $E = \{y, z\}$; stage 2 gives $C(\{y, z\}) = y$.

An RSM can violate WARP

Two rationales

Let $X = \{x, y, z\}$ with

- first rationale: $x \succ_1 y$ (and nothing else);
- second rationale: $y \succ_2 z \succ_2 x$ (and $y \succ_2 x$).

Two menus

- Menu $\{x, y, z\}$: stage 1 eliminates y, so $E = \{x, z\}$; stage 2 gives $C(\{x, y, z\}) = z$.
- Menu $\{y, z\}$: nothing is eliminated, so $E = \{y, z\}$; stage 2 gives $C(\{y, z\}) = y$.

z is chosen when y is available, and y when z is available: **WARP fails**. Shortlisting explains *some* inconsistencies of choice. But not every choice function is an RSM.

Characterization

Expansion

For all menus A, B: if $C(A) = C(B) = x$, then $C(A \cup B) = x$.

Characterization

Expansion

For all menus A, B: if $C(A) = C(B) = x$, then $C(A \cup B) = x$.

Weak WARP

If $\{x, y\} \subseteq B \subseteq A$ and $C(A) = C(\{x, y\}) = x$, then $C(B) \neq y$.

If x beats y in the pair and in a large menu, y cannot be chosen in any menu in between.

Characterization

Expansion

For all menus A, B: if $C(A) = C(B) = x$, then $C(A \cup B) = x$.

Weak WARP

If $\{x, y\} \subseteq B \subseteq A$ and $C(A) = C(\{x, y\}) = x$, then $C(B) \neq y$.

If x beats y in the pair and in a large menu, y cannot be chosen in any menu in between.

Theorem (Manzini and Mariotti, 2007)

A choice function is a rational shortlist method if and only if it satisfies Expansion and Weak WARP.

Characterization

Expansion

For all menus A, B: if $C(A) = C(B) = x$, then $C(A \cup B) = x$.

Weak WARP

If $\{x, y\} \subseteq B \subseteq A$ and $C(A) = C(\{x, y\}) = x$, then $C(B) \neq y$.

If x beats y in the pair and in a large menu, y cannot be chosen in any menu in between.

Theorem (Manzini and Mariotti, 2007)

A choice function is a rational shortlist method if and only if it satisfies Expansion and Weak WARP.

Weaker than WARP, but still **testable** on choice data alone.

The decoy effect is not a rational shortlist method

Lecture 1 data, with the decoy D never chosen

$$C(\{W,P\}) = W, \quad C(\{W,D\}) = W, \quad C(\{P,D\}) = P, \quad C(\{W,P,D\}) = P.$$

The decoy effect is not a rational shortlist method

Lecture 1 data, with the decoy D never chosen

$$C(\{W,P\}) = W, \quad C(\{W,D\}) = W, \quad C(\{P,D\}) = P, \quad C(\{W,P,D\}) = P.$$

Expansion

$C(\{W,P\}) = C(\{W,D\}) = W$
requires $C(\{W,P,D\}) = W$.
Violated.

The decoy effect is not a rational shortlist method

Lecture 1 data, with the decoy D never chosen

$$C(\{W,P\}) = W, \quad C(\{W,D\}) = W, \quad C(\{P,D\}) = P, \quad C(\{W,P,D\}) = P.$$

Expansion

$C(\{W,P\}) = C(\{W,D\}) = W$ requires $C(\{W,P,D\}) = W$.
Violated.

Directly

For P to win the full menu, W must be eliminated at stage 1. Not by P, since W wins the pair $\{W,P\}$. So $D \succ_1 W$, but then $C(\{W,D\}) = D$, a contradiction.

The decoy effect is not a rational shortlist method

Lecture 1 data, with the decoy D never chosen

$$C(\{W,P\}) = W, \quad C(\{W,D\}) = W, \quad C(\{P,D\}) = P, \quad C(\{W,P,D\}) = P.$$

Expansion

$C(\{W,P\}) = C(\{W,D\}) = W$ requires $C(\{W,P,D\}) = W$.
Violated.

Directly

For P to win the full menu, W must be eliminated at stage 1. Not by P, since W wins the pair $\{W,P\}$. So $D \succ_1 W$, but then $C(\{W,D\}) = D$, a contradiction.

Shortlisting cannot explain the decoy effect: we need a different structure on consideration.

3

Revealed attention

If removing an option she does not consider changes nothing, preference is revealed exactly by violations of WARP.

Attention filters (Masatlioglu et al., 2012)

Key assumption

Removing an alternative that is not considered does not change the consideration set.

$$E(A \setminus \{x\}) = E(A) \qquad \text{whenever } x \notin E(A). \tag{AF}$$

A mapping E satisfying (AF) is an **attention filter**.

Attention filters (Masatlioglu et al., 2012)

Key assumption

Removing an alternative that is not considered does not change the consideration set.

$$E(A \setminus \{x\}) = E(A) \qquad \text{whenever } x \notin E(A). \tag{AF}$$

A mapping E satisfying (AF) is an **attention filter**.

Plausible procedures that satisfy (AF)

- the top k alternatives according to some criterion (e.g. the k cheapest);
- the top alternative in each category;
- all alternatives of the most popular category.

Attention filters (Masatlioglu et al., 2012)

Key assumption

Removing an alternative that is not considered does not change the consideration set.

$$E(A \setminus \{x\}) = E(A) \qquad \text{whenever } x \notin E(A). \tag{AF}$$

A mapping E satisfying (AF) is an **attention filter**.

Plausible procedures that satisfy (AF)

- the top k alternatives according to some criterion (e.g. the k cheapest);
- the top alternative in each category;
- all alternatives of the most popular category.

Definition

C is a **choice with limited attention (CLA)** if there are a strict preference $\succ$ on X and an attention filter E such that $C(A)$ is the $\succ$-best alternative in $E(A)$, for every menu A.

Example: choice difficulty (Tversky and Shafir, 1992)

Data

S = a Sony CD player at a bargain price, A = a top Aiwa model, W = wait and keep looking.

$$C(\{S, W\}) = S, \qquad C(\{S, A, W\}) = W.$$

Example: choice difficulty (Tversky and Shafir, 1992)

Data

S = a Sony CD player at a bargain price, A = a top Aiwa model, W = wait and keep looking.

$$C(\{S, W\}) = S, \qquad C(\{S, A, W\}) = W.$$

Attempt 1

$E(\{S, W\}) = \{S, W\}$,
$E(\{S, A, W\}) = \{W\}$. Is E an attention filter?

Example: choice difficulty (Tversky and Shafir, 1992)

Data

S = a Sony CD player at a bargain price, A = a top Aiwa model, W = wait and keep looking.

$$C(\{S, W\}) = S, \qquad C(\{S, A, W\}) = W.$$

Attempt 1

$E(\{S, W\}) = \{S, W\}$,
$E(\{S, A, W\}) = \{W\}$. Is E an attention filter?

- **No.** $A \notin E(\{S, A, W\})$, so (AF) requires $E(\{S, W\}) = E(\{S, A, W\}) = \{W\}$, but $E(\{S, W\}) = \{S, W\}$.

Example: choice difficulty (Tversky and Shafir, 1992)

Data

S = a Sony CD player at a bargain price, A = a top Aiwa model, W = wait and keep looking.

$$C(\{S, W\}) = S, \qquad C(\{S, A, W\}) = W.$$

Attempt 1

$E(\{S, W\}) = \{S, W\}$,
$E(\{S, A, W\}) = \{W\}$. Is E an attention filter?

- **No.** $A \notin E(\{S, A, W\})$, so (AF) requires $E(\{S, W\}) = E(\{S, A, W\}) = \{W\}$, but $E(\{S, W\}) = \{S, W\}$.

Attempt 2

$E(\{S, W\}) = \{S\}$,
$E(\{S, A, W\}) = \{S, A, W\}$, with $W \succ S$ and $W \succ A$.

Example: choice difficulty (Tversky and Shafir, 1992)

Data

S = a Sony CD player at a bargain price, A = a top Aiwa model, W = wait and keep looking.

$$C(\{S, W\}) = S, \qquad C(\{S, A, W\}) = W.$$

Attempt 1

$E(\{S, W\}) = \{S, W\}$,
$E(\{S, A, W\}) = \{W\}$. Is E an attention filter?

- **No.** $A \notin E(\{S, A, W\})$, so (AF) requires $E(\{S, W\}) = E(\{S, A, W\}) = \{W\}$, but $E(\{S, W\}) = \{S, W\}$.

Attempt 2

$E(\{S, W\}) = \{S\}$,
$E(\{S, A, W\}) = \{S, A, W\}$, with $W \succ S$ and $W \succ A$.

- **Yes.** Everything is considered in the large menu; in $\{S, W\}$, removing the unconsidered W leaves $\{S\}$. A CLA.

Example: the attraction effect

Data (attraction effect, Lecture 4)

Two attributes, more is better. Competitor $(3, 1)$, target $(1, 3)$, decoy $(1, 2)$: dominated by the target only.

$$C(\{(1,3),(3,1)\}) = (3,1), \qquad C(\{(1,3),(1,2),(3,1)\}) = (1,3).$$

Example: the attraction effect

Data (attraction effect, Lecture 4)

Two attributes, more is better. Competitor $(3,1)$, target $(1,3)$, decoy $(1,2)$: dominated by the target only.

$$C(\{(1,3),(3,1)\}) = (3,1), \qquad C(\{(1,3),(1,2),(3,1)\}) = (1,3).$$

A CLA representation

- preference $(3,1) \succ (1,3) \succ (1,2)$;
- $E\big(\{(1,3),(3,1)\}\big) = \{(1,3),(3,1)\}$;
- $E\big(\{(1,3),(1,2),(3,1)\}\big) = \{(1,3),(1,2)\}$: the decoy draws attention to the target and away from the competitor.

Example: the attraction effect

Data (attraction effect, Lecture 4)

Two attributes, more is better. Competitor $(3,1)$, target $(1,3)$, decoy $(1,2)$: dominated by the target only.

$$C(\{(1,3),(3,1)\}) = (3,1), \qquad C(\{(1,3),(1,2),(3,1)\}) = (1,3).$$

A CLA representation

- preference $(3,1) \succ (1,3) \succ (1,2)$;
- $E\big(\{(1,3),(3,1)\}\big) = \{(1,3),(3,1)\}$;
- $E\big(\{(1,3),(1,2),(3,1)\}\big) = \{(1,3),(1,2)\}$: the decoy draws attention to the target and away from the competitor.
- (AF) holds: removing the unconsidered $(3,1)$ from the large menu gives $E\big(\{(1,3),(1,2)\}\big) = \{(1,3),(1,2)\}$.

Example: the attraction effect

Data (attraction effect, Lecture 4)

Two attributes, more is better. Competitor $(3,1)$, target $(1,3)$, decoy $(1,2)$: dominated by the target only.

$$C(\{(1,3),(3,1)\}) = (3,1), \qquad C(\{(1,3),(1,2),(3,1)\}) = (1,3).$$

A CLA representation

- preference $(3,1) \succ (1,3) \succ (1,2)$;
- $E\big(\{(1,3),(3,1)\}\big) = \{(1,3),(3,1)\}$;
- $E\big(\{(1,3),(1,2),(3,1)\}\big) = \{(1,3),(1,2)\}$: the decoy draws attention to the target and away from the competitor.
- (AF) holds: removing the unconsidered $(3,1)$ from the large menu gives $E\big(\{(1,3),(1,2)\}\big) = \{(1,3),(1,2)\}$.

Exercise

Compromise effect: exercise.

Revealing consideration and preference

How can choice data reveal that y was considered?

- Suppose $C(A) = x$, $y \in A$, $y \neq x$, and $C(A \setminus \{y\}) \neq x$.

Revealing consideration and preference

How can choice data reveal that y was considered?

- Suppose $C(A) = x$, $y \in A$, $y \neq x$, and $C(A \setminus \{y\}) \neq x$.
- Removing y changed the choice, so E must have changed.

Revealing consideration and preference

How can choice data reveal that y was considered?

- Suppose $C(A) = x$, $y \in A$, $y \neq x$, and $C(A \setminus \{y\}) \neq x$.
- Removing y changed the choice, so E must have changed.
- By (AF), this is impossible if $y \notin E(A)$. So $y \in E(A)$.

Revealing consideration and preference

How can choice data reveal that y was considered?

- Suppose $C(A) = x$, $y \in A$, $y \neq x$, and $C(A \setminus \{y\}) \neq x$.
- Removing y changed the choice, so E must have changed.
- By (AF), this is impossible if $y \notin E(A)$. So $y \in E(A)$.
- But x was chosen while y was considered: x **is revealed preferred to** y.

Revealing consideration and preference

How can choice data reveal that y was considered?

- Suppose $C(A) = x$, $y \in A$, $y \neq x$, and $C(A \setminus \{y\}) \neq x$.
- Removing y changed the choice, so E must have changed.
- By (AF), this is impossible if $y \notin E(A)$. So $y \in E(A)$.
- But x was chosen while y was considered: x **is revealed preferred to** y.

Revealing consideration and preference

How can choice data reveal that y was considered?

- Suppose $C(A) = x$, $y \in A$, $y \neq x$, and $C(A \setminus \{y\}) \neq x$.
- Removing y changed the choice, so E must have changed.
- By (AF), this is impossible if $y \notin E(A)$. So $y \in E(A)$.
- But x was chosen while y was considered: x **is revealed preferred to** y.

Definition

$x \, P \, y$ if there is a menu $A \ni x, y$ such that $C(A) = x$ and $C(A \setminus \{y\}) \neq x$.

Revealed preference in the examples

The examples

- **Choice difficulty:** $C(\{S, A, W\}) = W$ and $C(\{S, W\}) = S \neq W$, so $W \; P \; A$. Nothing is revealed about W versus S.

Revealed preference in the examples

The examples

- **Choice difficulty:** $C(\{S, A, W\}) = W$ and $C(\{S, W\}) = S \neq W$, so $W \; P \; A$. Nothing is revealed about W versus S.
- **Attraction effect:** $C(\{(1,3),(1,2),(3,1)\}) = (1,3)$ and $C(\{(1,3),(3,1)\}) \neq (1,3)$, so $(1,3) \, P \, (1,2)$: the target is better than its decoy. Nothing about target versus competitor.

Revealed preference in the examples

The examples

- **Choice difficulty:** $C(\{S, A, W\}) = W$ and $C(\{S, W\}) = S \neq W$, so $W \; P \; A$. Nothing is revealed about W versus S.
- **Attraction effect:** $C(\{(1,3),(1,2),(3,1)\}) = (1,3)$ and $C(\{(1,3),(3,1)\}) \neq (1,3)$, so $(1,3) \, P \, (1,2)$: the target is better than its decoy. Nothing about target versus competitor.
- **Binary data** never reveal anything: $C(\{x, y\}) = x = C(\{x\})$.

Revealed preference in the examples

The examples

- **Choice difficulty:** $C(\{S, A, W\}) = W$ and $C(\{S, W\}) = S \neq W$, so $W \; P \; A$. Nothing is revealed about W versus S.
- **Attraction effect:** $C(\{(1,3),(1,2),(3,1)\}) = (1,3)$ and $C(\{(1,3),(3,1)\}) \neq (1,3)$, so $(1,3) \, P \, (1,2)$: the target is better than its decoy. Nothing about target versus competitor.
- **Binary data** never reveal anything: $C(\{x, y\}) = x = C(\{x\})$.

Revealed preference in the examples

The examples

- **Choice difficulty:** $C(\{S, A, W\}) = W$ and $C(\{S, W\}) = S \neq W$, so $W \, P \, A$. Nothing is revealed about W versus S.
- **Attraction effect:** $C(\{(1,3),(1,2),(3,1)\}) = (1,3)$ and $C(\{(1,3),(3,1)\}) \neq (1,3)$, so $(1,3) \, P \, (1,2)$: the target is better than its decoy. Nothing about target versus competitor.
- **Binary data** never reveal anything: $C(\{x, y\}) = x = C(\{x\})$.

Under limited attention, a choice reveals a preference only when **removing** an option **changes** the choice. Revealed preference becomes much **sparser** than in the standard model.

Revealed attention: the characterization

Theorem (Masatlioglu et al., 2012)

C is a CLA if and only if P is acyclic: there are no $x_1, \ldots, x_n$ with $x_1 \, P \, x_2 \, P \cdots P \, x_n \, P \, x_1$.

Proof idea

- **Necessity:** in any CLA, $x \, P \, y$ implies $x \succ y$, and $\succ$ has no cycles.

Revealed attention: the characterization

Theorem (Masatlioglu et al., 2012)

C is a CLA if and only if P is acyclic: there are no $x_1, \ldots, x_n$ with $x_1 \, P \, x_2 \, P \cdots P \, x_n \, P \, x_1$.

Proof idea

- **Necessity:** in any CLA, $x \, P \, y$ implies $x \succ y$, and $\succ$ has no cycles.
- **Sufficiency (idea):** take a $\succ$ that extends the transitive closure of P; construct the "minimal" consideration sets compatible with the data, and check (AF).

Revealed attention: the characterization

Theorem (Masatlioglu et al., 2012)

C is a CLA if and only if P is acyclic: there are no $x_1, \ldots, x_n$ with $x_1 \, P \, x_2 \, P \cdots P \, x_n \, P \, x_1$.

Proof idea

- **Necessity:** in any CLA, $x \, P \, y$ implies $x \succ y$, and $\succ$ has no cycles.
- **Sufficiency (idea):** take a $\succ$ that extends the transitive closure of P; construct the "minimal" consideration sets compatible with the data, and check (AF).

Revealed attention: the characterization

Theorem (Masatlioglu et al., 2012)

C is a CLA if and only if P is acyclic: there are no $x_1, \ldots, x_n$ with $x_1 \, P \, x_2 \, P \cdots P \, x_n \, P \, x_1$.

Proof idea

- **Necessity:** in any CLA, $x \, P \, y$ implies $x \succ y$, and $\succ$ has no cycles.
- **Sufficiency (idea):** take a $\succ$ that extends the transitive closure of P; construct the "minimal" consideration sets compatible with the data, and check (AF).

A data set that is not a CLA

$$C(\{x,y,z\}) = x,\ C(\{x,y\}) = y \ \Rightarrow\ x \, P \, z; \qquad C(\{x,z,w\}) = z,\ C(\{z,w\}) = w \ \Rightarrow\ z \, P \, x.$$

WARP with limited attention

WARP (equivalent form)

For every menu A there is $x^* \in A$ such that, for every menu $B \ni x^*$,

$$C(B) \in A \implies C(B) = x^*.$$

WARP with limited attention

WARP (equivalent form)

For every menu A there is $x^* \in A$ such that, for every menu $B \ni x^*$,

$$C(B) \in A \implies C(B) = x^*.$$

WARP with limited attention

For every menu A there is $x^* \in A$ such that, for every menu $B \ni x^*$,

$$\left[C(B) \in A \text{ and } C(B) \neq C(B \setminus \{x^*\})\right] \implies C(B) = x^*.$$

WARP with limited attention

WARP (equivalent form)

For every menu A there is $x^* \in A$ such that, for every menu $B \ni x^*$,

$$C(B) \in A \implies C(B) = x^*.$$

WARP with limited attention

For every menu A there is $x^* \in A$ such that, for every menu $B \ni x^*$,

$$\left[C(B) \in A \text{ and } C(B) \neq C(B \setminus \{x^*\})\right] \implies C(B) = x^*.$$

Theorem (Masatlioglu et al., 2012)

C satisfies WARP with limited attention if and only if it is a CLA.

So, do we like this paper?

Pros

- A model of consideration sets with **testable implications**.
- Built on "natural" restrictions on how consideration works. *But how natural is (AF)?*

Cons

- The testable implications may be **weak**.
- If there are no violations of WARP, there is **no revealed preference** information at all.

So, do we like this paper?

Pros

- A model of consideration sets with **testable implications**.
- Built on "natural" restrictions on how consideration works. *But how natural is (AF)?*

Cons

- The testable implications may be **weak**.
- If there are no violations of WARP, there is **no revealed preference** information at all.

The same technique works for other assumptions on consideration, e.g. Lleras et al. (2017), where consideration sets shrink when the menu grows ("more is less").

Recap: deterministic consideration sets

Unrestricted

Unrestricted consideration sets explain **everything**, so nothing.

Shortlists

Manzini and Mariotti (2007): two criteria applied in sequence; characterized by Expansion and Weak WARP; cannot explain the decoy effect.

Attention filters

Masatlioglu et al. (2012): removing an unconsidered option changes nothing; CLA $\iff$ P acyclic $\iff$ WARP with limited attention.

Preference is revealed only by **WARP violations**: $x\,P\,y$ when removing y changes a choice of x.

4

Stochastic consideration sets

Random attention, rather than random tastes, can explain why choices from the same menu vary.

Stochastic choice and consideration sets

Deterministic consideration is **too deterministic**: in real data, the choice from a given menu varies.

Stochastic choice and consideration sets

Deterministic consideration is **too deterministic**: in real data, the choice from a given menu varies.

Stochastic choice

- Recall the **stochastic choice function** of Lecture 3: for each menu A and $x \in A$,

  $$p(x \mid A) = \text{probability that } x \text{ is chosen from } A.$$

- Recent papers explain this randomness by **random consideration sets**:
  - a **deterministic** preference $\succ$;
  - a **random** consideration set.

The random consideration set model

Assumptions

(Manzini and Mariotti, 2014)

- a strict preference $\succ$ on X;
- each x is considered with probability $\eta(x) \in (0, 1)$, **independently** across alternatives and **whatever the menu**;
- the DM picks the $\succ$-best considered alternative; if she considers nothing, she picks a **default** (e.g. no purchase).

The random consideration set model

Assumptions

(Manzini and Mariotti, 2014)

- a strict preference $\succ$ on X;
- each x is considered with probability $\eta(x) \in (0,1)$, **independently** across alternatives and **whatever the menu**;
- the DM picks the $\succ$-best considered alternative; if she considers nothing, she picks a **default** (e.g. no purchase).

Definition

p has a **random consideration set representation** $(\succ, \eta)$ if for every menu A and $x \in A$

$$p(x \mid A) = \eta(x) \prod_{y \in A:\ y \succ x} (1 - \eta(y)),$$

$$p(\text{default} \mid A) = \prod_{y \in A} (1 - \eta(y)).$$

The random consideration set model

Assumptions

(Manzini and Mariotti, 2014)

- a strict preference $\succ$ on X;
- each x is considered with probability $\eta(x) \in (0,1)$, **independently** across alternatives and **whatever the menu**;
- the DM picks the $\succ$-best considered alternative; if she considers nothing, she picks a **default** (e.g. no purchase).

Definition

p has a **random consideration set representation** $(\succ, \eta)$ if for every menu A and $x \in A$

$$p(x \mid A) = \eta(x) \prod_{y \in A:\ y \succ x} (1 - \eta(y)),$$

$$p(\text{default} \mid A) = \prod_{y \in A} (1 - \eta(y)).$$

$\eta(x)$: x is considered. $\prod_{y \succ x}(1 - \eta(y))$: no better alternative is considered.

Worked example

Parameters

$a \succ b \succ c$, with $\eta(a) = 0.5$, $\eta(b) = 0.4$, $\eta(c) = 0.8$.

Worked example

Parameters

$a \succ b \succ c$, with $\eta(a) = 0.5$, $\eta(b) = 0.4$, $\eta(c) = 0.8$.

Choice probabilities

Menu	$p(a \mid A)$	$p(b \mid A)$	$p(c \mid A)$	default
$\{a, b, c\}$	0.5	$0.4 \times 0.5 = 0.20$	$0.8 \times 0.5 \times 0.6 = 0.24$	$0.5 \times 0.6 \times 0.2 = 0.06$
$\{b, c\}$	–	0.40	$0.8 \times 0.6 = 0.48$	$0.6 \times 0.2 = 0.12$
$\{c\}$	–	–	0.80	0.20

Worked example

Parameters

$a \succ b \succ c$, with $\eta(a) = 0.5$, $\eta(b) = 0.4$, $\eta(c) = 0.8$.

Choice probabilities

Menu	$p(a \mid A)$	$p(b \mid A)$	$p(c \mid A)$	default
$\{a, b, c\}$	0.5	$0.4 \times 0.5 = 0.20$	$0.8 \times 0.5 \times 0.6 = 0.24$	$0.5 \times 0.6 \times 0.2 = 0.06$
$\{b, c\}$	–	0.40	$0.8 \times 0.6 = 0.48$	$0.6 \times 0.2 = 0.12$
$\{c\}$	–	–	0.80	0.20

The most **attention-grabbing** option c is chosen often, although it is the worst.

Identification

$\eta(c) = p(c \mid \{c\})$; and $b \succ c$ because adding b lowers c's probability ($0.48 < 0.80$) while adding c leaves b's unchanged ($0.40 = \eta(b)$).

Do we like the random consideration set model?

Pros

- A simple, closed-form link between attention η, preference $\succ$ and choice probabilities.
- Separates **taste** ($\succ$) from **attention** (η); both are identified.
- A tractable alternative (or complement) to random utility models (Lecture 3).

Cons and caveats

- Independence and menu-independence of η are **strong** assumptions.
- **Regularity:** adding an alternative never raises $p(x \mid A)$, as in every RUM (Lecture 3). Context effects (attraction, compromise) need menu-dependent or correlated attention.

The random attention model (Cattaneo et al., 2020)

Random attention model (RAM)

The **random attention model** (RAM) drops independence.

- An **attention rule** μ: $\mu(B \mid A)$ = probability that the consideration set is B, for nonempty $B \subseteq A$.
- With a strict preference $\succ$:

$$p(x \mid A) = \sum_{B \subseteq A} \mathbf{1}\{x \text{ is the } \succ\text{-best element of } B\}\, \mu(B \mid A).$$

The random attention model (Cattaneo et al., 2020)

Random attention model (RAM)

The **random attention model** (RAM) drops independence.

- An **attention rule** μ: $\mu(B \mid A)$ = probability that the consideration set is B, for nonempty $B \subseteq A$.
- With a strict preference $\succ$:

$$p(x \mid A) = \sum_{B \subseteq A} \mathbf{1}\{x \text{ is the } \succ\text{-best element of } B\}\, \mu(B \mid A).$$

Without further restrictions, this model is vacuous: set $\mu(\{x\} \mid A) = p(x \mid A)$.

The random attention model (Cattaneo et al., 2020)

Random attention model (RAM)

The **random attention model** (RAM) drops independence.

- An **attention rule** μ: $\mu(B \mid A)$ = probability that the consideration set is B, for nonempty $B \subseteq A$.
- With a strict preference $\succ$:

$$p(x \mid A) = \sum_{B \subseteq A} \mathbf{1}\{x \text{ is the } \succ\text{-best element of } B\}\, \mu(B \mid A).$$

Without further restrictions, this model is vacuous: set $\mu(\{x\} \mid A) = p(x \mid A)$.

Monotonic attention

For every $x \in A$ and $B \subseteq A \setminus \{x\}$,
$\mu(B \mid A) \leq \mu(B \mid A \setminus \{x\})$.
Removing an alternative cannot make a consideration set that does not contain it less likely.

Revealed preference in the random attention model

Revealed preference

If $p(x \mid A) > p(x \mid A \setminus \{y\})$, then $x \succ y$.
Why: removing y can only raise the probability of consideration sets without y. If x's probability falls, x must win some consideration sets that contain y.

Revealed preference in the random attention model

Revealed preference

If $p(x \mid A) > p(x \mid A \setminus \{y\})$, then $x \succ y$.
Why: removing y can only raise the probability of consideration sets without y. If x's probability falls, x must win some consideration sets that contain y.

Theorem (Cattaneo et al., 2020)

p has a RAM representation if and only if this revealed preference relation is acyclic.

Revealed preference in the random attention model

Revealed preference

If $p(x \mid A) > p(x \mid A \setminus \{y\})$, then $x \succ y$.
Why: removing y can only raise the probability of consideration sets without y. If x's probability falls, x must win some consideration sets that contain y.

Theorem (Cattaneo et al., 2020)

p has a RAM representation if and only if this revealed preference relation is acyclic.

Example (stochastic decoy)

$p(T \mid \{T, K\}) = 0.5$ and $p(T \mid \{T, K, D\}) = 0.6$ for target T, competitor K, decoy D. Then $T \succ D$; regularity fails, so no RUM (Lecture 3) explains the data, yet a RAM does.

Recap: stochastic consideration sets

Randomness in choice can come from **random attention** instead of random preferences.

Random consideration sets

Manzini and Mariotti (2014): independent attention $\eta(x)$; closed-form choice probabilities; η and $\succ$ identified; satisfies regularity.

Random attention model

Cattaneo et al. (2020): any attention rule with monotonic attention; preference revealed when removing y lowers x's choice probability; can explain the attraction effect.

5

Application: consideration sets and marketing

When consumers look at a new product only after a campaign, firms compete for attention as well as for quality.

Marketing aims at consideration

Examples from Eliaz and Spiegler (2011):

Advertising content

An ad that highlights

- a drawback of a group of rival products, or
- one of the good features of the firm's own product

may attract the attention of consumers who currently buy the rival products.

Argumentation by a salesperson

A consumer enters a car dealership with a particular model in mind; the salesperson tries to convince him to **consider** a different model.

Competitive marketing: the setting

Eliaz and Spiegler (2011) model competition between firms when consumers use consideration sets.

Simplified version: firms

Each firm i chooses a pair (x^i, m^i):

- **quality** $x^i \in \{0, 1\}$;
- **marketing** $m^i \in \{0^m, 1^m\}$, where 0^m = no campaign, 1^m = campaign.

A consumer

has a **default** product (x^s, m^s) (her status quo) and meets a **new** product (x^n, m^n).

Consideration-driven choice

Two steps

Choice depends on a **consideration function** $\phi(x^s, m^n) \in \{0, 1\}$ and proceeds in two steps:

1. **The consideration set is formed:**

$$E = \begin{cases} \{x^s, x^n\} & \text{if } \phi(x^s, m^n) = 1, \\ \{x^s\} & \text{if } \phi(x^s, m^n) = 0. \end{cases}$$

2. The consumer chooses the **highest-quality** product in E; if qualities are equal she keeps her default.

Consideration-driven choice

Two steps

Choice depends on a **consideration function** $\phi(x^s, m^n) \in \{0, 1\}$ and proceeds in two steps:

1. **The consideration set is formed:**

$$E = \begin{cases} \{x^s, x^n\} & \text{if } \phi(x^s, m^n) = 1, \\ \{x^s\} & \text{if } \phi(x^s, m^n) = 0. \end{cases}$$

2. The consumer chooses the **highest-quality** product in E; if qualities are equal she keeps her default.

Why does ϕ depend only on x^s and m^n?

- A consumer may look around if her default quality x^s is **low**;
- or if the new product has a **strong campaign** m^n.

Consideration-driven choice

Two steps

Choice depends on a **consideration function** $\phi(x^s, m^n) \in \{0, 1\}$ and proceeds in two steps:

1. **The consideration set is formed:**

$$E = \begin{cases} \{x^s, x^n\} & \text{if } \phi(x^s, m^n) = 1, \\ \{x^s\} & \text{if } \phi(x^s, m^n) = 0. \end{cases}$$

2. The consumer chooses the **highest-quality** product in E; if qualities are equal she keeps her default.

Why does ϕ depend only on x^s and m^n?

- A consumer may look around if her default quality x^s is **low**;
- or if the new product has a **strong campaign** m^n.

Here: $\phi(x^s, 1^m) = 1$ and $\phi(x^s, 0^m) = 0$: only a campaign makes the consumer consider the new product.

The market

Setup

- **Two identical firms** $i = 1, 2$ choose (x^i, m^i) simultaneously.
- **Consumers** differ only in their default: for half of them firm 1 is the default, for the other half firm 2.
- **Fixed costs:** (x, m) costs $c_x + c_m$, with $c_1 > c_0 \geq 0$, $c_{0^m} = 0 < c_{1^m}$, and $c_1 - c_0 + c_{1^m} < \frac{1}{2}$.

The market

Setup

- **Two identical firms** $i = 1, 2$ choose (x^i, m^i) simultaneously.
- **Consumers** differ only in their default: for half of them firm 1 is the default, for the other half firm 2.
- **Fixed costs:** (x, m) costs $c_x + c_m$, with $c_1 > c_0 \geq 0$, $c_{0^m} = 0 < c_{1^m}$, and $c_1 - c_0 + c_{1^m} < \frac{1}{2}$.

Profits of firm i against j (market shares minus costs)

$$\pi_i = \begin{cases} \frac{1}{2}\left[1 + \phi(x^j, m^i)\right] - c_{x^i} - c_{m^i} & \text{if } x^i > x^j \quad \text{(wins } j\text{'s defaults if they look)}, \\ \frac{1}{2}\left[1 - \phi(x^i, m^j)\right] - c_{x^i} - c_{m^i} & \text{if } x^i < x^j \quad \text{(loses its defaults if they look)}, \\ \frac{1}{2} - c_{x^i} - c_{m^i} & \text{if } x^i = x^j. \end{cases}$$

The payoff matrix

Notation

Write $L = (0, 0^m)$, $L^A = (0, 1^m)$, $H = (1, 0^m)$, $H^A = (1, 1^m)$; $\Delta = c_1 - c_0$, $a = c_{1^m}$; net of c_0.

Payoff of firm i

Firm i \ firm j	L	L^A	H	H^A
L	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	0
L^A	$\frac{1}{2} - a$	$\frac{1}{2} - a$	$\frac{1}{2} - a$	$-a$
H	$\frac{1}{2} - \Delta$	$\frac{1}{2} - \Delta$	$\frac{1}{2} - \Delta$	$\frac{1}{2} - \Delta$
H^A	$1 - \Delta - a$	$1 - \Delta - a$	$\frac{1}{2} - \Delta - a$	$\frac{1}{2} - \Delta - a$

The payoff matrix

Notation

Write $L = (0, 0^m)$, $L^A = (0, 1^m)$, $H = (1, 0^m)$, $H^A = (1, 1^m)$; $\Delta = c_1 - c_0$, $a = c_{1^m}$; net of c_0.

Payoff of firm i

Firm i \ firm j	L	L^A	H	H^A
L	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	0
L^A	$\frac{1}{2} - a$	$\frac{1}{2} - a$	$\frac{1}{2} - a$	$-a$
H	$\frac{1}{2} - \Delta$	$\frac{1}{2} - \Delta$	$\frac{1}{2} - \Delta$	$\frac{1}{2} - \Delta$
H^A	$1 - \Delta - a$	$1 - \Delta - a$	$\frac{1}{2} - \Delta - a$	$\frac{1}{2} - \Delta - a$

Reading the matrix

- L^A is **strictly dominated** by L: advertising a low-quality product never pays.
- H guarantees $\frac{1}{2} - \Delta$; H^A pays off only against low quality; L is hurt only by H^A.

No symmetric pure equilibrium: against L, play H^A; against H^A, play H; against H, play L.

Mixed strategies and equilibrium

Definitions

- A **mixed strategy** σ assigns a probability $\sigma(x, m)$ to each $(x, m) \in S = \{0, 1\} \times \{0^m, 1^m\}$.
- **Expected payoff:**

$$\Pi_i(\sigma_i, \sigma_j) = \sum_{s, s' \in S} \sigma_i(s)\, \sigma_j(s')\, \pi_i(s, s').$$

- **Nash equilibrium:** $\Pi_i(\sigma_i, \sigma_j) \geq \Pi_i(\sigma, \sigma_j)$ for every σ and $i = 1, 2$.

Mixed strategies and equilibrium

Definitions

- A **mixed strategy** σ assigns a probability $\sigma(x, m)$ to each $(x, m) \in S = \{0, 1\} \times \{0^m, 1^m\}$.
- **Expected payoff:**

$$\Pi_i(\sigma_i, \sigma_j) = \sum_{s, s' \in S} \sigma_i(s)\, \sigma_j(s')\, \pi_i(s, s').$$

- **Nash equilibrium:** $\Pi_i(\sigma_i, \sigma_j) \geq \Pi_i(\sigma, \sigma_j)$ for every σ and $i = 1, 2$.

Solving for a symmetric mixed equilibrium

With probabilities $(\sigma_L, \sigma_H, \sigma_{H^A})$:

- H and H^A indifferent:
  $\frac{1}{2} - \Delta = \frac{1}{2} - \Delta - a + \frac{1}{2}\sigma_L \;\Rightarrow\; \sigma_L = 2a.$

Mixed strategies and equilibrium

Definitions

- A **mixed strategy** σ assigns a probability $\sigma(x, m)$ to each $(x, m) \in S = \{0, 1\} \times \{0^m, 1^m\}$.
- **Expected payoff:**

$$\Pi_i(\sigma_i, \sigma_j) = \sum_{s, s' \in S} \sigma_i(s)\, \sigma_j(s')\, \pi_i(s, s').$$

- **Nash equilibrium:** $\Pi_i(\sigma_i, \sigma_j) \geq \Pi_i(\sigma, \sigma_j)$ for every σ and $i = 1, 2$.

Solving for a symmetric mixed equilibrium

With probabilities $(\sigma_L, \sigma_H, \sigma_{H^A})$:

- H and H^A indifferent:
  $\frac{1}{2} - \Delta = \frac{1}{2} - \Delta - a + \frac{1}{2}\sigma_L \;\Rightarrow\; \sigma_L = 2a.$
- L and H indifferent:
  $\frac{1}{2}(1 - \sigma_{H^A}) = \frac{1}{2} - \Delta \;\Rightarrow\; \sigma_{H^A} = 2\Delta.$

The equilibrium

Proposition 1

There is a unique symmetric Nash equilibrium:

$$\sigma(0, 0^m) = 2c_{1^m}, \qquad \sigma(1, 1^m) = 2(c_1 - c_0), \qquad \sigma(1, 0^m) = 1 - 2(c_1 - c_0 + c_{1^m}).$$

The equilibrium

Proposition 1

There is a unique symmetric Nash equilibrium:

$$\sigma(0,0^m) = 2c_{1^m}, \qquad \sigma(1,1^m) = 2(c_1 - c_0), \qquad \sigma(1,0^m) = 1 - 2(c_1 - c_0 + c_{1^m}).$$

Example

$c_1 - c_0 = 0.10$, $c_{1^m} = 0.15 \Rightarrow \sigma_L = 0.3$, $\sigma_H = 0.5$, $\sigma_{H^A} = 0.2$.

- Each strategy in the support earns 0.4 (net of c_0): e.g. $\Pi(L) = \frac{1}{2}(1 - 0.2) = 0.4$.
- A consumer keeps a low-quality default although the rival offers high quality with probability $0.3 \times 0.5 = 0.15$.

Interpretation

Four properties of the equilibrium

1. The equilibrium is **mixed**: consumers end up with an **inferior** product with positive probability.

Interpretation

Four properties of the equilibrium

1. The equilibrium is **mixed**: consumers end up with an **inferior** product with positive probability.
2. Firms **advertise** with positive probability, $2(c_1 - c_0)$.

Interpretation

Four properties of the equilibrium

1. The equilibrium is **mixed**: consumers end up with an **inferior** product with positive probability.
2. Firms **advertise** with positive probability, $2(c_1 - c_0)$.
3. **Advertising signals quality:** the only product that is advertised in equilibrium is the high-quality one.

Interpretation

Four properties of the equilibrium

1. The equilibrium is **mixed**: consumers end up with an **inferior** product with positive probability.
2. Firms **advertise** with positive probability, $2(c_1 - c_0)$.
3. **Advertising signals quality:** the only product that is advertised in equilibrium is the high-quality one.
4. Cheaper advertising ($c_{1^m} \downarrow$) means **less** low quality in the market ($\sigma_L = 2c_{1^m}$).

Interpretation

Four properties of the equilibrium

1. The equilibrium is **mixed**: consumers end up with an **inferior** product with positive probability.
2. Firms **advertise** with positive probability, $2(c_1 - c_0)$.
3. **Advertising signals quality:** the only product that is advertised in equilibrium is the high-quality one.
4. Cheaper advertising ($c_{1^m} \downarrow$) means **less** low quality in the market ($\sigma_L = 2c_{1^m}$).

Interpretation

Four properties of the equilibrium

1. The equilibrium is **mixed**: consumers end up with an **inferior** product with positive probability.
2. Firms **advertise** with positive probability, $2(c_1 - c_0)$.
3. **Advertising signals quality:** the only product that is advertised in equilibrium is the high-quality one.
4. Cheaper advertising ($c_{1^m} \downarrow$) means **less** low quality in the market ($\sigma_L = 2c_{1^m}$).

When consumers consider alternatives only after a campaign, competition is for **attention**, not only for quality: firms randomize, and some consumers stay with inferior products.

6

Wrapping up

A recap of the models of today, exercises, and questions for discussion.

Recap

Consideration sets

- The DM maximizes $\succ$ (or u) over **the options she considers**, not over the whole menu.
- Without restrictions on consideration, the model is **empty**; each restriction gives testable content:
  - **shortlists** (Expansion, Weak WARP);
  - **attention filters** (acyclic revealed preference P, WARP with limited attention);
  - **random consideration** (closed form, regularity) and **random attention** (monotonic attention).

Revealed preference

Preference is revealed by **WARP violations**: when removing y changes the choice of x.

In markets, firms compete for **attention**: advertising and quality are mixed and correlated.

Exercises (1/2)

Exercise 1 · Easy

Random consideration sets with $a \succ b \succ c$ and $\eta(a) = 0.6$, $\eta(b) = \eta(c) = 0.5$. Compute $p(\cdot \mid \{a, b, c\})$, $p(\cdot \mid \{b, c\})$ and the probability of the default in each menu.

Exercise 2 · Intermediate

Compromise effect. Let $x = (1, 3)$, $y = (2, 2)$, $z = (3, 1)$ with $C(\{x, y\}) = x$ and $C(\{x, y, z\}) = y$. Find a CLA representation. What does P reveal?

Exercises (2/2)

Exercise 3 · Intermediate

In the marketing game, let $c_1 - c_0 = 0.05$ and $c_{1^m} = 0.2$. Compute the equilibrium, the equilibrium profit, and the probability that a consumer buys low quality although high quality is on the market. What happens to the share of low-quality products if advertising becomes more expensive?

Exercise 4 · Advanced

(a) Show that every rational shortlist method satisfies Expansion. (b) Construct an RSM on $\{x, y, z\}$ with a **cycle** in binary choices: $C(\{x, y\}) = x$, $C(\{y, z\}) = y$, $C(\{x, z\}) = z$.

Questions for discussion

Rationality

Do bounded rationality models explain **why** people deviate from rationality, or do they offer a **richer account** of rationality?

Applications

Think of an application of consideration set models (labour markets, politics, online platforms. . .).

Mechanisms

Which **mechanisms** could explain how consideration sets are formed?

Measurement

How could we **measure** consideration sets directly, instead of inferring them from choices? (Hint: Lecture 5's choice process and search data.)

Final take-aways

1. A consideration set model separates **what you like** ($\succ$) from **what you look at** (E).
2. It has empirical content only if consideration is **structured**: shortlists, attention filters, monotonic random attention.
3. Preferences are then revealed by **WARP violations**, not by every choice.
4. In markets, firms compete for **attention**, and some consumers end up with inferior products.

References I

Cattaneo, M. D., Ma, X., Masatlioglu, Y., and Suleymanov, E. (2020). A random attention model. *Journal of Political Economy*, 128(7):2796–2836.

Eliaz, K. and Spiegler, R. (2011). Consideration sets and competitive marketing. *The Review of Economic Studies*, 78(1):235–262.

Kahneman, D. (2011). *Thinking, Fast and Slow*. Farrar, Straus and Giroux, New York.

Lleras, J. S., Masatlioglu, Y., Nakajima, D., and Ozbay, E. Y. (2017). When more is less: Limited consideration. *Journal of Economic Theory*, 170:70–85.

Manzini, P. and Mariotti, M. (2007). Sequentially rationalizable choice. *American Economic Review*, 97(5):1824–1839.

Manzini, P. and Mariotti, M. (2014). Stochastic choice and consideration sets. *Econometrica*, 82(3):1153–1176.

Masatlioglu, Y., Nakajima, D., and Ozbay, E. Y. (2012). Revealed attention. *American Economic Review*, 102(5):2183–2205.

Tversky, A. and Shafir, E. (1992). Choice under conflict: The dynamics of deferred decision. *Psychological science*, 3(6):358–361.