Limited Consideration and Stochastic Choice: Theory and Evidence

Victor Aguiar[1] María José Boccardi[2] Nail Kashaev[1] Jeongbin Kim[3]

[1]University of Western Ontario

[2]Amazon

[3]NUS

Bristol-Warwick

2022

Motivation and Contribution

- The Random Utility Model (RUM) is the standard to describe population behavior: Heterogeneous preferences + Full Consideration.
- Decision makers (DMs) may not consider all alternatives. We extend RUM to a framework with **limited consideration**: Random Attention and Utility Model (RAUM).
- RAUM is a general and testable framework. Possible curse of dimensionality. One solution: restriction to attention-indexes.
- We run an experiment and do statistical testing for RAUM consistency with the attention-index restriction.
- **Preview of the Findings**: Reject RUM/ Fail to reject a RAUM that is an attention-index.

Our Model: Random Attention and Utility Model (RAUM)

- Dataset $P = (p(a, A))_{A \in \mathcal{A}}$ describing a population of DMs (vectorized p). $A \in \mathcal{A}$: $\emptyset \subset A \subseteq X$ with X finite.
- P has RAUM representation if:

$$p(a, A) = \sum_{(\succ, \phi) \in U \times \phi} \boldsymbol{\pi(\succ, \phi)} 1(a \succ b, \ \forall b \in \phi(A)),$$

for some distribution over preferences and consideration filters $\pi \in \Delta(U \times \Phi)$.

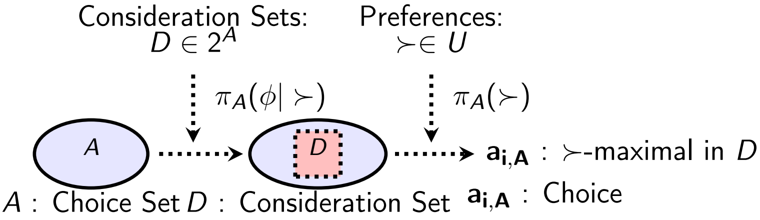
- U is the set of strict preferences. Φ is thes set of filters.
- For a fixed D, $\phi(A) = D$ when $D \subseteq A$ and $\phi(A) = \emptyset$ otherwise.

Lemma

Every P is a RAUM.

Our Model: Sources of Randomness

- Each DM is endowed with a preference $\succ$ and a random consideration rule $\pi_A(\phi|\succ)$.



- Sleeping agent: $X = \{a, b, c\}$, $b \succ^1 a \succ^1 c$, $a \succ^2 b \succ^2 c$, $\phi^1(X) = \{a, b, c\}$, $\phi^2(X) = \{a\}$. $\pi_X(\phi^1|\succ^1) = 1$, $\pi_X(\phi^2|\succ^2) = 1$. $\pi_X(\succ^1) = \frac{1}{2}$.
- An inattentive DM $p^*_\succ(a, A) = \sum_{\phi\in\Phi} \pi_A(\phi|\succ)1(a \succ b \forall b \in \phi(A))$. RAUM is a mixture of inattentive DMs.

Set-monotone and stable RAUM

- *Stability*: $\pi_A(\succ) = \pi_B(\succ)$ for all A, B and $\succ$, with $\pi_A(\succ) = \sum_{\phi \in \Phi} \pi(\succ, \phi)$.
- *Set-monotonicity*: For $A \subseteq B$, $\pi_A(\phi | \succ) \geq \pi_B(\phi | \succ)$ with $\phi(A) \neq \emptyset$.
- Not testable: Stable RAUM/ Set-monotone RAUM.
- Stable and testable RAUM is testable!
- Examples: Search and Satisfy, Rational Inattention, and *attention-index* models.

Theorem

TFAE:

- *P admits a set-monotone and stable RAUM representation.*
- *There exists a $\nu \in \mathbb{R}^{d_g}$, such that $g = G\nu$ (with a known matrix G of $1, 0, -1$, and $g = (p', 1_{dm}, 0_{dr})$, with $d_g = d_\rho + d_m + d_r$).*

- We generalize RUM ([McFadden and Richter, 1990]) and Random Attention Model (RAM, [Cattaneo et al., 2017]). Testing Kitamura and Stoye [2018].

A special case: Attention-index/Link (L-RAUM)

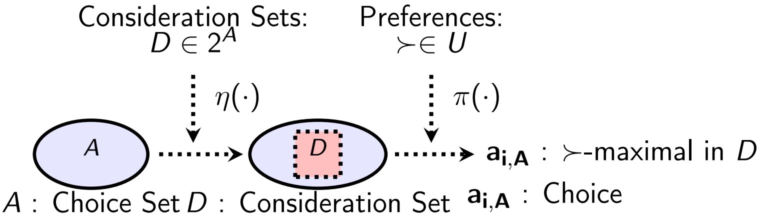
- P has an attention-index representation:

$$p(a, A) = \sum_{\succ \in U} \boldsymbol{\pi(\succ)} \sum_{D \subseteq A} \boldsymbol{m_A(D)} 1(a \succ b, \ \forall b \in D),$$

for some distribution over preferences $\pi \in \Delta(U)$ and some distribution over consideration sets $m_A \in \Delta(2^A)$.

- Attention index: $\eta : 2^X \to [0, 1]$. Analogous to utility but over menus.
- $m_A(D) = \psi(\eta(D), \sum_{C \in g(D,A)} \eta(C))$, where ψ is a link-function and g is an index. (Examples coming)
- This framework is stable and set-monotone (monotone ψ).
- Additional assumptions: (i) Independence: Preferences $\perp$ Consideration. (ii) (default) $o \in A$ such that $m_A(D) = 0$ if $o \notin D$.

Our Model: Sources of Randomness



- Two-stage system: First attention is determined, then DMs use preferences to decide.
- Simplify then choose.

Representative Agent: Optimal Consideration

- A representative DM with random utility with mean utility $u : X \to \mathbb{R}$, such that $u(x) + \xi_x$.
- McFadden's surplus: $\alpha(D) = \mathbb{E}[\max_{x \in D} u(x) + \xi_x]$.
- $m_A = argmax_{m \in \Delta(2^A)} \sum_{D \subseteq A} [m(D)\alpha(D) - K(m(D))]$.
- $K(t) = -t \log(t)/\theta$ then $m_A(D) = \frac{exp(\theta\alpha(D))}{\sum_{C \subseteq A} exp(\theta\alpha(C))}$.
- $\eta(D) = exp(\theta\alpha(D))$, and $\psi(t, s) = \frac{t}{s}$.
- Reduction of dimensionality from $\sum_{A \subseteq X} 2^{|A|} - 1$ to $2^{|X|} - 1$.

Attention index: Survival Race

- $\eta \in \Delta(2^X)$, attention index. $\eta(A)$ captures how attractive is A.
- LA-Logit Attention (Brady and Rehbeck 2016)

$$m_A^{LA}(D) = \frac{\eta(D)}{\sum_{C \subseteq A} \eta(C)}$$

- MM-Manzini-Mariotti (2014)/Independent attention

$$m_X^{MM}(D) = \eta(D) = \prod_{a \in D} \gamma(a) \prod_{a \in X \setminus D} (1 - \gamma(a))$$

- EBA Elimination-by-Aspects (Tversky, 1972)

$$m_A^{EBA}(D) = \sum_{C : C \cap A = D} \eta(C)$$

Underlying Full Consideration RUM

- $p_\pi(a, C) = \sum_{\succ \in U} \pi(\succ) 1(a \succ b \forall b \in C)$ (RUM/FC).
- We can re-write the L-RAUM as:

$$p(a, A) = \sum_{C \subseteq A} m_A(C) \boldsymbol{p_\pi(a, C)},$$

- Our main theoretical result is a decomposition of P into m_A, p_π.
- Note that:

$$p(o, A) = 1 - \sum_{C \subseteq A : C \neq \emptyset} m_A(C).$$

It does not depend on p_π: key assu. preferences and consideration are independent. We can identify $\eta^{\mathbf{L}}$ from $\{p(o, A)\}_{A \in \mathcal{A}}$.

Identification of the distribution of consideration

- If P admits a L-RAUM representation then: $m_A^L(C) = F^L(\{p(o, A \cup o)\}_{A \subseteq X})$.
- We recover m_A^L uniquely for $L \in \{LA, MM, EBA\}$.

L	m_A^L	η^L
$\mathcal{M}^{LA}$	$m_A^{LA}(D) = \frac{\eta(D)}{\sum_{C \subseteq A} \eta(C)} > 0$	$\eta^{LA}(D) = \sum_{B \subseteq D} (-1)^{\|D \setminus B\|} \frac{p(o,X)}{p(o,B)}$
$\mathcal{M}^{MM}$	$\eta(\cdot) = \prod_{a \in X \setminus \cdot} (1 - \gamma(a)) \prod_{b \in \cdot} \gamma(b)$	$\gamma^{MM}(a) = 1 - \frac{p(o,A)}{p(o,A \setminus \{a\})}$
$\mathcal{M}^{EBA}$	$m_A(D) = \sum_{C : C \cap A = D} \eta(C)$	$\eta^{EBA}(D) = \sum_{A \subseteq D : D \in \mathcal{A}} (-1)^{\|D \setminus A\|} p(o, X \setminus A)$

Identification of the distribution of preferences

- Once we have identified m_A^L, we can identify p_π^L uniquely:
- $p_\pi^L(a, A) = \frac{p(a,A) - \sum_{C \subset A} m_A^L(C) p_\pi^L(a,C)}{\mathbf{m_A^L(A)}}$ ($P_\pi^L = \{p_\pi^L(a, A)\}_{a \in A, A \in \mathcal{A}}$).

Theorem

Suppose that for given $L \in \{LA, MM, EBA\}$ and P, (i) m^L is well-defined, (ii) $m_A^L(A) > 0$ for all $A \in \mathcal{A}$. Then the following are equivalent.

1. *P is a L-RAUM-rule;*
2. *P_π^L is a FC-RAUM-rule (i.e., RUM).*

- We have reduced our problem to testing RUM! (Kitamura and Stoye 2018).
- m_A, p_π are identified uniquely!

Testing Procedure (I)

Theorem (Testing)

The following are equivalent.

1. P^L *is RUM and* m^L *is well-defined;*
2. $\inf_{v \in \mathbb{R}_+^d} \|g^L - Gv\| = 0$*, where* $g^L = (P^{L'}, m^{L'})'$.

$$G = \begin{bmatrix} B & 0_{d_p \times d_m} \\ 0_{d_m \times \|X\|!} & I_{d_m} \end{bmatrix}, \qquad B_{k,l} = 1(a \in A)1(a \succ_l c, : \forall : c \in A)$$

We can apply [Kitamura and Stoye, 2018].

Testing Procedure (I)

Theorem (Testing)

The following are equivalent.

1. *P^L is RUM and m^L is well-defined;*
2. $\inf_{v \in \mathbb{R}^d_+} \|g^L - Gv\| = 0$, *where* $g^L = (P^{L'}, m^{L'})'$.

$$G = \begin{bmatrix} B & 0_{d_p \times d_m} \\ 0_{d_m \times \|X\|!} & I_{d_m} \end{bmatrix}, \qquad B_{k,l} = 1(a \in A)1(a \succ_l c, : \forall : c \in A)$$

We can apply [Kitamura and Stoye, 2018].

Testing Procedure (I)

Theorem (Testing)

The following are equivalent.

1. P^L *is RUM and* m^L *is well-defined;*
2. $\inf_{v \in \mathbb{R}^d_+} \|g^L - Gv\| = 0$, *where* $g^L = (P^{L'}, m^{L'})'$.

$$G = \begin{bmatrix} B & 0_{d_p \times d_m} \\ 0_{d_m \times \|X\|!} & I_{d_m} \end{bmatrix}, \qquad B_{k,l} = 1(a \in A)1(a \succ_l c, : \forall : c \in A)$$

We can apply [Kitamura and Stoye, 2018].

Testing Procedure (I)

Theorem (Testing)

The following are equivalent.

1. *P^L is RUM and m^L is well-defined;*
2. $\inf_{v \in \mathbb{R}^d_+} \|g^L - Gv\| = 0$, *where* $g^L = (P^{L'}, m^{L'})'$.

$$G = \begin{bmatrix} B & 0_{d_p \times d_m} \\ 0_{d_m \times \|X\|!} & I_{d_m} \end{bmatrix}, \qquad B_{k,l} = 1(a \in A)1(a \succ_l c, : \forall : c \in A)$$

We can apply [Kitamura and Stoye, 2018].

Testing Procedure (II)

Given testing theorem, a natural test statistic is

$$T_n = n \min_{[v - \tau_n \iota / d] \in \mathbb{R}^d_+} (\hat{g}^L - Gv)'(\hat{g}^L - Gv)$$

where $\hat{g}^L = (\hat{P}^{L'}, \hat{m}^{L'})'$; τ_n is a tuning parameter; and ι is a vector of ones of dimension d.

Testing Procedure (III)

Let $\hat{g}_l^{L,*}$, $l = 1, \ldots, L$ be bootstrap replications of $\hat{g}^L$. To compute CV of T_n we follow the bootstrap procedure proposed in [Kitamura and Stoye, 2018]:

1. Compute $\hat{\eta}_{\tau_n} = Gv_{\tau_n}$, where v_{τ_n} solves

$$n \min_{[v - \tau_n \iota / d] \in \mathbb{R}_+^d} (\hat{g}^L - Gv)'(\hat{g}^L - Gv);$$

2. Compute

$$\hat{g}_l^{L,*} = \hat{g}_l^{L,*} - \hat{g}^T + \hat{\eta}_{\tau_n},$$

3. Compute the bootstrap test statistics

$$T_{n,l}^* = n \min_{[v - \tau_n \iota / d] \in \mathbb{R}_+^d} (\hat{g}_l^{L,*} - Gv)'(\hat{g}_l^{L,*} - Gv), \quad l = 1, \ldots, L;$$

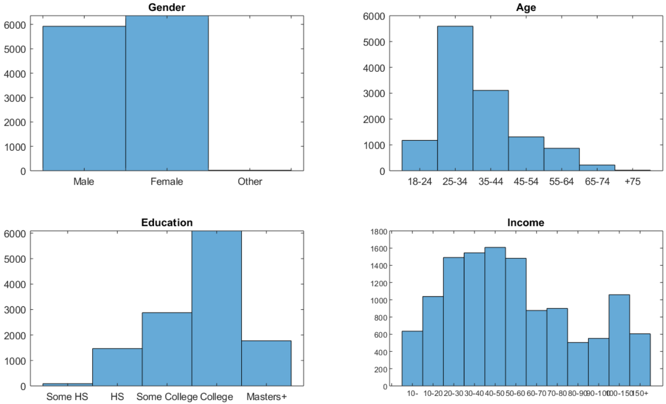
4. Use the empirical distribution of the bootstrap statistic to compute critical values of T_n. Reject if $T_n > \hat{c_{1-\alpha}}$, $\alpha \in (0, 1/2)$ significance level.

The Experiment

- We design and implement an (mTurk) experiment to collect a **standard stochastic choice dataset** with a default from a **large** sample of individuals.
- We require full choice set variation, but no requirement to have repeated individual choices. We vary exogenously the cost of consideration and the choice set.
- Our experiment: 12297 choices from 2135 individuals (3 levels of complexity/ 3 frames).

Sample

- Run between Aug 25th and Sept 17th 2018 on Amazon MTurk
- 2135 individuals with 12,297 independent decisions
- Avg. payment \$1.09 (\$ 0.25 participation fee)
- Avg duration 251.68 secs - Hourly rate $\sim$ \$15



Design - Consideration Cost H/M/L

Select (only) one of the following alternatives. Click over the selected option and hit next [-->].

Get 12 tokens for sure.

Prize	Probability
(3-8+22+4 +9) tokens	50%
(5-7 + 41 +10 -39) tokens	50%

Prize	Probability
(27-10 -2 +11 +24) tokens	25%
(25+6-1+3+15) tokens	20%
(1-18 +3 +17 +11) tokens	15%
(9-6 -7 -4 + 8) tokens	40%

Design - Consideration Cost H/M/L

Select (only) one of the following alternatives. Click over the selected option and hit next [->].

Get 12 tokens for sure.

Prize	Probability
(18+2+10) tokens	50%
(4+5+1) tokens	50%

Prize	Probability
(66+5-21) tokens	25%
(28-24+44) tokens	20%
(10-3+7) tokens	15%
(2-5+3) tokens	40%

Design - Consideration Cost H/M/L

Select (only) one of the following alternatives. Click over the selected option and hit next [-->].

Get 12 tokens for sure.

Prize	Probability
50 tokens	25%
48 tokens	20%
14 tokens	15%
0 tokens	40%

Prize	Probability
30 tokens	50%
10 tokens	50%

Evidence against RUM

Under the null of RUM-consistency there should be no effect of the Costly Consideration Treatments on behavior.

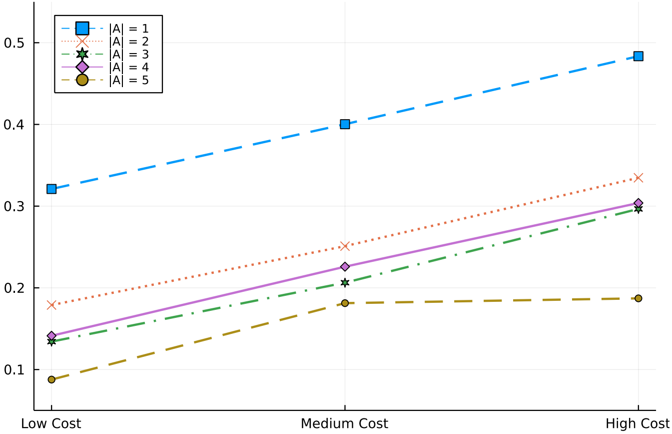


Figure: Estimated frequency of the Outside Option: Average by menu size.

RUM vs LA:RAUM

- We test RUM, LA-RAUM
- Rich menu variation increases the statistical power.
- We have 5 lotteries and one certain outcome.
- The certain outcome is always available and dominated. It is also pre-selected. (Default).
- The number of possible rankings over the choice set is 120 under LA-RAUM and 720 under RUM.
- We reject RUM. We cannot reject LA-RAUM.

Stability of Preferences across frames/cost treatments

- Unique design allows to test RUM (stability of preferences).
- RUM is description independent (implicit assumptions: *consequentialism*+ full consideration.
- *Consequentialism* means preferences do not change with descriptions/frames.
- We cannot reject LA under consequentialism. Different attention-index per frame, but same stable preference.

Model	T_n	p-value
RUM	3231.59	<0.001
LA	24959.06	0.524
EBA	24840.23	0.001

Notes: Number of bootstrap replications=1000.

Conclusion

- We extend existing econometrics tools for RUM, to allow for consideration-mediated choice: testing and identification results
- We have designed a novel experiment that allows us to discern among competing models for population behavior.
- We find that:
  1. RUM fails
  2. LA with preference heterogeneity describes behavior across all costs.
  3. Preferences are stable across costs for LA-RAUM
- [Abaluck and Adams, 2017] have used our findings to choose a semiparametric specification in their setup for insurance buying.
- Take-home message: we need to model inattention.

Abaluck, J. and Adams, A. (2017).
What do consumers consider before they choose? identification from asymmetric demand responses.
Technical report, National Bureau of Economic Research.

Cattaneo, M. D., Ma, X., Masatlioglu, Y., and Suleymanov, E. (2017).
A random attention model.
arXiv preprint arXiv:1712.03448.

Kitamura, Y. and Stoye, J. (2018).
Nonparametric analysis of random utility models.
Econometrica, 86(6):1883–1909.

McFadden, D. and Richter, M. K. (1990).
Stochastic rationality and revealed stochastic preference.
Preferences, Uncertainty, and Optimality, Essays in Honor of Leo Hurwicz, Westview Press: Boulder, CO, pages 161–186.