Bounded rationality is asymptotically rare

Alfio Giarlotta*, Angelo Petralia†, Stephen Watson‡

August 5, 2021

Abstract

We define a class of properties of choices, called hereditary finitely violated (HFV), which encompasses several declinations of bounded rationality proposed in the literature. HFV properties asymptotically fail to hold for all choices. It follows that almost all finite choices cannot be explained by most known models of bounded rationality. We provide numerical estimates confirming the rarity of bounded rationality even for relatively small sets of alternatives.

JEL Classification: D01.

1 Rationalizability and bounded rationality

According to the theory of revealed preferences pioneered by Samuelson (1938), choice is observed and preference is revealed. In this approach, rationality coincides with the possibility to justify the choice behavior of a decision maker (DM) by maximizing the binary relation of revealed preference. However, this notion of rationality, called *rationalizability*, fails to explain many observed phenomena. Following the inspiring analysis of Simon (1955), in the last twenty years rationalizability has been weakened by forms of *bounded rationality*, which aim to explain a larger portion of choices by means of more flexible paradigms.

Within this stream of research, Kalai, Rubinstein, and Spiegler (2002) describe an approach of 'multi-rationalization', which uses distinct rationales (linear orders) to explain choice behavior. According to this model of *rationalization by multiple rationales (RMR)*, selection from a menu is justified by maximizing one of the available orders, with no need to semantically link the rationale to the menu itself. By its own nature, the RMR model is *universal*: any choice is multi-rationalizable, because the number of rationales can be increased as needed to explain any type of behavior. Thus, the authors study minimal

*Department of Economics and Business, University of Catania, Italy. alfio.giarlotta@unict.it *(corresponding author)*

†University of Turin & Collegio Carlo Alberto, Turin, Italy. angelo.petralia@carloalberto.org

‡Dept. of Math. and Stat., York University, Toronto, Canada. watson@mathstat.yorku.ca

representations. Their Proposition 1 shows that any choice c on n elements requires at most $n-1$ rationales. It follows that the *degree of rationality* of c ranges from 1 (for rationalizable choices) to $n-1$ (for *chaotic choices*).[1] Furthermore, their Proposition 2 determines the asymptotic behavior of this degree of rationality: as n tends to infinity, almost all choices become chaotic.

Alternative models of bounded rationality aim to identify regularity properties entailed by a theory of choices. Without being exhaustive, we mention some of them (see also Section 2A.) In the work of Masatlioglu and Ok (2005), later elaborated by Apesteguia and Ballester (2013), the DM restricts her attention to alternatives that are superior to her *status quo*. Manzini and Mariotti (2007) propose an approach in which the DM selects from each menu the unique item that survives after the *sequential* application of distinct criteria (asymmetric relations). Xu and Zhou (2007) characterize a rationalization method which justifies the selection from any menu as the *subgame perfect Nash equilibrium* outcome of an associated extensive game. Rubinstein and Salant (2008) investigate a *post-dominance rationality* choice rule: the DM first discards any alternative which is dominated by another alternative in the menu, and then chooses the best alternative from the remaining ones. According to the choice procedure of Manzini and Mariotti (2012), the DM only considers those alternatives that belong to some salient *categories*. Masatlioglu, Nakajima, and Ozbay (2012) describe a DM *with limited attention*, who is unable to take into account all the alternatives in a menu. A similar argument is used by Lleras, Masatlioglu, Nakajima, and Ozbay (2017) to analyze *overwhelming choices*. Cherepanov, Feddersen, and Sandroni (2013) present a *theory of rationalization*, in which the DM discards items not satisfying some psychological constraint. In the work of Apesteguia and Ballester (2013) the DM's choice is guided by *routes*. Yildiz (2016) discusses a choice rule based on a pairwise comparison of items according to an ordered *list*.

The bounded rationality methods mentioned above explain choice behavior by 'sequential procedures' that appeal to different categories of tools (several binary relations, a relation and a choice correspondence, game trees, etc.). However, these methods do have a common feature: they are *non-universal*, in the sense that some choice behaviors are boundedly rationalizable, but (many) others are not. The following query arises:

Question. *As the number of items tends to infinity, what is the fraction of boundedly rationalizable choices?*

In other words: *What is the asymptotic explanatory power of non-universal models of bounded rationality?* This note answers this question for all discussed approaches (in fact, also for others). To that end, we study suitable properties of choices, which we call

[1] The expressions 'degree of rationality' and 'chaotic choice' are ours.

hereditary finitely violated (HFV). Our main result shows that HFV properties fail to hold for almost all choices. Then we derive what one may expect from the very concept of bounded rationality: non-universal models explain less and less as the size of the ground set grows larger and larger. To make our analysis more concrete, we also provide some numerical estimates that confirm the rarity of bounded rationalizability.

Our findings reinforce the belief that if an observed choice behavior is explained by a bounded rationality model, then we ought to be confident that that behavior has a sound justification. As a byproduct of our approach, we finally derive a striking analogy between universal and non-universal models: as the number of items approaches infinity, not only 'chaotic rationality' but also 'bounded irrationality' rule supreme.

2 Hereditary finitely violated properties

Let X be a nonempty finite set of options available to the decision maker. A nonempty set $A \subseteq X$ is a *menu*, and $\mathscr{X} = 2^X \setminus \{\varnothing\}$ denotes the family of all menus. A *choice correspondence* on X is a map $c \colon \mathscr{X} \to \mathscr{X}$, which selects at least one item from each menu, that is, $\varnothing \neq c(A) \subseteq A$ for any $A \in \mathscr{X}$. In particular, a *choice function* selects a unique item from each menu; thus, we identify it with a map $c \colon \mathscr{X} \to X$ such that $c(A) \in A$ for any $A \in \mathscr{X}$. Here we mostly deal with choice functions: unless unclear from context, we refer to them as *choices*.

A binary relation $\succ$ on X is a subset of X^2. It is *asymmetric* if $x \succ y$ implies $\neg(y \succ x)$ for all $x, y \in X$, *transitive* if $(x \succ y) \wedge (y \succ z)$ implies $x \succ z$ for all $x, y, z \in X$, and *complete* if either $x \succ y$ or $y \succ x$ holds for all distinct $x, y \in X$. An asymmetric, transitive, and complete binary relation is a *linear order*, often denoted by $\rhd$.

Given an asymmetric relation $\succ$ on X and a menu $A \in \mathscr{X}$, the set of *maximal* (or *non-dominated*) elements of A is $\max(A, \succ) = \{x \in X : y \succ x \text{ for no } y \in A\}$. The theory of revealed preferences pioneered by Samuelson (1938) studies when and how a binary relation justifies choice behavior by maximization. Formally, a choice $c \colon \mathscr{X} \to X$ is *rationalizable* (or *binary*) if there is an asymmetric relation (in fact, a linear order) $\succ$ on X such that, for any $A \in \mathscr{X}$, $c(A)$ is the unique element of the set $\max(A, \succ)$; in this case, we slightly abuse notation, and write $c(A) = \max(A, \succ)$. If $\succ$ is an asymmetric relation on X, and A is a nonempty subset of X, we denote by $\succ_{\restriction A}$ the restriction of $\succ$ to A, that is, for any $a, b \in A$, $a \succ_{\restriction A} b$ holds if and only if $a \succ b$. Note that if $\rhd$ is a linear order on X, then $\rhd_{\restriction A}$ is a linear order on $A \subseteq X$.

Definition 1. Let $c \colon \mathscr{X} \to X$ be a choice function. For any $A \in \mathscr{X}$, let $\mathscr{A}$ be the family of nonempty subsets of A. The choice *induced by c on A* is $c_{\restriction A} \colon \mathscr{A} \to A$, defined by $c_{\restriction A}(B) = c(B)$ for any $B \in \mathscr{A}$. Given two choices $c \colon \mathscr{X} \to X$ and $c' \colon \mathscr{A} \to A$, where

$A \in \mathscr{X}$, we say that c *contains* c' (or c' is a *subchoice* of c) if the equality $c'(B) = c_{\restriction A}(B)$ holds for any $B \in \mathscr{A}$.

The central notion of this paper is a category of properties of choice functions that extend to all of their subchoices, and are non-universally satisfied on finite sets. We shall show that these properties are 'asymptotically rare', that is, the fraction of choices satisfying any of them is negligible as the size of the ground set grows larger and larger.

DEFINITION 2. A *property* $\mathscr{P}$ of choice functions is a second order logic formula (with quantification over elements and sets) involving a constant symbol for choice functions. A property $\mathscr{P}$ is called:

- *hereditary* if whenever $\mathscr{P}$ holds for a choice function c, then $\mathscr{P}$ also holds for all subchoices $c_{\restriction A}$, with A ranging over $\mathscr{X}$;
- *finitely violated* if there is a choice function for which $\mathscr{P}$ fails to hold.[2]

Then $\mathscr{P}$ is *hereditary finitely violated* (*HFV*) if it is both hereditary and finitely violated. Furthermore, we say that $\mathscr{P}$ is *asymptotically rare* if, as the size of the ground set tends to infinity, the fraction of choices satisfying $\mathscr{P}$ tends to 0.

Definitions 1 and 2 are given for choice functions, but they extend to choice correspondences: see Cantone, Giarlotta, and Watson (2021, Sections 2 and 3).

EXAMPLE 1. Many well known properties of choice functions/correspondences are HFV. For instance, Axiom α, originally introduced by Chernoff (1954) and so called by Sen (1971), is hereditary finitely violated.[3] Additional HFV properties are the following axioms of choice consistency: β, γ, ρ, *path independence*, and *WARP*.[4] On the other hand, some properties of choices are not finitely violated, and some others are not hereditary. For instance, the rationalizability of choice functions by multiple rationales à la Kalai, Rubinstein, and Spiegler (2002) is not finitely violated. Furthermore, the property of being non-chaotic[5] is non-hereditary.[6]

[2] We employ the term 'finitely' because one can consider choices defined on infinite ground sets. However, in this note we do not wish to address this more complicated issue.

[3] Recall that a choice correspondence $c\colon \mathscr{X} \to \mathscr{X}$ satisfies Axiomα whenever for any $x \in X$ and $A, B \in \mathscr{X}$, if $x \in A \subseteq B$ and $x \in c(B)$, then $x \in c(A)$.

[4] See Cantone, Giarlotta, and Watson (2021, Section 3.2), and references therein.

[5] A choice function on n items is *non-chaotic* if at most $n - 2$ linear orders rationalize it.

[6] Another property of choice correspondences that fails to be hereditary is *CWDE (Choosing Without Dominated Elements)*, which is used by García-Sanz and Alcantud (2015) to (partially) extend the characterization of the *rational shortlist method* of Manzini and Mariotti (2007) to choice correspondences: see Cantone, Giarlotta, and Watson (2021, Section 3.5).

A Non-universal models of bounded rationality

We briefly describe the non-universal[7] bounded rationality models mentioned in Section 1, and show that the associated notions of rationalizability are HFV properties.

A choice function $c\colon \mathscr{X} \to X$ is called:

(i) *Rationalizable* if there is a linear order $\rhd$ on X such that $c(A) = \max(A, \rhd)$ for all $A \in \mathscr{X}$.

(ii) *With status quo bias* (Apesteguia and Ballester, 2013, p.92) if there is a triple $(\rhd, d, Q)$, with $\rhd$ linear order on X, $d \in X$, and $Q \subseteq \{x \in X : x \rhd d\}$, such that, for any $S \in \mathscr{X}$, either properties (1)-(2)-(3) or properties (1)-(2)-(3′) hold:

  (1) if $d \notin S$, then $c(S) = \max(S, \rhd)$;

  (2) if $d \in S$ and $Q \cap S = \varnothing$, then $c(S) = d$;

  (3) if $d \in S$ and $Q \cap S \neq \varnothing$, then $c(S) = \max(Q \cap S, \rhd)$;

  (3′) if $d \in S$ and $Q \cap S \neq \varnothing$, then $c(S) = \max(S \setminus \{d\}, \rhd)$.[8]

(iii) *Sequentially rationalizable* (Manzini and Mariotti, 2007) if there exists an ordered list $(\succ_1, \ldots, \succ_n)$ of asymmetric relations on X such that, for all $A \in \mathscr{X}$, defining recursively $M_0(A) := A$ and $M_i(A) := \max(M_{i-1}(A), \succ_i)$ for each $i = 1, \ldots, n$, we have $c(A) = M_n(A)$.

(iv) *Rationalizable by game trees* (Xu and Zhou, 2007) if there is a game tree (G, R), where (1) G is a rooted tree whose terminal nodes are bijectively mapped onto X, and (2) R is a list $(R_1, \ldots, R_n)$ of linear orders on X, each of which is associated to one of the n players of the game, such that $c(A) = \mathit{SPNE}(G|A; R)$ for all $A \in \mathscr{X}$. (Here $G|A$ is the reduced tree derived from G by retaining only paths that lead to terminal nodes in A, and $\mathit{SPNE}(\Gamma)$ is the subgame perfect Nash equilibrium outcome of the game Γ.)

(v) *Rationalizable by a post-dominance rationality procedure* (Rubinstein and Salant, 2008) if there are (1) an acyclic *dominance* relation R on X, and (2) a *post-dominance* relation $\succ$ on X (which is transitive and complete whenever restricted on sets of undominated elements) such that $c(A) = \max(\max(A, R), \succ)$ for all $A \in \mathscr{X}$.

(vi) *Categorize-then-choose* (Manzini and Mariotti, 2012) if there are (1) an asymmetric relation $\succ^*$ on $\mathscr{X}$ (*shading* relation), and (2) an asymmetric complete relation $\succ$

[7] Recall that a choice model is *universal* if it rationalizes any choice, and *non-universal* otherwise.

[8] This notion, defined for a collection of 'choice problems', is due to Masatlioglu and Ok (2005). Apesteguia and Ballester (2013) call *extreme endogenous status quo biased* choices satisfying (1)-(2)-(3) for all $S \in \mathscr{X}$, and *weak endogenous status quo biased* those satisfying (1)-(2)-(3′) for all $S \in \mathscr{X}$.

on X such that $c(A) = \max(\max(A, \succ^*), \succ)$ for all $A \in \mathscr{X}$, where $\max(A, \succ^*) = \bigcup\{\max(\mathscr{A}, \succ^*)\}$.[9]

(vii) *With limited attention* (Masatlioglu, Nakajima, and Ozbay, 2012) if there are (1) a choice correspondence $\Gamma\colon \mathscr{X} \to \mathscr{X}$ (*attention filter*) such that $x \notin \Gamma(A)$ implies $\Gamma(A) = \Gamma(A \backslash \{x\})$, and (2) a linear order $\rhd$ on X such that $c(A) = \max(\Gamma(A), \rhd)$, where $A \in \mathscr{X}$ and $x \in X$.

(viii) *Consistent with basic rationalization theory* (Cherepanov, Feddersen, and Sandroni, 2013) if there are (1) a correspondence ψ on X satisfying Axiom α (*psychological constraint*), and (2) an asymmetric relation $\succ$ on X such that $c(A) = \max(\psi(A), \succ)$ for all $A \in \mathscr{X}$.[10]

(ix) *A sequential procedure guided by a set of routes* (Apesteguia and Ballester, 2013) if there is a collection $\mathscr{R}$ of linear orders (*routes*), defined on the family $\mathscr{B}$ of all binary menus of $\mathscr{X}$, such that $c(A) = c(A \backslash \{r(A^{\lhd})\})$ for all $A \in \mathscr{X}$ and $\rhd \in \mathscr{R}$, with $r(A^{\lhd})$ being the item discarded in the first (according to $\rhd$) binary submenu of A.[11]

(x) *List-rational* (Yildiz, 2016) if there is a linear order f (*list*) on X such that $c(A) = c(\{c(A \backslash \{x\}), x\})$ for all $A \in \mathscr{X}$, where $x = \max(A, f)$.

(xi) *Overwhelming* (Lleras, Masatlioglu, Nakajima, and Ozbay, 2017) if there are (1) a choice correspondence ψ on X satisfying Axiom α, and (2) a linear order $\rhd$ on X such that $c(A) = \max(\psi(A), \rhd)$ for all $A \in \mathscr{X}$.

As announced, we have:

Lemma 1. *All properties (i)–(xi) are HFV.*

Proof. All choice models (i)–(ix) are characterized by behavioral axioms, which are finitely violated. Thus, it suffices to prove hereditariness.

(i) Let $c\colon \mathscr{X} \to X$ be a choice rationalizable by some linear order $\rhd$ on X. Since $c_{\upharpoonright A}(B) = c(B) = \max(B, \rhd) = \max\left(B, \rhd_{|A}\right)$ for any $B \in \mathscr{A}$, the claim follows. (See Definition 1.)

(ii) Let $c\colon \mathscr{X} \to X$ be a choice with status quo bias. Fix $A \in \mathscr{X}$ and $a_0 \in A$. Set

[9] Given a binary relation $\succ^*$ on $\mathscr{X}$ and a menu $A \in \mathscr{X}$, the set $\max(\mathscr{A}, \succ^*)$ is the collection of (non-dominated) menus $\{B \subseteq A \mid B' \succ B \text{ for no } B' \subseteq A\}$.

[10] If $\succ$ is a linear order, c is said to be *consistent with order rationalization theory*.

[11] For any $B \in \mathscr{B}$, denote by $r(B)$ the item rejected in B. For any $A \in \mathscr{X}$, with $|X| \geqslant 2$, and any route $\rhd$ on $\mathscr{B}$, the first binary submenu of A according to the route $\rhd$ is the menu $A^{\lhd}$ such that either $A^{\lhd} = A$ if $|A| = 2$ or, otherwise, $A^{\lhd} \in \mathscr{B}$, $A^{\lhd} \subsetneq A$, and, for any $B \in \mathscr{B}$ distinct from $A^{\lhd}$ such that $B \subsetneq A$, we have that $B \rhd A^{\lhd}$.

$$d_A := \begin{cases} d & \text{if } d \in A \\ a_0 & \text{otherwise.} \end{cases} \qquad \text{and} \qquad Q_A := \begin{cases} Q \cap A & \text{if } d \in A \\ \{a \in A \colon a \rhd d_A\} & \text{otherwise.} \end{cases}$$

We claim that $(\rhd_A, d_A, Q_A)$ witnesses that $c_{\restriction A}$ is an (extreme or weak) endogenous status quo biased choice (see Footnote 8). Let $S \in \mathscr{A}$. There are 3 cases.

(I) Suppose $d \notin A$. Since c satisfies (1), we get $c(T) = \max(T, \rhd)$ for any $T \in \mathscr{X}$. By definition, d_A is an item $a_0 \in A$, and Q_A is the set of all items in A that dominate a_0. Fix an arbitrary menu $S \in \mathscr{A}$. If $d_A = a_0 \notin S$, then $c_{\restriction A}(S) = c(S) = \max(S, \rhd) = \max(S, \rhd_{\restriction A})$, hence (1) holds for $c_{\restriction A}$. If $d_A = a_0 \in S$ and $Q_A \cap S = \varnothing$, then $c_{\restriction A}(S) = c(S) = \max(S, \rhd) = \max(S, \rhd_{\restriction A}) = d_A$, and so (2) holds for $c_{\restriction A}$ as well. Finally, suppose $d_A = a_0 \in S$ and $Q_A \cap S \neq \varnothing$. Thus $c_{\restriction A}(S) = c(S) = \max(S, \rhd) = \max(S, \rhd_{\restriction A}) = \max(Q_A \cap S, \rhd_{\restriction A}) = \max(S \setminus \{d_A\}, \rhd_{\restriction A})$, which says that both (3) and (3′) hold for $c_{\restriction A}$.

(II) Suppose $d \in A$ and $Q \cap A = \varnothing$. By the definition of d_A and Q_A, we have $d_A = d$ and $Q_A = \varnothing$. Fix $S \in \mathscr{A}$. If $d_A = d \notin S$, then property (1) of c yields $c(S) = \max(S \rhd)$, and so $c_{\restriction A}(S) = c(S) = \max(S, \rhd) = \max(S, \rhd_{\restriction A})$, that is, (1) holds for $c_{\restriction A}$. On the other hand, if $d_A = d \in S$, then $c_{\restriction A}(S) = c(S) = d_A$ by property (2) for c, which proves (2) for $c_{\restriction A}$. (Note that both (3) and (3′) vacuously hold for $c_{\restriction A}$ in this case.)

(III) To complete the proof, let $d \in A$ and $Q \cap A \neq \varnothing$, hence $d_A = d$ and $Q_A = Q \cap A \neq \varnothing$ by definition. Fix $S \in \mathscr{A}$. If $d_A \notin S$, then $c_{\restriction A}(S) = c(S) = \max(S, \rhd) = \max(S, \rhd_{\restriction A})$ by property (1) for c, and so (1) holds for $c_{\restriction A}$ as well. Next, suppose $d_A \in S$. Note that since $S \subseteq A$, we have $Q_A \cap S = (Q \cap A) \cap S = Q \cap S$. Now, if $Q_A \cap S = \varnothing$, then $Q \cap S = \varnothing$, hence $c(S) = d = d_A$ by property (2) for c. It follows that $c_{\restriction A}(S) = c(S) = d = d_A$, and so (2) holds for $c_{\restriction A}$ as well. Finally, let $d_A \in S$ and $Q_A \cap S \neq \varnothing$. Since $S \subseteq A$, we get $Q \cap S = Q_A \cap S \neq \varnothing$, whence properties (3) and (3′) apply for c, yielding $c(S) = \max(Q \cap S, \rhd)$ and $c(S) = \max(S \setminus \{d\}, \rhd)$, respectively. Thus, either $c_{\restriction A}(S) = c(S) = \max(Q \cap S, \rhd) = \max(Q_A \cap S, \rhd) = \max(Q_A \cap S, \rhd_{\restriction A})$ or $c_{\restriction A}(S) = c(S) = \max(S \setminus \{d\}, \rhd) = \max(S \setminus \{d_A\}, \rhd_{\restriction A})$, that is, either (3) or (3′) holds for $c_{\restriction A}$.

(iii) Let $c \colon \mathscr{X} \to X$ be a sequentially rationalizable choice, and $(\succ^1, \ldots, \succ^n)$ an ordered list of asymmetric relations that sequentially rationalizes c. Let $A \in \mathscr{X}$. One can readily check that $\left(\succ^1_{\restriction A}, \ldots, \succ^n_{\restriction A}\right)$ is an ordered list of asymmetric relations, and sequentially rationalizes the subchoice $c_{\restriction A} \colon \mathscr{A} \to A$.

(iv) Let $c \colon \mathscr{X} \to X$ be a choice *rationalizable by a game tree*. It is known that c is rationalizable by a game tree if and only if it satisfies 'weak separability' and 'divergence consistency', defined as follows. *Weak separability* requires that for any menu $A \in \mathscr{X}$ of size at least two, there is a partition $\{B, D\} \subseteq \mathscr{A}$ of A such that $c(S \cup T) = c(\{c(S), c(T)\})$ for any $S \subseteq B$ and $T \subseteq D$.[12] For each $x, y, z \in X$, let $x \circlearrowleft \{y, z\}$ stand for $c(\{x, y, z\}) = x$ but x, y, z give rise to a *cyclic binary selection*, that is, either (i) $c(\{x, y\}) = x$, $c(\{y, z\}) = y$, and $c(\{x, z\}) = z$, or (ii) $c(\{x, y\}) = y$, $c(\{y, z\}) = z$, and $c(\{x, z\}) = x$. Then *divergence consistency* requires that for any $x_1, x_2, y_1, y_2 \in X$, if $x_1 \circlearrowleft \{y_1, y_2\}$ and $y_1 \circlearrowleft \{x_1, x_2\}$, then $c(\{x_1, y_1\}) = x_1$ if and only if $c(\{x_2, y_2\}) = y_2$. For any $A \in \mathscr{X}$, one can readily check that $c_{\upharpoonright A}$ satisfies weak separability and divergence consistency.

(v) Suppose $c \colon \mathscr{X} \to X$ is rationalizable by a post-dominance rationality procedure. Rubinstein and Salant (2008) show that such a choice is characterized by *exclusion consistency*, which requires that for any $A \in \mathscr{X}$ and any $x \in X \backslash A$, if $c(A \cup \{x\}) \notin \{c(A), x\}$, then there is no set A' containing x such that $c(A') = c(A)$. Toward a contradiction, suppose $c_{\upharpoonright A}$ violates exclusion consistency for some $A \in \mathscr{X}$. Thus, there are $B \in \mathscr{A}$ and $y \in A \backslash B$ such that $c_{\upharpoonright A}(B \cup \{y\}) \notin \{c_{\upharpoonright A}(B), y\}$, and $c_{\upharpoonright A}(B) = c_{\upharpoonright A}(B')$ for some $B' \in \mathscr{A}$. Since $A \subseteq X$, we get $B, B' \in \mathscr{X}$, $y \in X \backslash B$, $c(B \cup \{y\}) \notin \{c(B), y\}$, and $c(B) = c(B')$, which is impossible.

(vi) Suppose $c \colon \mathscr{X} \to X$ is categorize-then-choose. The authors show that c is categorize-then-choose if and only if c satisfies *Weak WARP (WWARP)*. This property, introduced by Manzini and Mariotti (2007), requires that for any $A, B \in \mathscr{X}$ and distinct $x, y \in X$, if $x, y \in A \subseteq B$ and $c(x, y) = c(B) = x$, then $c(A) \neq y$. Cantone, Giarlotta, and Watson (2019) argue that WWARP is hereditary. This suffices to prove that categorize-then-choose is hereditary.

(vii) Let $c \colon \mathscr{X} \to X$ be with limited attention. Given $A \in \mathscr{X}$, the choice correspondence $\Gamma_{\upharpoonright A}$ induced by the attention filter Γ on X is an attention filter on A. Since $c_{\upharpoonright A}(B) = c(B) = \max(\Gamma(B), \rhd) = \max\left(\Gamma_{\upharpoonright A}(B), \rhd_{\upharpoonright A}\right)$ for any $B \in \mathscr{A}$, the claim holds.

(viii) Suppose $c \colon \mathscr{X} \to X$ is consistent with basic rationalization theory. Given $A \in \mathscr{X}$, the choice correspondence $\psi_{\upharpoonright A}$ induced by the psychological constraint ψ on X is a psychological constraint on A. Thus the claim follows from $c_{\upharpoonright A}(B) = c(B) = \max(\psi(B), \succ) = \max\left(\psi_{\upharpoonright A}(B), \succ_{\upharpoonright A}\right)$ for any $B \in \mathscr{A}$.[13]

(ix) Apesteguia and Ballester (2013) show that a choice function is a procedure guided by a set of routes if and only if it is sequentially rationalizable. Thus, the claim follows from part (iii).[14]

[12] Weak separability is a restricted version of *path independence*, defined by Plott (1973), which requires that $c(A \cup B) = c(c(A), c(B))$ for any (not necessarily disjoint) $A, B \in \mathscr{X}$.

[13] We can also use (vi) to prove the claim. Indeed, consistency with basic rationalization theory is equivalent to 'categorize-then-choose' (Manzini and Mariotti, 2012), since both are characterized by WWARP.

[14] The authors prove that any choice with status quo bias is rationalizable by game trees, and, in turn,

(x) Let $c\colon \mathscr{X} \to X$ be a list-rational choice. For any $B \in \mathscr{A}$, we have that $c_{\upharpoonright A}(B) = c(B) = c(\{c(B\backslash\{x\}), x\})$, where $x = \max(B, f) = \max\big(B, f_{|A}\big)$.

(xi) Overwhelming choices are equivalent to choices consistent with order rationalization theory due to Cherepanov, Feddersen, and Sandroni (2013). Thus, the claim follows from part (viii).

This completes the proof of Lemma 1. ∎

Remark 1. Lemma 1 can be proved by using a model-theoretic approach. Let a property $\mathscr{P}$ of choice functions be *unboundedly universal* if $\mathscr{P}$ can be written as a sequence of (unbounded) universal quantifiers over items or menus, followed by a formula Q all of whose quantifiers are bounded (by the variables of those unbounded universal quantifiers). Now observe that any unboundedly universal property is hereditary, and all properties of the models (i)–(xi) (with the only exception of (ii)) are 'unboundedly universal properties'.

3 Rarity of bounded rationality

The main result of this note is the following:

Theorem 1. *Any HFV property is asymptotically rare.*

In the path to prove Theorem 1, we need some preliminary results and notions.

Lemma 2. *Let X be a finite set of size $n \geqslant 1$. For any family $(X_j)_{j=1}^{p}$ of pairwise disjoint menus in $\mathscr{X}$ having size $m_j \geqslant 1$, the number of choice functions on X is*

$$\prod_{j=1}^{p} \prod_{\varnothing \neq A \subseteq X_j} |A| \times \prod_{\substack{\varnothing \neq B \subseteq X \\ (\forall j) B \nsubseteq X_j}} |B| \;=\; \prod_{j=1}^{p} \prod_{k=1}^{m_j} k^{\binom{m_j}{k}} \times \prod_{k=1}^{n} k^{\binom{n}{k} - \sum_{j=1}^{p} \binom{m_j}{k}}. \tag{1}$$

Proof. The number of choice functions on X is

$$\prod_{\varnothing \neq A \subseteq X} |A| = \prod_{k=1}^{n} k^{\binom{n}{k}}.$$

A straightforward computation yields[15]

$$\prod_{j=1}^{p} \prod_{k=1}^{m_j} k^{\binom{m_j}{k}} \times \prod_{k=1}^{n} k^{\binom{n}{k} - \sum_{j=1}^{p} \binom{m_j}{k}} = \prod_{k=1}^{n} k^{\sum_{j=1}^{p} \binom{m_j}{k}} \times \prod_{k=1}^{n} k^{\binom{n}{k} - \sum_{j=1}^{p} \binom{m_j}{k}} = \prod_{k=1}^{n} k^{\binom{n}{k}}.$$

The claim follows. ∎

any choice rationalizable by game trees is sequentially guided by a set of routes. Thus, parts (ii) and (iv) readily follow from (iii). However, for the sake of completeness, we have included a full proof of (ii).

[15] Recall that $\binom{r}{s}$ is equal to zero by definition whenever $r < s$.

Remark 2. Lemma 2 can be stated requiring that $(X_j)_{j=1}^p$ is such that $|X_i \cap X_j| \leqslant 1$ for distinct i, j. For such a family, equation (1) still holds, since singletons may be counted more than once. This observation will be crucial to obtain finer estimates: see Section 3B.

Next, we introduce the notion of isomorphic choices.

Definition 3. Two choice correspondences $c\colon \mathscr{X} \to \mathscr{X}$ and $c'\colon \mathscr{X}' \to \mathscr{X}'$, respectively having X and X' as ground sets, are *isomorphic* if there is a bijection $\sigma\colon X \to X'$ such that $\sigma(c(A)) = c'(\sigma(A))$ for any $A \in \mathscr{X}$, where $\sigma(A)$ is the set $\{\sigma(a) : a \in A\}$. Whenever c' has a subchoice that is isomorphic to c, we shall say that c' *contains a copy of* c.

A final notion is needed:

Definition 4. Let $c\colon \mathscr{X} \to \mathscr{X}$ be a choice correspondence on X. For any permutation π of X, let $c_\pi\colon \mathscr{X} \to \mathscr{X}$ be the choice correspondence defined by $c(A) := \pi^{-1}(c(\pi(A)))$ for all $A \in \mathscr{X}$. We say that c is *dynamic* whenever all c_π's are pairwise distinct.

In other words, a choice function/correspondence is dynamic whenever any relabeling of the elements of the ground set produces a new choice. However, for the special case of choice functions, the situation is straightforward:[16]

Lemma 3. *All choice functions are dynamic.*

Proof. Toward a contradiction, suppose $c\colon \mathscr{X} \to X$ is a non-dynamic choice function. Thus, there are two distinct permutations π and σ of X such that $c_\pi = c_\sigma$. Without loss of generality, assume σ is the identity and π is not, that is, $c_\pi = c$ with $\pi \neq \mathrm{id}_X$. Set $A := \{x \in X : \pi(x) \neq x\} \in \mathscr{X}$, and so $\pi(A) = A$. We claim that $c(A)$ is a fixed point of π, that is, $\pi(c(A)) = c(A)$. Indeed,

$$c(A) = c_\pi(A) = \pi^{-1}(c(\pi(A))) \implies \pi(c(A)) = \pi(\pi^{-1}(c(\pi(A)))) = c(\pi(A)) = c(A)\,.$$

Now the definition of A yields $c(A) \notin A$, which is impossible. ∎

Finally, we get what we were after:

Corollary 1. *Given a choice function c on a ground set X of size $m \geqslant 1$, there are exactly $m!$ choice functions on X that are isomorphic to c.*

Proof. This number is clearly at most $m!$, and is at least $m!$ by Lemma 3. ∎

[16] The situation is more involved for a choice correspondence c, because the existence of nontrivial *indiscernible menus* (Cantone, Giarlotta, and Watson, 2019) makes c non-dynamic (and the correspondent group of automorphisms nontrivial). Moreover, the analysis is further complicated by the existence of non-dynamic choice correspondences with no indiscernible menus other than singletons. Details are available upon request.

Corollary 1 plays a key role in the next result:

LEMMA 4. *Let X be such that $|X| \geqslant pm \geqslant m$, and c a choice on Y such that $|Y| = m \geqslant 1$. The fraction of choices on X not containing a copy of c is $\leqslant \left(\frac{\phi(m)-m!}{\phi(m)}\right)^p$.*

Proof. Let $(X_j)_{j=1}^p$ be p disjoint menus in $\mathscr{X}$ such that $|X_j| = m$ for all j's. Set

$$\phi(m) := \prod_{\varnothing \neq A \subseteq Y} |A| = \prod_{k=1}^{m} k^{\binom{m}{k}} \qquad \text{and} \qquad \theta(m,n) := \prod_{\substack{\varnothing \neq B \subseteq X \\ (\forall j) B \nsubseteq X_j}} |B| = \prod_{k=1}^{n} k^{\binom{n}{k} - p\binom{m}{k}}.$$

Since $|X_j| = |Y|$ for all j's, Lemma 2 yields that the total number of choices on X is equal to $(\phi(m))^p \cdot \theta(m,n)$. By Corollary 1, there are exactly $m!$ choices on each X_j that are copies of c, hence the total number of choices on X not containing a copy of c is $\leqslant (\phi(m) - m!)^p \cdot \theta(m,n)$. It follows that the fraction of choices that do not contain a copy of c is at most

$$\frac{(\phi(m) - m!)^p \cdot \theta(m,n)}{(\phi(m))^p \cdot \theta(m,n)} = \left(\frac{\phi(m) - m!}{\phi(m)}\right)^p,$$

as claimed. ∎

REMARK 3. According to Remark 2, in Lemma 4 it suffices to consider a family $(X_j)_{j=1}^p$ of menus that pairwise intersect in at most one item, which can be obtained from a ground set X of size $\geqslant m + (p-1)(m-1)$. To see why, fix a linear order $\rhd$ on X such that $|X| = m + (p-1)(m-1)$, and list its elements according to $\rhd$. Define a family $\mathscr{F} := (X_j)_{j=1}^p$ of subsets of X all having size m by setting $X_j := \{x_{(j-1)m-(j-2)}, \ldots, x_{jm-(j-1)}\}$ for any $j \in \{1, \ldots p\}$. Each X_j is disjoint from all other elements of $\mathscr{F}$, with the possible exception of X_{j-1} and X_{j+1} (if they exist), for which the size of the intersection is one. Thus, in Lemma 4 we can require $|X| \geqslant p(m-1) + 1$.

LEMMA 5. *Let c be a choice on a set of size $m \geqslant 1$. For any $\epsilon > 0$, there is $N \in \mathbb{N}$ such that, for any $n > N$, the fraction of all choices on a set of size n not containing a copy of c is $< \epsilon$.*

Proof. Fix $\epsilon > 0$. Let $p \in \mathbb{N}$ be such that $\left(\frac{\phi(m)-m!}{\phi(m)}\right)^p < \epsilon$. We claim that $N := pm$ works. Let X be a set such that $|X| = n > N$. Therefore, X has at least p disjoint subsets of size m. By Lemma 4, the fraction of all choices on X not containing a copy of c is at most $\left(\frac{\phi(m)-m!}{\phi(m)}\right)^p < \epsilon$. ∎

We are finally ready to prove our main result.

Proof of Theorem 1. Let $\mathscr{P}$ be an HFV property. As $\mathscr{P}$ is finitely violated, there is a choice c on a finite set Y for which $\mathscr{P}$ fails. Fix $\epsilon > 0$. By Lemma 5, there is an integer

N such that, for any $n > N$, the fraction of all choices on a set X of cardinality n not containing a copy of c is less than ϵ. Since $\mathscr{P}$ is hereditary, $\mathscr{P}$ does not holds for any choice on X containing a copy of c. Thus, the fraction of all choices defined on X satisfying $\mathscr{P}$ is less than ϵ. ∎

The next result readily follows from Lemma 1 and Theorem 1.

COROLLARY 2. *All properties (i)–(xi) are asymptotically rare.*

Corollary 2 establishes the limited rationalizability of most bounded rationality approaches proposed in the literature. When the cardinality of the ground set increases, all these methods gradually loose their predictive accuracy, and justify few choice behaviors observed on that set. This fact supports the claim that all bounded rationality methods introduced in the literature are explicative of 'rational' behavior.

REMARK 4. The argument used in the proof of Theorem 1 also applies to choice correspondences, provided that Lemmata 2 and 4 are suitably reformulated. To start, note that the number of choice correspondences on a set X of size n is (Aleskerov, Bouyssou, and Monjardet, 2007, p.29)

$$\prod_{\varnothing \neq A \subseteq X} \left(2^{|A|} - 1\right) = \prod_{k=1}^{n} \left(2^k - 1\right)^{\binom{n}{k}}.$$

Thus, given a family $(X_j)_{j=1}^{p}$ of p disjoint menus in $\mathscr{X}$ having size $m_j \geqslant 1$, the number of choice correspondences on X is

$$\prod_{j=1}^{p} \prod_{\varnothing \neq A \subseteq X_j} \left(2^{|A|} - 1\right) \times \prod_{\substack{\varnothing \neq B \subseteq X \\ (\forall j) B \nsubseteq X_j}} \left(2^{|B|} - 1\right) = \prod_{j=1}^{p} \prod_{k=1}^{m_j} \left(2^k - 1\right)^{\binom{m_j}{k}} \times \prod_{k=1}^{n} \left(2^k - 1\right)^{\binom{n}{k} - \sum_{j=1}^{p} \binom{m_j}{k}}.$$

For a ground set X of size $n \geqslant mp$, there is a family $(X_j)_{j=1}^{p}$ of p disjoint menus in $\mathscr{X}$ all having size m. Thus, we can set (by adapting the definitions of ϕ and θ for choice correspondences)

$$\phi(m) := \prod_{k=1}^{m} \left(2^k - 1\right)^{\binom{m}{k}} \quad \text{and} \quad \theta(m,n) := \prod_{k=1}^{n} \left(2^k - 1\right)^{\binom{n}{k} - p\binom{m}{k}}$$

to conclude that the number of all choice correspondences on X is $\phi^p(m) \cdot \theta(m,n)$. Let c be a choice correspondence on a set Y of cardinality m not satisfying some HFV property $\mathscr{P}$. Since there is at least a copy of c on each X_j, the fraction of choices on X which do not contain a copy of c is at most $\left(\frac{\phi(m)-1}{\phi(m)}\right)^p$. This suffices to prove, using Lemma 5 and the remainder of Theorem 1's proof, that any HFV property of choice correspondences is asymptotically rare. A better estimate could be obtained by computing the number of choice correspondences that are isomorphic to a given one, which however appears to be a rather challenging combinatorial problem.

A Numerical estimates

Lemma 4 allows us to obtain an upper bound for the probability to catch a boundedly rationalizable choice on a ground set of a given size. Indeed, we have:

Corollary 3. *Let c be a choice on a set Y of size m not satisfying an HFV property $\mathscr{P}$. Furthermore, let X be a set of size at least $\zeta(\epsilon, m) := \left\lceil \frac{\log \epsilon}{\log\left(\frac{\phi(m)-m!}{\phi(m)}\right)} \right\rceil (m-1)+1$, where $0<\epsilon<1$. The fraction of choices on X satisfying $\mathscr{P}$ is less than ϵ.*[17]

Proof. Fix $\epsilon > 0$, and let $p \in \mathbb{N}$ such that $\left(\frac{\phi(m)-m!}{\phi(m)}\right)^p < \epsilon$. Taking logs, we get

$$p > \frac{\log \epsilon}{\log\left(\frac{\phi(m)-m!}{\phi(m)}\right)} \,. \tag{2}$$

Let p^* be the ceiling of $\frac{\log \epsilon}{\log\left(\frac{\phi(m)-m!}{\phi(m)}\right)}$. By Remark 3 and inequality (2), the fraction of choices on X not containing a copy of c is less than ϵ. Since $\mathscr{P}$ is hereditary, the fraction of choices on X satisfying $\mathscr{P}$ is less than ϵ. ∎

By Corollary 3 and Lemma 1, for any $0 < \epsilon < 1$ and bounded rationality model not explaining a choice on a set of minimal size m, the probability of finding a non-rationalizable choice on a set X of cardinality at least $\zeta(\epsilon, m)$ is less than ϵ. Note that if m is large, then, for a given ϵ, the value of $\zeta(\epsilon, m)$ may be exceptionally high. However, this never happens for all the models examined in this paper, because the corresponding minimum value of m is either 3 or 4:

Lemma 6. *The minimum size m of the ground set of a choice that fails to be boundedly rationalizable by models (i)–(xi) is*

(a) *$m = 3$ for (i), (ii), (iii), (iv), (v), (ix), and (x),*

(b) *$m = 4$ for (vi), (vii), (viii), and (xi).*

Proof. (a) Clearly, we have $m \geqslant 3$ for models (i), (ii), (iii), (iv), (v), (ix), and (x). To show that $m \leqslant 3$ for all of them, note any sequentially rationalizable choice satisfies the axiom *Always Chosen* (Manzini and Mariotti, 2007, p. 1831),[18] whereas a list-rational choice is

[17] Here $\lceil \log \epsilon / \log\left((\phi(m)-m!)/\phi(m)\right)\rceil$ denotes the *ceiling* of $\log \epsilon / \log\left((\phi(m)-m!)/\phi(m)\right)$.

[18] *Always Chosen* holds for c if for any $A \in \mathscr{X}$ and $x \in A$, if $c(\{x,y\}) = x$ for all $y \in A \setminus \{x\}$, then $c(A) = x$.

characterized by the asymmetry and acyclicity of the binary relation F_c of *revealed-to-follow* (Yildiz, 2016, p. 594).[19] Consider the choice function $c\colon \mathscr{X} \to X$ on $X = \{x, y, z\}$ defined by

$$xy\underline{z}\,,\ \underline{x}z\ ,\ \underline{y}z\,,\ \underline{x}y\,.$$

Always Chosen does not hold for c, hence c is not sequentially rationalizable, and it also fails to be with status quo bias, rationalizable by game trees, and guided by a set of routes. Furthermore, c violates exclusion consistency,[20] thus is not rationalizable by a post-dominance rationality procedure. Finally, c is not list rational, because F_c is not asymmetric ($xF_c\,z$ and $zF_c\,x$).

(b) Again, $m \geqslant 3$ for models (vi), (vii), (viii), and (ix). We show that $m > 3$ for all of them. A choice with limited attention is characterized by WARP(LA) (Masatlioglu, Nakajima, and Ozbay, 2012, p. 2193),[21] whereas choices categorize-then-choose and those consistent with basic rationalization theory are characterized by WWARP (Cherepanov, Feddersen, and Sandroni, 2013, p. 780): both axioms always hold on three items. Overwhelming choices are consistent with order rationalization theory, which rationalizes any choice on three items. Finally, we prove $m \leqslant 4$ for all these models. Let $c\colon \mathscr{X} \to X$ be the choice on $X = \{w, x, y, z\}$ defined by

$$wxy\underline{z}\,, \qquad wx\underline{y}\,,\ \underline{w}xz\,,\ w\underline{y}z\,,\ \underline{x}yz\,, \qquad \underline{w}x\,,\ \underline{w}y\,,\ w\underline{z}\,,\ \underline{x}y\,,\ x\underline{z}\,,\ \underline{y}z.$$

WARP(LA) fails for c (take $A = \{x, y, z\}$), hence c it is not with limited attention. Moreover, c is neither categorize-then-choose, nor consistent with basic rationalization theory, nor overwhelming, because it does not satisfy WWARP.[22] ■

Table 1 displays $\zeta(\epsilon, m)$ for the special cases $m \in \{3, 4\}$ and $\epsilon \in \{10\%, 5\%, 1\%\}$.

---

[19] For distinct $x, y \in X$, $xF_c\,y$ holds if there is $A \in \mathscr{X}$ such that either (i) $c(A \cup \{y\}) = x$ and $(c(\{x, y\}) = y$ or $c(A) \neq x)$, or (ii) $c(A \cup \{y\}) \neq x$ and $(c(\{x, y\}) = x$ and $c(A) = x)$. Recall that *acyclic* means that there are no distinct $p \geqslant 3$ items $x_1\, x_2, \ldots, x_p \in X$ such that $x_1 F_c\, x_2\, F_c \ldots F_c\, x_n F_c\, x_1$. (Note that 'acyclic' in Yildiz (2016) stands for 'asymmetric and acyclic'.)

[20] Take $A = \{x, z\}$, $a = y$, and $A' = \{x, y\}$.

[21] A choice $c\colon \mathscr{X} \to X$ satisfies *WARP(LA)* if for any $A \in \mathscr{X}$, there is $x^* \in A$ such that, for any B containing x^*, if $c(B) \in A$ and $c(B) \neq c(B \setminus \{x^*\})$, then $c(B) = x^*$.

[22] Take $w, z \in \{w, x, z\} \subsetneq X$.

ϵ	$m = 3$	$m = 4$
0.1	19	5968
0.05	23	7762
0.01	35	11932

Table 1: A rough test for rationalizability of bounded rationality models. In the first column, three values of ϵ are considered. In the other two columns, we compute the size $\zeta(\epsilon, m)$ of the ground set X for m equal to 3 and 4. The numbers in the second column apply to models (i), (ii), (iii), (iv), (v), (ix), and (x) in Lemma 1, whereas the numbers in the third column apply to models (vi), (vii), (viii), and (xi).

B Finer estimates

The values of $\zeta(\epsilon, m)$ in Table 1 provide rough estimates of the minimum size $\zeta(\epsilon, m)$ of the ground set X such that the fraction of boundedly rationalizable choices on X is less than ϵ, whenever the minimum size m of a counterexample is either 3 or 4. Using a combinatorial approach, here we exhibit much finer estimates, substantially reducing — *ceteris paribus* — the size of X (or, equivalently the value of ϵ). To that end, we introduce the following notion:

Definition 5. Let X be a nonempty finite set of size n, and $m \in \mathbb{N}$ such that $1 \leqslant m \leqslant n$. A collection $\mathscr{F}$ of subsets of X is an *m-leveled almost disjoint family* (a *$\mathsf{LAD}(m)$ family*) if $|A| = m$ for all $A \in \mathscr{F}$, and $|A \cap B| \leqslant 1$ for any distinct $A, B \in \mathscr{F}$.

As said in Remark 3, the family $(X_j)_{j=1}^{p}$ of menus considered in the proof of Lemma 4 may be any $\mathsf{LAD}(m)$ family $\mathscr{F}$ such that $|\mathscr{F}| = p$. Arguing as in Corollary 3, we derive:

Corollary 4. *Let $\mathscr{P}$ be an HFV property, and c a choice on a set Y of size m such that $\mathscr{P}$ fails for it. Furthermore, let X be a set for which there is a $\mathsf{LAD}(m)$ family $\mathscr{F}$ such that $|\mathscr{F}| = \left\lceil \frac{\log \epsilon}{\log\left(\frac{\phi(m) - m!}{\phi(m)}\right)} \right\rceil$, with $0 < \epsilon < 1$. The fraction of choices on X satisfying $\mathscr{P}$ is less than ϵ.*

By virtue of Lemma 1, Corollary 4 applies to all models (i)-(xi). Since the larger a $\mathsf{LAD}(m)$ family is and the better the estimate becomes, we can remarkably improve the numbers given in Table 1 by exhibiting suitable $\mathsf{LAD}(m)$ families.

For instance, for $\epsilon = 0.032$ and $m = 3$, we have $\left\lceil \frac{\log 0.032}{\log\left(\frac{\phi(3) - 3!}{\phi(3)}\right)} \right\rceil = 12$. Since each set X of size $|X| \geqslant 9$ has a $\mathsf{LAD}(3)$ family of 12 menus,[23] Corollary 4 yields that the fraction of boundedly rationalizable choices on a set of size at least 9 is less than 3.2%: see Table 2.

[23] To prove this, display the elements of X in a 3×3 matrix, and take as subsets all rows, columns, and diagonals (at 45° and −45° degrees, in a determinant-like fashion).

For another example, take $\epsilon = 1.009 \cdot 10^{-8}$ and $m = 3$, hence $\left\lceil \frac{\log 1.009 \cdot 10^{-8}}{\log\left(\frac{\phi(3)-3!}{\phi(3)}\right)} \right\rceil = 64$. Any set X such that $|X| \geqslant 21$ has a $\mathsf{LAD}(3)$ family of 64 menus.[24] By Corollary 4, the probability of catching a boundedly rationalizable choice on a set of size 21 becomes trifling ($< 0.000011\%$).

ϵ	$m = 3$
0.032	9
0.0000001009	21

Table 2: Refined estimates for $m = 3$. The first column gives an upper bound for the probability of having a choice justified by models (i), (ii), (iii), (iv), (v), (ix), and (x) on sets of size 9 and 21.

We conclude by providing better estimates for $m = 4$. The next results are needed.

Lemma 7. *WWARP does not hold for at least $\frac{3}{5}$ of all choice functions on 4 elements.*

Lemma 8. *WARP(LA) does not hold for at least $\frac{5}{12}$ of all choice functions on 4 elements.*

Proof of Lemma 7. Let $Y = \{a, b, c, d\}$. Without loss of generality, suppose $\underline{a}bcd$.[25] The selection from the menus $\{a, b\}$, $\{a, c\}$, and $\{a, d\}$ can only be one of the following:

(1) exactly one of $\underline{a}b$, $\underline{a}c$, $\underline{a}d$ holds (which happens for a fraction $\frac{3}{8}$ of all choices);

(2) exactly two of $\underline{a}b$, $\underline{a}c$, $\underline{a}d$ hold (which happens for a fraction $\frac{3}{8}$ of all choices);

(3) all of $\underline{a}b$, $\underline{a}c$, $\underline{a}d$ hold (which happens for a fraction $\frac{1}{8}$ choices);

(4) none of $\underline{a}b$, $\underline{a}c$, $\underline{a}d$ holds (which happens for a fraction $\frac{1}{8}$ of all choices).

We examine the first three cases.

(1) Without loss of generality suppose $\underline{a}b$. If either $a\underline{b}c$ or $a\underline{b}d$ holds, then WWARP fails, and the fraction for which this happens is $1 - \left(\frac{2}{3}\right)^2 = \frac{5}{9}$.

(2) Without loss of generality suppose $\underline{a}b$ and $\underline{a}c$ hold. If either $a\underline{b}c$, or $ab\underline{c}$, or $a\underline{b}d$, or $a\underline{c}d$ hold, then WWARP fails, and the fraction for which this happens is $1 - \frac{1}{3} \cdot \left(\frac{2}{3}\right)^2 = \frac{23}{27}$.

(3) If either $a\underline{b}c$, or $ab\underline{c}$, or $a\underline{b}d$, or $ab\underline{d}$, or $ac\underline{d}$, or $a\underline{c}d$ holds, then WWARP fails, and the fraction for which this happens is $1 - \left(\frac{1}{3}\right)^3 = \frac{26}{27}$.

It follows that the total fraction of choices on Y such that WWARP fails is at least

$$\frac{3}{8} \cdot \frac{5}{9} + \frac{3}{8} \cdot \frac{23}{27} + \frac{1}{8} \cdot \frac{26}{27} = \frac{35}{54} > \frac{3}{5},$$

[24] The proof of this fact uses modular arithmetic, and is available upon request.

[25] The argument given for these choices can be replicated for all choices such that $a\underline{b}cd$, or $ab\underline{c}d$, or $abc\underline{d}$ holds. The percentage of choices is always the same for each of these distinct cases, so the same holds at a global level.

as claimed. ■

Proof of Lemma 8. Recall that $c' \colon \mathscr{X} \to X$ satisfies WARP(LA) if and only if the relation P on X, defined by xPy if there is $A \in \mathscr{X}$ such that $x = c'(A) \neq c'(A \backslash \{y\})$, is asymmetric and acyclic (Masatlioglu, Nakajima, and Ozbay, 2012, Lemma 1 and Theorem 3). Let $Y = \{a, b, c, d\}$. Without loss of generality, suppose $\underline{a}bcd$. The selection from the menus $\{a, b, c\}$, $\{a, b, d\}$, and $\{a, c, d\}$ can only one of the following:

(1) exactly one of $\underline{a}bc$, $\underline{a}bd$, $\underline{a}cd$ holds (which happens for a fraction $\frac{4}{9}$ of all choices);

(2) exactly two of $\underline{a}bc$, $\underline{a}bd$, $\underline{a}cd$ hold (which happens for a fraction $\frac{2}{9}$ of all choices);

(3) all of $\underline{a}bc$, $\underline{a}bd$, $\underline{a}cd$ hold (which happens for a fraction $\frac{1}{27}$ of all choices);

(4) none of $\underline{a}bc$, $\underline{a}bd$, $\underline{a}cd$ holds (which happens for a fraction $\frac{8}{27}$ of all choices).

Next, we obtain the fraction of choices for which WARP(LA) fails in (1) and (4).

(1) Without loss of generality, let $\underline{a}bc$. Note that aPc and aPb. Consider the selection from $\{a, b, d\}$ and $\{a, c, d\}$. Exactly one of the following subcases holds:

(1.1) $a\underline{b}d$ and $a\underline{c}d$ (which happens for a fraction $\frac{1}{4}$ of all choices);
(1.2) $a\underline{b}d$ and $ac\underline{d}$ (which happens for a fraction $\frac{1}{4}$ of all choices);
(1.3) $ab\underline{d}$ and $a\underline{c}d$ (which happens for a fraction $\frac{1}{4}$ of all choices);
(1.4) $ab\underline{d}$ and $ac\underline{d}$ (which happens for a fraction $\frac{1}{4}$ of all choices).

We examine the first three subcases.

(1.1) If $b\underline{d}$ or $c\underline{d}$, then bPa or cPa, hence WARP(LA) fails. Within this subcase, the fraction of choices for which this happens is $\frac{3}{4}$.
(1.2) If $b\underline{d}$, then bPa, hence WARP(LA) fails. Within this subcase, the fraction of choices for which this happens is $\frac{1}{2}$.
(1.3) If $c\underline{d}$, then cPa, hence WARP(LA) fails. Within this subcase, the fraction of choices for which this happens is $\frac{1}{2}$.

We conclude that, within case (1), the fraction of choices for which WARP(LA) fails to hold is at least $\frac{3}{16} + \frac{1}{8} + \frac{1}{8} = \frac{7}{16}$.

(4) Note that aPd, aPc, and aPb. Without loss of generality, let $a\underline{b}c$. As in case (1), consider the selection from $\{a, b, d\}$ and $\{a, c, d\}$, for which exactly one of following subcases happens:

(4.1) $a\underline{b}d$ and $a\underline{c}d$ (for a fraction $\frac{1}{4}$ of choices);
(4.2) $a\underline{b}d$ and $ac\underline{d}$ (for a fraction $\frac{1}{4}$ of choices);
(4.3) $ab\underline{d}$ and $a\underline{c}d$ (for a fraction $\frac{1}{4}$ of choices);

(4.4) $ab\underline{d}$ and $ac\underline{d}$ (for a fraction $\frac{1}{4}$ of choices).

We examine separately each of the four subcases.

(4.1) If $b\underline{d}$ or $c\underline{d}$, then bPa or cPa, so WARP(LA) fails (for a fraction $\frac{3}{4}$).
(4.2) If $b\underline{d}$ or $\underline{c}d$, then bPa or dPa, so WARP(LA) fails (for a fraction $\frac{3}{4}$).
(4.3) If $\underline{b}d$ or $c\underline{d}$, then dPa or cPa, so WARP(LA) fails (for a fraction $\frac{3}{4}$).
(4.4) If $\underline{b}d$ or $\underline{c}d$, then dPa, so WARP(LA) fails (for a fraction $\frac{3}{4}$).

Thus, within case (2), the fraction for which WARP(LA) fails is at least $\frac{3}{4}$.

It follows that the total fraction of choices for which WARP(LA) fails is at least

$$\frac{4}{9}\cdot\frac{7}{16}+\frac{8}{27}\cdot\frac{3}{4} \;=\; \frac{5}{12}\,,$$

as claimed. ■

By Remark 2, the number of choices on X of size n is $(\phi(4))^p \cdot \theta(4,n)$. Lemma 7 yields that the fraction of choices on X satisfying WWARP is less than $\frac{\left(\frac{2}{5}\phi(4)\right)^p\cdot\theta(4,n)}{(\phi(4))^p\cdot\theta(4,n)} = \left(\frac{2}{5}\right)^p$. Using Lemma 8, we can argue similarly for WARP(LA), and finally obtain:

Corollary 5. *For any $\mathsf{LAD}(m)$ family $\mathscr{F}$ of subsets of X such that $|\mathscr{F}| = p$,*

- *the fraction of choices on X satisfying WWARP is less than $\left(\frac{2}{5}\right)^p$, and*
- *the fraction of choices on X satisfying WARP(LA) is less than $\left(\frac{7}{12}\right)^p$.*

Corollary 5 entails better estimates for models (vi), (vii), (viii), and (xi), which have a minimal counterexample of size $m = 4$. For instance, a set X such that $|X| = 28$ has a $\mathsf{LAD}(4)$ family of size 57. By Corollary 5, the probability of finding a choice that is rationalizable by (vi), (viii), and (xi) on a set of cardinality 28 is less than $\left(\frac{2}{5}\right)^{57} = 2.07\cdot 10^{-23}$, whereas the probability of catching a choice on the same set explained by (vii) is less than $\left(\frac{7}{12}\right)^{57} = 4.5\cdot 10^{-14}$: see Table 3.

	ϵ	
$\lvert X\rvert$	Weak WARP	WARP(LA)
28	$2.07\cdot 10^{-23}$	$4.5\cdot 10^{-14}$

Table 3: Refined estimates for $m = 4$. The last two columns give an upper bound for the probability that a choice on X of size 28 is rationalizable by, respectively, models (vi)-(viii)-(xi) and model (vii).

Remark 5. The fraction ϵ of non-rationalizable choices on a ground set of cardinality $\zeta(\epsilon, m)$, and the associated refinements discussed here are an *ex-ante* approximation of the *hit rate*, as defined by Selten (1991, p. 194). This score, which gives the relative frequency of correct predictions, is a component of a global measure of predictive success of a theory.

Starting from Afriat (1974), several attempts have been made to identify a *measure of rationality*, which may take into account deviations of individual behavior from the maximization principle. In this respect, Apesteguia and Ballester (2017) define the so-called *swap index*, which is the sum, across all the observed menus, of the number of alternatives that must be swapped with the chosen one in order to obtain a choice rationalizable by the linear order(s) maximizing this sum. Our numerical estimates may be seen as a benchmark to investigate performances of rationality indices as the number of available grows larger and larger.

References

Afriat, S. 1974. On a system of inequalities in demand analysis: an extension of the classical method. *International Economic Review* 14: 460–472.

Aleskerov, F., Bouyssou, D., and Monjardet, B., 2007. *Utility Maximization, Choice and Preference*. Berlin: Springer, 2nd edition.

Apesteguia, J., and Ballester, M. A., 2013. Choice by sequential procedures. *Games and Economic Behavior* 77: 90–99.

Apesteguia, J., and Ballester, M. A., 2017. A measure of rationality and welfare. *Journal of Political Economy* 123: 1278–1310.

Cantone, D., Giarlotta, A., and Watson, S., 2019. Congruence relations on a choice space. *Social Choice and Welfare* 52: 247–294.

Cantone, D., Giarlotta, A., and Watson, S., 2021. Choice resolutions. *Social Choice and Welfare* 56: 713–753.

Cherepanov, V., Feddersen, T., and Sandroni, A., 2013. Rationalization. *Theoretical Economics* 8: 775–800.

Chernoff, H., 1954. Rational selection of decision functions. *Econometrica* 22: 422–443.

García-Sanz, M., and Alcantud J. C. R., 2015. Sequential rationalization of multivalued choice. *Mathematical Social Sciences* 74: 29–33.

Kalai, G., Rubinstein, A., and Spiegler, R., 2002. Rationalizing choice functions by multiple rationales. *Econometrica* 70: 2481–2488.

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: 1824–1839.

Manzini, P., and Mariotti, M., 2012. Categorize then choose: Boundedly rational choice and welfare. *Journal of the European Economic Association* 10: 1141–1165.

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

Masatlioglu, Y., and Ok, E. A., 2005. Rational choice with status quo bias. *Journal of Economic Theory* 121: 1–29.

Plott, C., 1973. Path independence, rationality and social choice. *Econometrica* 41: 1075–1091.

Rubinstein, A., and Salant, Y., 2008. Some thoughts on the principle of revealed preference. In: Schotter, A., and Caplin, A., *The Foundations of Positive and Normative Economics: A Handbook* pp. 116–124. Oxford University Press.

Samuelson, A. P., 1938. A note on the pure theory of consumer's behaviour. *Economica* 17: 61–71.

Selten, R., 1991. Properties of a measure of predictive success. *Mathematical Social Sciences* 21: 153–167.

Sen, A. K., 1971. Choice functions and revealed preference. *The Review of Economic Studies* 38: 307–317.

Simon, H. A., 1955. A behavioral model of rational choice. *Quarterly Journal of Economics* 69: 99–118.

Xu, Y., and Zhou, L., 2007. Rationalizability of choice functions by game trees. *Journal of Economic Theory* 134: 548–556.

Yildiz, K., 2016. List-rationalizable choice. *Theoretical Economics* 11: 587–599.