

Preference for flexibility in the absence of learning: the risk attitude effect

Stefano Ficco · Vladimir A. Karamychev

Received: 3 August 2007 / Accepted: 20 May 2008 / Published online: 10 June 2008
© The Author(s) 2008

Abstract Individuals' preferences over opportunity sets may display "preference for flexibility" which prescribes to gradually eliminate alternatives from a given set until a final choice is made. One rationale for this preference for flexibility is individuals' incentive to postpone the final choice in order to better learn their underlying preferences over basic alternatives. In this paper we show that even in the absence of learning, preference for flexibility arises if individuals are risk-averse or, at least, are not very risk-seeking. Thus, individual's attitude towards risk provides yet another rationale for preference for flexibility. One of our results is that in the absence of learning, risk-neutral as well as risk-averse individuals display the same, maximal preference for flexibility.

Keywords Individual choice · Preference for flexibility · Risk attitude

JEL Classification D01 · D81

We thank Han Bleichrodt, Robert Dur, Chaim Fershtman, Maarten Janssen, Peran van Reeve, Peter Wakker, and Timothy van Zandt for helpful comments to and inspiring discussions. We are very grateful to the anonymous referee for very constructive comments and suggestions.

S. Ficco
Europe Economics, Chancery House,
55-64 Chancery Lane, London WC2A 1QU, UK
e-mail: Ficcos@europe-economics.com

V. A. Karamychev (✉)
Department of Economics, Erasmus School of Economics and Business Economics,
Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands
e-mail: Karamychev@FEW.EUR.NL

1 Introduction

The traditional economic approach views individual choice as a contingent choice of the best alternative from an opportunity set. In certain situations, however, the choice is made in more than one stage. A well known example (due to [Kreps 1979](#)) is that of making reservations at a restaurant. Prior to choosing a specific meal, an individual chooses a restaurant, i.e., a menu, from which he will later choose a meal.

Analyzing individuals' multi-stage choice behavior, economists have identified individual's desire for "flexibility" or for "freedom of choice". In the former example, even though the individual may prefer a menu containing only *steak* to a menu containing only *chicken*, he may, nevertheless, strictly prefer a menu containing both *steak* and *chicken* to either of the first two menus because it gives him greater *flexibility* of choice in the second stage. Kreps rationalizes this choice behavior by individual's uncertainty about future tastes (cf. [Koopmans 1964](#)).^{1,2} The decision-maker optimally postpones his choice as this allows him to learn his preferences over basic alternatives.

Typically, individual learning requires gathering and processing relevant information. In complex environments, however, these activities can be far beyond humans' cognitive abilities.³ In this paper we show that even in the absence of learning, when postponing a choice does not improve individuals' information, there might still be incentives to do so. In particular, individual's attitude towards risk creates incentives to postpone the final choice, i.e., leads to preference for flexibility. Thus, individual's degree of risk aversion provides yet another rationale for preference for flexibility.

We show that when the decision-maker is either risk-averse or risk-neutral, his ranking of opportunity sets displays maximal desire for flexibility: in each selection stage only one alternative is eliminated, all other alternatives are passed to the next stage. When, to the contrary, the decision-maker is very risk-seeking, his ranking of opportunity sets shows absolute flexibility aversion, and the decision-maker makes his final choice in the very first stage by choosing only one alternative. We also show that for all other shapes of Bernoulli utility functions, the ranking of opportunity sets may display all intermediate degrees of preference for flexibility: in contrast to very risk-seeking decision-makers, the choice may require more than one selection stage, and in contrast to risk-averse and risk-neutral decision-makers, more than one alternative must be eliminated at some stages.

We use the two-stage choice model of [Kreps \(1979\)](#) and extend it to multiple stages. This allows us to abstract from "the last-stage effect" which forces a decision-maker to choose a single alternative rather than a set thereof. Starting with an initial set of alternatives, an individual sequentially eliminates alternatives until only one alterna-

¹ [Pattanaik and Xu \(1990\)](#) propose a similar cardinality criterion for ranking menus in terms of "freedom of choice".

² [Puppe \(1995, 1996\)](#) suggests a characterization of "freedom of choice" which combines individual preferences over basic alternatives with individual's desire to choose from a larger opportunity set.

³ For studies that explore the role of humans' cognitive limitations in economics and in psychology see, e.g., [Simon \(1955\)](#) and [Miller \(1956\)](#). For an exhaustive survey of imperfect information processing and bounded rationality see, e.g., [Lipman \(1995\)](#).

tive remains in the set. In each stage, the decision-maker decides which alternatives he passes to the next stage, and which alternatives he eliminates.

We follow [Kreps \(1979\)](#) and assume that, although the decision-maker has a rational preference relation over basic alternatives, he is uncertain about it. In each selection stage the decision-maker gets a signal which provides him imprecise information about his preferences. This signal can be interpreted as a result of individual's personal perceptive comparison of alternatives, and, therefore, it represents his subjective preference relation at that stage.

We disable learning by assuming that in each stage the decision-maker compares opportunity sets only using a signal from that stage. We may also view this imperfect recall feature as re-labeling of alternatives in each stage. With such an interpretation, this reflects high complexity of the environment of a given choice problem.⁴ Having observed his subjective preferences at a stage, the decision-maker compares all subsets of the current set of alternatives and selects an optimal subset for the next stage. The final choice is made when the subset contains only one alternative.

[Puppe \(1995\)](#) points out that there exists an intrinsic inconsistency between even a "mild nontrivial preference for flexibility" and the rationality of underlying preferences. This impossibility result manifests that if preferences are rational, the ranking of opportunity sets must be degenerate. It is the individual's uncertainty about preferences that allows us to circumvent this impossibility result.

The literature on choice flexibility takes a normative viewpoint and does not assume any specific information structure and any specific individual learning process. In contrast, we take a positive approach to preference for flexibility by assuming a particular information structure. In our model, preference for flexibility arises in the form of individual's incentives to postpone the final choice by selecting an opportunity set with multiple alternatives. We show that individual's incentives to postpone final choices may arise even in the absence of learning. In other words, individuals may want to delay choices not only because they expect *better* information in the future, but also because they expect *different* information. In the latter case, individuals' attitude towards risk becomes, for the chosen information structure, the only factor which determines preference for flexibility.

The reason why *different* information is desirable for individuals in the middle of a choice process is that the current set of alternatives contains intrinsic information about ranks of all its alternatives.⁵ Although in the absence of learning, individuals are not able to explicitly combine current signal with previous ones, they are still able to implicitly combine current signal with intrinsic information that the current sample carries. In this respect, the choice process is a process of combining intrinsic information with the current signal in order to improve intrinsic information in the next stage. That is why the impossibility result of [Puppe \(1995\)](#) does not hold here: passing all alternatives to the next stage does not improve intrinsic information in the next stage because the sample remains the same. Eliminating some alternatives, to the

⁴ [Simon \(1955\)](#) points out difficulties in distinguishing individuals' cognitive limitations from environmental complexities in modeling boundedly rational behavior.

⁵ In other words, the distribution of ranks in opportunity sets at any subsequent selection stages differs from the *ex ante* uniform distribution.

contrary, may improve that information, even though the sample will contain fewer alternatives.

This paper can also be related to the literature on choice rationalization, in which individuals sequentially use different rationales (preference relations over alternatives) and gradually eliminate all but one alternative. One of the pioneering papers in viewing the choice as a result of sequential elimination is of [Tversky \(1972\)](#). This “choice-by-elimination” approach has recently been used in the literature on choice rationalization and proven to be useful in rationalizing choices by multiple rationales.⁶ Our paper differs from this literature in that our decision-maker does not use given rationales as choice heuristics. He does not need to follow an exogenously fixed procedure of maximizing his subjective preferences over alternatives, i.e., rationales, at each stage. We show that a decision-maker may (and will, if he is risk-averse) find it strictly sub-optimal to select only the best (in accordance with the corresponding rationale) alternatives at a stage. Instead, we give him the freedom to derive his preference over subsets of alternatives and eliminate alternatives accordingly using informational links between the rationale (signal) at a stage and the underlying preference relation over alternatives. In this respect, we rationalize choice procedures rather than individuals’ actual choices.

For the chosen informational structure, we show that “extremely”⁷ risk-seeking individuals have a unique optimal choice process: to use all the information available at a stage in order to identify the best alternative, and to select only that best alternative. The final choice is made in the very first stage by using only one subjective preference. As a result, this choice satisfies all rationality criteria and is rationalized by classic utility maximization. On the opposite side of the spectrum, all risk-averse individuals, including risk-neutral ones, also have a unique and uniform optimal choice process: to identify the worst alternative in the current sample and to eliminate it, which we refer to as to “Single Worst Elimination” choice procedure. This procedure induces choices that although satisfy transitivity, generally violate the Weak Axiom of Revealed Preferences (see [Ficco et al. 2006](#)).

Consequently, choices of risk-averse individuals are not rationalizable by utility maximization. In the middle, choice processes of individuals, which utility functions are neither concave nor extremely convex, can show all intermediate degrees of preference for flexibility, and, as long as these choice processes involve more than one selection stage, the induced choices violate the Weak Axiom. In this way, we establish a link between individual’s degree of risk-aversion and the rationality of the individual’s choice through the degree of preference for flexibility of choice.

The rest of the paper is organized as follows. Section 2 introduces the choice model, which is then analyzed in Sect. 3. Section 4 concludes. Appendices contain some proofs.

⁶ See, for instance, [Kalai et al. \(2002\)](#), and [Manzini, Mariotti \(2007\)](#). [Krause \(2008\)](#) extends this framework to intertemporal setting.

⁷ We define an “extremely” risk-seeking individual as an individual who gets positive utility level only by choosing the best alternative from the initial opportunity set, and zero utility level otherwise.

2 The choice model

A decision-maker with a rational but yet unknown preference relation over basic alternatives, must choose a single alternative from an initial set that consists of $N > 1$ alternatives.⁸ He makes his choice in a sequence of stages. In each stage t when the opportunity set consists of $N_t > 1$ alternatives, the decision-maker observes a ranking of alternatives, from 1, the best, to N_t , the worst. With probability $q \in (0, 1)$, this ranking represents his underlying true preference relation over alternatives. With the remaining probability $(1 - q)$, the ranking is a purely random ranking. Having observed ranks of each of N_t alternatives, perceived ranks hereinafter, the decision-maker selects a subset of alternatives of size N_{t+1} , $N_{t+1} = 1, \dots, N_t - 1$, which becomes his opportunity set in stage $t + 1$. Thus, in each stage, he decides how many alternatives and alternatives of which ranks he selects for the next stage. This decision represents the choice process of the decision-maker.

We prevent the decision-maker from learning by assuming that he is prospective but not retrospective so that he cannot aggregate information from the past. In every stage, he either does not use or does not remember signals that he has been using in the previous stages, or, alternatively, these signals contain no relevant information at that stage due to re-labeling. Thus, the signal that the decision-maker receives at a stage represents *all* his information about his true preferences that he may use at that stage.

When $N_t = 1$, the choice process is over, and the unique alternative in the opportunity set becomes the final choice of the decision-maker. The rank of the chosen alternative within the initial set is random. By p_N^k we denote the probability that the rank of this alternative is equal to k , and by the column vector \mathbf{p}_N we denote a distribution of the rank of the finally chosen alternative:

$$\mathbf{p}_N = (p_N^1, \dots, p_N^N)^T, \quad \sum_{k=1}^N p_N^k = 1.$$

Probabilities p_N^k are determined, amongst other factors, by the choice process and are to some extent under the control of the decision-maker.

In the standard expected utility theory framework, the risk attitude of an individual corresponds to the shape of his Bernoulli utility function. Typically, in evaluating the expected utility function, a risk-averse individual attaches higher weights to less favorable states of the world. In the limit, when the decision-maker becomes “extremely” risk-averse, he puts all the weight to the worst possible state. In our model, the utility levels that alternatives yield are exogenously fixed by their true ranks in the initial set of alternatives, and the decision-maker can only affect the probabilities \mathbf{p}_N , with which different ranks, i.e., states, occur. Hence, in our model, an extremely risk-averse decision-maker is an individual who minimizes the probability of choosing the worst alternative, i.e., the alternative with the highest rank. Consequently, his objective func-

⁸ Due to the finiteness of the initial set of alternatives, there is no necessity to make a distinction between a preference relation and representing it utility function.

tion F^{RA} to be maximized can be taken to be

$$F^{RA} = -p_N^N.$$

Similarly, in evaluating the expected utility function, a risk-seeking individual attaches higher weights to more favorable states of the world. In the limit, when the decision-maker becomes “extremely” risk-seeking, he puts all the weight to the best state. Hence, an extremely risk-seeking decision-maker is an individual who maximizes the probability of choosing the best alternative, and his objective function F^{RS} can be taken to be

$$F^{RS} = p_N^1.$$

All intermediate degrees of risk-aversion in our model correspond to the objective function of the expected utility form:

$$F = \sum_{k=1}^N p_N^k U(k),$$

where $U(k)$ is the corresponding weakly decreasing function of the rank. This U function itself can be viewed as individuals’ von Neumann-Morgenstern (expected) utility function over lotteries that the alternative of rank k yields. For notational convenience, we denote utility function U by the row utility vector

$$\mathbf{u}_N = (U(1), U(2), \dots, U(N)),$$

where the k th element of \mathbf{u}_N is the (expected) utility derived from alternative of rank k , i.e., $(\mathbf{u}_N)_{1,k} = U(k)$. The objective function of an individual with the utility function U can now be written in a compact form as the following inner product:

$$F = \mathbf{u}_N \mathbf{p}_N.$$

An extremely risk-averse decision-maker can be represented by a concave function $U^{RA}(k)$ with the corresponding utility vector:

$$\begin{aligned} \mathbf{u}_N^{RA} &= (0, \dots, 0, -1), \text{ i.e., } U^{RA}(k) = 0 \text{ for } k = 1, \dots, N - 1 \\ &\text{and } U^{RA}(N) = -1. \end{aligned}$$

To the contrary, an extremely risk-seeking decision-maker can be represented by a convex function $U^{RS}(k)$ with the corresponding utility vector:

$$\mathbf{u}_N^{RS} = (1, 0, \dots, 0), \text{ i.e., } U^{RS}(1) = 1 \text{ and } U^{RS}(k) = 0 \text{ for } k = 2, \dots, N.$$

In each selection stage, the decision-maker ranks subsets of alternatives from the current opportunity set in accordance with his objective function, and selects an opportunity set for the next stage.

3 Analysis of the model

We begin the analysis by arguing that if some alternatives are eliminated at a certain stage, these alternatives must be the alternatives that are perceived as being the worst at that stage. The proof of this claim is trivial and, therefore, is omitted.⁹ Now, let us suppose that the decision-maker has already derived his optimal choice procedure for all numbers $n \leq N$ of initial alternatives, and, therefore, knows all corresponding probabilities p_n . Then, when he faces a set of N alternatives and eliminates $m \geq 1$ worst-perceived alternatives, p_N can be written as follows:

$$p_N = \mathbf{D}_n^m p_{N-m},$$

where the probability transition matrix \mathbf{D}_n^m is derived in Appendix 1. Thus, when DM eliminates m_t worst alternatives at stage t , the resulting probability distribution p_N can be recursively written as follows:

$$p_N = \mathbf{D}_{N_1}^{m_1} p_{N_1-m_1} = \mathbf{D}_{N_1}^{m_1} \mathbf{D}_{N_2}^{m_2} p_{N_2-m_2} = \dots = \prod_t \mathbf{D}_{N_t}^{m_t},$$

where $N_{t+1} = N_t - m_t$. The problem of the decision-maker is to identify optimal values of m_t as a function of N_t and q that maximize his objective function $F = u_N p_N = u_N \prod_t \mathbf{D}_{N_t}^{m_t}$.¹⁰

3.1 Choice process of extremely risk-averse individual

First, we consider an extremely risk-averse decision-maker who maximizes $F^{\text{RA}} = -p_N^N$. His optimal choice process is derived in the following proposition.

Proposition 1 *An extremely risk-averse decision-maker eliminates only the worst-perceived alternative at a stage, i.e., $m = 1$. His pay-off is $F^{\text{RA}} = -(1 - q)^{N-1} / N$.*

Proof of Proposition 1 Eliminating $m \geq 2$ alternatives yields the pay-off $F^{\text{RA}}(m)$:

$$F^{\text{RA}}(m) = -p_N^N = -(\mathbf{D}_N^m p_{N-m})_{N,1} = - \sum_{j=1}^{N-m} ((\mathbf{D}_N^m)_{N,j} p_{N-m}^j).$$

Using the fact that $(\mathbf{D}_N^m)_{r,k} = 0$ for $r = N$ and all $k < N - m$, and $(\mathbf{D}_N^m)_{N,N-m} = (1 - q)(N - m) / N$ simplifies this expression to

$$F^{\text{RA}}(m) = (\mathbf{D}_N^m)_{N,N-m} p_{N-m}^{N-m} = - \frac{(1 - q)(N - m)}{N} p_{N-m}^{N-m}.$$

⁹ It is easy to show that by replacing an eliminated alternative by a not eliminated alternative with higher ranks improves the distribution of ranks in the next stage in the first-order stochastic dominance sense.

¹⁰ The described selection process is stationary so that m_t do not explicitly depend on t .

Similarly, eliminating $(m - 1)$ alternatives yields the pay-off $F^{RA}(m - 1)$. If, after eliminating $(m - 1)$ alternatives, the decision-maker eliminates one alternative at the next stage, his pay-off will be

$$\begin{aligned}
 \tilde{F}^{RA}(m) &= -(\mathbf{D}_N^{m-1} \mathbf{p}_{N-m+1})_{N,1} \\
 &= -(\mathbf{D}_N^{m-1} \mathbf{D}_{N-m+1}^1 \mathbf{p}_{N-m})_{N,1} \\
 &= -\sum_{j=1}^{N-m} (\mathbf{D}_N^{m-1} \mathbf{D}_{N-m+1}^{m-1})_{N,j} p_{N-m}^j \\
 \tilde{F}^{RA}(m) &= -(\mathbf{D}_N^{m-1} \mathbf{D}_{N-m+1}^1)_{N,N-m} p_{N-m}^{N-m} \\
 &= -\sum_{j=1}^{N-m+1} (\mathbf{D}_N^{m-1})_{N,j} (\mathbf{D}_{N-m+1}^1)_{j,N-m} p_{N-m}^{N-m} \\
 &= -(\mathbf{D}_N^{m-1})_{N,N-m+1} (\mathbf{D}_{N-m+1}^1)_{N-m+1,N-m} p_{N-m}^{N-m} \\
 &= -\frac{(1 - q^2)(N - m)}{N} p_{N-m}^{N-m}
 \end{aligned}$$

Clearly, $\tilde{F}^{RA}(m) - F^{RA}(m) > 0$ due to $q < 1$ and $p_{N-m}^{N-m} > 0$. This implies that eliminating $m \geq 2$ alternatives at a stage is strictly dominated by eliminating $(m - 1)$ alternatives at that stage followed by elimination of one alternative at the next stage. Thus, elimination of $m = 1$ alternatives at a stage is the unique optimal choice process. Then, the pay-off of the decision-maker is

$$F^{RA} = -\left(\prod_{N_t=N}^2 \mathbf{D}_{N_t}^1 \right)_{N,1} = -\prod_{N_t=N}^2 \frac{(1 - q)(N_t - 1)}{N_t} = -\frac{(1 - q)^{N-1}}{N}.$$

This ends the proof. □

In accordance with Proposition 1, an extremely risk-averse decision-maker follows the Single Worst Elimination process (see Ficco et al. 2006). He makes his final choice in $(N - 1)$ stages by eliminating a single alternative in each stage. The reason why it is the case is the following. In order to eliminate the worst alternative with the largest possible probability, the decision-maker wants to use as many imprecise signals, which imperfectly reveal his underlying preferences, as possible. Eliminating only one alternative at a stage allows him to use the maximum possible number $(N - 1)$ of such signals in order to identify and subsequently eliminate the worst alternative. As a result, the choice of an extremely risk-averse decision-maker shows the strongest preference for flexibility.

3.2 Choice process of extremely risk-seeking individual

Next, we consider an extremely risk-seeking decision-maker who maximizes $F^{RS} = p_N^1$. His optimal choice process is derived in the following proposition.

Proposition 2 *An extremely risk-seeking decision-maker eliminates all but one alternative at the very first stage, i.e., $m = N - 1$. His pay-off is $F^{RS} = q + (1 - q)/N$.*

Proof of Proposition 2 Eliminating $m \geq 2$ alternatives yields the pay-off $F^{RS}(m)$:

$$\begin{aligned}
 F^{RS}(m) &= p_N^1 = (\mathbf{D}_N^m \mathbf{p}_{N-m})_{1,1} = \sum_{j=1}^{N-m} \left((\mathbf{D}_N^m)_{1,j} p_{N-m}^j \right) = (\mathbf{D}_N^m)_{1,1} p_{N-m}^1 \\
 &= \left(1 - \frac{(1-q)m}{N} \right) p_{N-m}^1
 \end{aligned}$$

Eliminating $(m - 1)$ alternatives now and one alternative at the next stage yields $\tilde{F}^{RS}(m)$:

$$\begin{aligned}
 \tilde{F}^{RS}(m) &= (\mathbf{D}_N^{m-1} \mathbf{p}_{N-m+1})_{1,1} = (\mathbf{D}_N^{m-1} \mathbf{D}_{N-m+1}^1 \mathbf{p}_{N-m})_{1,1} \\
 &= \sum_{j=1}^{N-m} (\mathbf{D}_N^{m-1} \mathbf{D}_{N-m+1}^1)_{1,j} p_{N-m}^j \\
 \tilde{F}^{RS}(m) &= (\mathbf{D}_N^{m-1} \mathbf{D}_{N-m+1}^1)_{1,1} p_{N-m}^1 = \sum_{j=1}^{N-m+1} (\mathbf{D}_N^{m-1})_{1,j} (\mathbf{D}_{N-m+1}^1)_{j,1} p_{N-m}^1 \\
 &= (\mathbf{D}_N^{m-1})_{1,1} (\mathbf{D}_{N-m+1}^1)_{1,1} p_{N-m}^1 \\
 &= \left(q + (1-q) \frac{C_{N-1}^{m-1}}{C_N^{m-1}} \right) \left(q + (1-q) \frac{C_{N-m}^1}{C_{N-m+1}^1} \right) p_{N-m}^1 \\
 &= \left(1 - \frac{(1-q)(m-1)}{N} \right) \left(1 - \frac{1-q}{N-m+1} \right) p_{N-m}^1
 \end{aligned}$$

It is easy to see that $F^{RS}(m) - \tilde{F}^{RA}(m) > 0$ for all $m \geq 2$:

$$F^{RS}(m) - \tilde{F}^{RA}(m) = \frac{q(1-q)(m-1)}{N(N-m+1)} p_{N-m}^1 > 0.$$

This implies that eliminating more alternatives at a stage is strictly beneficial to the decision-maker. Thus, the only optimal choice process is to eliminate $m = N - 1$ alternatives at the very first stage. The pay-off of the decision-maker then is

$$F^{RS} = \left(\mathbf{D}_N^{N-1} \mathbf{p}_1 \right)_{1,1} = 1 - \frac{(1-q)(N-1)}{N} = q + \frac{1-q}{N}.$$

This ends the proof. □

In accordance with Proposition 2, an extremely risk-seeking decision-maker chooses the best perceived alternative right away, in the spirit of Simon’s satisficing criterion (Simon 1955). The intuition behind this result is as follows. Elimination of an alternative at a stage brings a risk that this alternative is the first best in the initial sample. The probability of this event decreases with the size of the current opportunity set. Consequently, the elimination of all but one alternative at the very first stage minimizes this probability and, consequently, maximizes the chances that the chosen alternative is the first best alternative. As a result, the choice of an extremely risk-seeking decision-maker shows no preference for flexibility whatsoever. He makes his choice immediately at the very first stage.

Propositions 1 and 2 shed some light on how individuals’ risk attitude affects their incentives to postpone the final choice. Individuals’ desire to avoid the worst alternative, i.e., risk aversion, creates individuals’ desire for flexibility. In contrast, individuals’ desire to choose the first-best alternative, i.e., risk seeking attitude, destroys it.

A natural question¹¹ that arises is whether the two procedures identified by Propositions 1 and 2 are the only optimal choice processes for an arbitrary utility function. The answer to this question is negative, and we demonstrate this in the following example.

Example 1 Let the decision-maker choose amongst $N = 4$ alternatives and let him have a utility function with the corresponding utility vector $\mathbf{u}_4^* = (1, y, x, 0)$, where $0 \leq x \leq y \leq 1$. For four alternative choice problems, there are also four different selection procedures:¹²

- (a) $m_1 = 3$, which is the procedure used by an extremely risk-seeking decision-maker. We denote this procedure by $S^{4 \rightarrow 1}$.
- (b) $m_1 = 2, m_2 = 1$, which is a procedure where two alternatives are eliminated in stage 1, and then one more alternative is eliminated in stage 2. We denote this procedure by $S^{4 \rightarrow 2 \rightarrow 1}$.
- (c) $m_1 = 1, m_2 = 2$, which is a procedure where one alternative is eliminated in stage 1, and then two more alternatives are eliminated in stage 2. We denote this procedure by $S^{4 \rightarrow 3 \rightarrow 1}$.
- (d) $m_1 = m_2 = m_3 = 1$, which is the procedure used by an extremely risk-averse decision-maker. We denote this procedure by $S^{4 \rightarrow 3 \rightarrow 2 \rightarrow 1}$.

It is a routine to calculate pay-offs of the individual $F^* = \mathbf{u}_4^* \prod_t \mathbf{D}_{N_t}^{m_t}$ for all four selection procedures and make comparisons for $q = 1/2$. The results of this investigation are shown in Fig. 1.

¹¹ We thank the anonymous referee for raising this point.

¹² In general, a choice from N alternatives allows for 2^{N-2} different selection procedures.

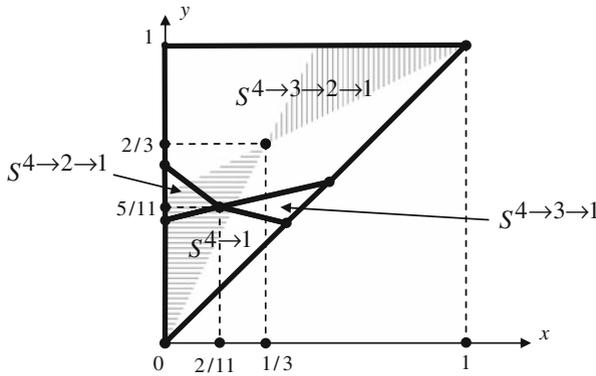


Fig. 1 Regions of parameters x and y , for which each selection procedure is optimal

Depending on values of x and y , i.e., on the utility function of the decision-maker, any choice process can be optimal. In general, individuals may display not only maximal or no preferences for flexibility but also all intermediate degrees (the procedure $S^{4 \to 2 \to 1}$ is an example where two alternatives are kept and the other two are eliminated). Moreover, the preference for flexibility of a given individual may depend on the number of alternatives in the current sample. If, for example, $S^{4 \to 3 \to 1}$ is the optimal choice process of an individual, he displays the maximal preference for flexibility facing $N = 4$ alternatives, and, at the same time, he displays no preference for flexibility facing $N = 3$ alternatives. It points out potential difficulties in trying to rank individuals' choice procedures with respect to degrees of preference for flexibility.

Interestingly, there exists a convex utility function with $x = 2/11$ and $y = 5/11$, for which all four procedures are pay-off equivalent. Thus, an individual with such a utility function is indifferent between all procedures, all choice processes are optimal, i.e., rationalizable, and his preference for flexibility remains undetermined.

Another interesting observation is that for values $y \geq 0.5(x + 1)$ and $x \geq 0.5y$, which correspond to concave and linear utility functions and which are denoted in Fig. 1 by vertically shaded area, procedure $S^{4 \to 3 \to 2 \to 1}$, i.e., Single Worst Elimination, is strictly better than any other procedure. To the contrary, for values $y \leq 0.5(x + 1)$ and $x \leq 0.5y$, which correspond to convex utility functions and which are denoted in Fig. 1 by horizontally shaded area, any of the procedures can be optimal, including $S^{4 \to 3 \to 2 \to 1}$. Thus, all risk-averse and risk-neutral individuals in our model make choices in the same way, whereas risk-seeking individuals may differ in a way they make their choices thereby displaying all possible degrees of preference for flexibility.

From a procedural point of view, Example 1 shows that the expected utility maximization rationalizes different choice processes, depending on individuals' attitudes towards risk. In particular, a sequential maximization of multiple rationales, the procedure that is often used in the literature on choice rationalization, is rationalizable but only for relatively risk-seeking individuals. If an individual is risk-averse, the sequential maximization of multiple rationales is not justified by the expected utility maximization.

3.3 Choice process of a risk-averse individual

A natural question that Example 1 poses is whether the optimality of elimination of the single worst for weakly risk-averse individuals can be generalized for all other values of N and q . Here we answer this question in the affirmative.

We begin with a decision-maker whose utility function $U^K(k)$ is weakly decreasing and weakly concave, and which is represented by a utility vector \mathbf{u}_N^K :

$$\mathbf{u}_N^K \equiv (0, \dots, 0, -1, -2, \dots, -K + 1, -K).$$

In other words, $(\mathbf{u}_N^K)_{1,k} = 0$ for $k = 1, \dots, N - K$ and $(\mathbf{u}_N^K)_{1,k} = -k - K + N$ for $k = N - K + 1, \dots, N$. For any $K = 1, \dots, N - 1$, $U^K(k)$ is zero for $(N - K)$ best ranks and is linearly decreasing over $(K + 1)$ worst ranks. In the following proposition, we derive the optimal choice process of such individuals.

Proposition 3 *A decision-maker with a utility vector \mathbf{u}_N^K , $K = 1, \dots, N - 1$ eliminates only the worst-perceived alternative at a stage, i.e., $m = 1$.*

Proof of Proposition 3 The proof follows the same reasoning as proofs of Propositions 1 and 2. Eliminating $m \geq 2$ alternatives yields the pay-off $F^K(m)$:

$$F^K(m) = \mathbf{u}_N^K \mathbf{p}_N = \mathbf{u}_N^K \mathbf{D}_N^m \mathbf{p}_{N-m}.$$

Eliminating $(m - 1)$ alternatives now and one alternative at the next stage yields $\tilde{F}^K(m)$:

$$\tilde{F}^K(m) = \mathbf{u}_N^K \mathbf{D}_N^{m-1} \mathbf{D}_{N-m+1}^1 \mathbf{p}_{N-m}.$$

We define:

$$\begin{aligned} E_N^K(m) &\equiv \frac{mC_N^m}{(1-q)q} \left(\tilde{F}^K(m) - F^K(m) \right) \\ &= \frac{mC_N^m}{(1-q)q} \mathbf{u}_N^K \left(\mathbf{D}_N^{m-1} \mathbf{D}_{N-m+1}^1 - \mathbf{D}_N^m \right) \mathbf{p}_{N-m}. \end{aligned}$$

In Appendix 2 we show that $E_N^K(m)$ can be written as follows:

$$E_N^K(m) = \mathbf{g}(K, m, N) \mathbf{p}_{N-m},$$

where the row vector $\mathbf{g}(K, m, N) = (g_1, \dots, g_{N-m})$ has the following components:

$$\begin{aligned} g_k(K, m, N) &= k \left(\sum_{r=N-K+1}^N (r + K - N) (C_{r-1}^{r-k-1} C_{N-r}^{m-r+k} - C_{r-1}^{r-k} C_{N-r}^{m-1-r+k}) - C_{k-N+K}^0 C_N^{m-1} \right). \end{aligned}$$

In the rest of the proof that we have placed in Appendix 3, we show that $g(K, m, N) \geq 0$ and for some k $g_k(K, m, N) > 0$, for any $N \geq 3, m = 2, \dots, N - 1$, and $K = 1, \dots, N - 1$. Together with $p_{N-m} \gg 0$, which holds for any $N - m \geq 1$ and $q \in (0, 1)$, this implies that $E_N^K(m) > 0$ and, hence, $\tilde{F}^K(m) > F^K(m)$. Therefore, eliminating $m \geq 2$ alternatives at a stage is strictly dominated by eliminating $(m - 1)$ alternatives at that stage followed by elimination of one alternative at the next stage. Thus, elimination of $m = 1$ alternatives at a stage is the unique optimal choice process. \square

In accordance with Proposition 3, not only extremely risk-averse individuals display the maximal desire for flexibility, but also individuals with weakly decreasing utility function that is linear in rank for some number of worst ranks and constant for all the other remaining (best) ranks.

One interesting property of the system of utility vectors $\{u_N^K\}, K = 1, \dots, N - 1$, is that they are linearly independent. Together with the constant vector $u_N^0 \equiv (1, \dots, 1)$, vectors $\{u_N^K\}, K = 0, \dots, N - 1$ form a basis so that a utility vector u_N that represents an arbitrary utility function $U(k)$ can be uniquely represented by its spectrum $\alpha \equiv (\alpha_0, \dots, \alpha_{N-1})$:

$$u_N = \sum_{K=0}^{N-1} \alpha_i u_N^K.$$

Using properties of $\{u_N^K\}$, in the following proposition we derive the optimal choice process of an individual with an arbitrary concave utility function.

Proposition 4 *A decision-maker with an arbitrary concave utility function $U(k)$ eliminates only the worst-perceived alternative at a stage, i.e., $m = 1$.*

Proof of Proposition 4 It can be directly verified that an arbitrary utility function $U(k)$ yields the following spectrum of the corresponding vector u_N :

$$\alpha_0 = U(1),$$

$$\alpha_K = 2U(N - K) - U(N - K + 1) - U(N - K - 1), \quad K = 1, \dots, N - 2, \quad \text{and}$$

$$\alpha_{N-1} = U(1) - U(2).$$

Hence, an arbitrary concave and decreasing in rank utility function $U(k)$ has strictly positive spectrum for $K = 1, \dots, N - 1$. The rest of the proof follows from Proposition 3. Eliminating $m \geq 2$ alternatives yields the pay-off $F(m)$:

$$F(m) = u_N p_N = \sum_{K=0}^{N-1} \alpha_i (u_N^K p_N) = \sum_{K=0}^{N-1} \alpha_i (u_N^K D_N^m p_{N-m}) = \sum_{K=0}^{N-1} \alpha_i F^K(m).$$

Eliminating $(m - 1)$ alternatives now and one alternative at the next stage yields $\tilde{F}(m)$:

$$\begin{aligned} \tilde{F}(m) &= \mathbf{u}_N \mathbf{D}_N^{m-1} \mathbf{D}_{N-m+1}^1 \mathbf{p}_{N-m} = \sum_{K=0}^{N-1} \alpha_i \left(\mathbf{u}_N^K \mathbf{D}_N^{m-1} \mathbf{D}_{N-m+1}^1 \mathbf{p}_{N-m} \right) \\ &= \sum_{K=0}^{N-1} \alpha_i \tilde{F}^K(m). \end{aligned}$$

Taking the difference $\tilde{F}(m) - F(m)$ yields:

$$\tilde{F}(m) - F(m) = \sum_{K=0}^{N-1} \alpha_i \left(\tilde{F}^K(m) - F^K(m) \right)$$

In accordance with Proposition 3, $\tilde{F}^K(m) > F^K(m)$ for all $K = 1, \dots, N - 1$. On the other hand, it is obvious that $\tilde{F}^0(m) = F^0(m) = 1$ because an individual with the utility vector $\mathbf{u}_N^0 \equiv (1, \dots, 1)$ gets a unit of utility from any procedure. This implies that $\tilde{F}(m) - F(m) > 0$ and ends the proof of the proposition. \square

Proposition 4 shows that all risk-neutral and risk-averse decision-makers display the same, maximal preference for flexibility—they eliminate only one alternative at a stage. This procedure is the only rationalizable choice process as it is strictly better than any other choice process. By the continuity argument, it also remains optimal for small perturbations of the utility function. It can be seen in Example 1 that the region of parameters' values for which procedure $S^{4 \rightarrow 3 \rightarrow 2 \rightarrow 1}$, i.e., Single Worst Elimination, is optimal is wider than the region of concave utilities (vertically shaded area). Thus, the optimality of this choice process goes beyond concave utilities.

The reason that non-concave utility function might lead to more aggressive elimination of alternatives is that the spectrum of the corresponding utility vector must necessarily have a negative entry $\alpha_K = 2U(N - K) - U(N - K + 1) - U(N - K - 1)$ for some $K = 1, \dots, N - 2$. For such a utility vector, there will be at least one negative term in the decomposition of $\tilde{F}(m) - F(m)$. This has a simple interpretation: local convexity at rank K of the utility function $U(k)$ creates the opposite incentive to eliminate all but one alternative. The optimal procedure then results from the trade-off between the desire to postpone the choice (due to concave segments of the utility function) and the desire to make a rapid choice (due to convex segments). Consequently, individuals with more local convexities, and with larger convexities defined by the absolute value of α_k , will eliminate alternatives more aggressively and converge in their choice behavior in the limit to the extremely risk-seeking individuals.

4 Conclusion

In this paper, we identify risk attitude of an individual as the only factor which, in the absence of learning, determines preference for flexibility. This result is obtained

by rationalizing choice processes of individuals in the presence of imperfect recall and under a particular information structure. Risk-averse individuals have the highest incentive to postpone the choice and, therefore, they show the strongest preference for flexibility: they eliminate only one alternative at a time. This suggests that risk-averse individuals will generally take the longest time to make a decision and, therefore, will be more sensitive to discounting. To the contrary, extremely risk-seeking individuals have no incentives to postpone the choice and, therefore, they show no desire for flexibility whatsoever: they make choices immediately and seem to be not subject to discounting. All other individuals, who are neither risk-seeking nor risk-averse, can display all possible degrees of desire for freedom. These theoretical predictions can be experimentally tested in a laboratory environment, as well as in field experiments. From a normative point of view, our results justify Single Worst Elimination as a choice process of risk-averse individuals and ask for further investigation of choices that are induced by this selection rule.

The results also suggest that the analysis of individuals' choice behavior does not allow us to make a distinction between underlying risk neutrality and risk aversion. This implies that the procedural approach to individuals' choice that we have taken seems to overestimate individuals' degree of risk aversion relative to the aversion that they show in making a choice by evaluating lotteries. This bias could also be empirically tested.

Our results follow from the corresponding inequalities that are strict. We can, therefore, claim that these results are robust to small perturbations of the model's primitives and underlying modeling assumptions. Hence, even if learning is possible but is very limited, if signals' accuracies across stages are different but the differences are small, and if there are costs associated with the timing of the final decision (discounting) but these costs are not large, individuals' degree of risk aversion will still positively affect individuals' desire for flexibility.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

Appendix 1: Derivations of D_N^1 and its generalization D_N^m

If the rating of alternatives is correct, then $p_N = A_N^1 p_{N-1}$ where A_N^1 is the $N \times (N - 1)$ matrix with a generic (r, k) -element $(A_N^1) = \delta_{r,k}$. If the rating of alternatives is random, then $p_N = B_N^1 p_{N-1}$ where B_N^1 is the $N \times (N - 1)$ matrix with a generic (r, k) -element $(B_N^1)_{r,k} = C_{r-1}^{r-k} C_{N-r}^{1-r-k} / N$:

$$A_N^1 = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \\ 0 & 0 & \dots & 0 \end{pmatrix} \quad \text{and} \quad B_N^1 = \frac{1}{N} \begin{pmatrix} N-1 & 0 & \dots & 0 & 0 \\ 1 & N-2 & \dots & 0 & 0 \\ 0 & 2 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & N-2 & 1 \\ 0 & 0 & \dots & 0 & N-1 \end{pmatrix}$$

Here C_y^x is the binomial coefficient, which we define to be zero if $x < 0$, $y < 0$, or $x > y$. Hence, if the ranking is correct with probability q and random with the remaining probability $(1 - q)$, \mathbf{p}_N can be written as $\mathbf{p}_N = \mathbf{D}_N^1 \mathbf{p}_{N-1}$ where $\mathbf{D}_N^1 \equiv q\mathbf{A}_N^1 + (1 - q)\mathbf{B}_N^1$. Similarly, when the decision-maker eliminates m worst alternatives at stage t , we define $N \times (N - m)$ matrices \mathbf{A}_N^m and \mathbf{B}_N^m as $(\mathbf{A}_N^m)_{rk} = \delta_{rk}$ and $(\mathbf{B}_N^m)_{r,k} = C_{r-1}^{r-k} C_{N-r}^{m-r+k} / C_N^m$.

It can be directly verified that the $N \times (N - m)$ matrix $\mathbf{D}_N^m \equiv q\mathbf{A}_N^m + (1 - q)\mathbf{B}_N^m$ with generic elements $(\mathbf{D}_N^m)_{r,k} = q\delta_{r,k} + (1 - q)C_{r-1}^{r-k} C_{N-r}^{m-r+k} / C_N^m$ is indeed the probability transition matrix so that $\mathbf{p}_N = \mathbf{D}_N^m \mathbf{p}_{N-m}$ by definition.

Appendix 2: Derivations of $E_N^K(m)$

First, using the definition of \mathbf{D}_N^m we show that each column of \mathbf{D}_{N-m+1}^1 has at most two nonzero entries:

$$\begin{aligned} (\mathbf{D}_{N-m+1}^1)_{r,r} &= q + \frac{(1 - q)(N - m + 1 - r)}{M - m + 1} = 1 - \frac{(1 - q)r}{N - m + 1} \quad \text{and} \\ (\mathbf{D}_{N-m+1}^1)_{r,r-1} &= \frac{(1 - q)(r - 1)}{N - m + 1} \end{aligned}$$

Then, we calculate the product $\mathbf{D}_N^{m-1} \mathbf{D}_{N-m+1}^1$:

$$\begin{aligned} (\mathbf{D}_N^{m-1} \mathbf{D}_{N-m+1}^1)_{r,k} &= (\mathbf{D}_N^{m-1})_{r,k+1} (\mathbf{D}_{N-m+1}^1)_{k+1,k} + (\mathbf{D}_N^{m-1})_{r,k} (\mathbf{D}_{N-m+1}^1)_{k,k} \\ &= \left(q\delta_{r,k+1} + \frac{(1 - q)}{C_N^{m-1}} C_{r-1}^{r-k-1} C_{N-r}^{m-r+k} \right) \left(\frac{(1 - q)k}{N - m + 1} \right) \\ &\quad + \left(q\delta_{r,k} + \frac{(1 - q)}{C_N^{m-1}} C_{r-1}^{r-k} C_{N-r}^{m-1-r+k} \right) \left(1 - \frac{(1 - q)k}{N - m + 1} \right) \end{aligned}$$

Using the above expression, $E_N^K(m)$ can now be written as follows:

$$E_N^K(m) = \frac{m C_N^m}{(1 - q)q} \mathbf{u}_N^K \left(\mathbf{D}_N^{m-1} \mathbf{D}_{N-m+1}^1 - \mathbf{D}_N^m \right) \mathbf{p}_{N-m} = \mathbf{g}(K, m, N) \mathbf{p}_{N-m},$$

where we define $\mathbf{g}(K, m, N)$ as follows:

$$\mathbf{g}(K, m, N) = (g_1, \dots, g_{N-m}) \equiv \frac{m C_N^m}{(1 - q)q} \mathbf{u}_N^K \left(\mathbf{D}_N^{m-1} \mathbf{D}_{N-m+1}^1 - \mathbf{D}_N^m \right).$$

The k th component g_k of g can be simplified in the following way:

$$\begin{aligned}
 g_k &= \frac{mC_N^m}{(1-q)q} \sum_{r=N-K+1}^N (r+K-N) \left((\mathbf{D}_N^m)_{r,k} - (\mathbf{D}_N^{m-1} \mathbf{D}_{N-m+1}^1)_{r,k} \right) \\
 &= \frac{mC_N^m}{(1-q)q} \sum_{r=N-K+1}^N (r+K-N) \left(q\delta_{r,k} + \frac{1-q}{C_N^m} C_{r-1}^{r-k} C_{N-r}^{m-r+k} \right) \\
 &\quad - \frac{mC_N^m}{(1-q)q} \sum_{r=N-K+1}^N (r+K-N) \left(q\delta_{r,k+1} + \frac{(1-q)}{C_N^{m-1}} C_{r-1}^{r-k-1} C_{N-r}^{m-r+k} \right) \\
 &\quad \times \frac{(1-q)k}{N-m+1} - \frac{mC_N^m}{(1-q)q} \sum_{r=N-K+1}^N (r+K-N) \\
 &\quad \times \left(q\delta_{r,k+1} + \frac{(1-q)}{C_N^{m-1}} C_{r-1}^{r-k} C_{N-r}^{m-1-r+k} \right) \left(1 - \frac{(1-q)k}{N-m+1} \right)
 \end{aligned}$$

and then:

$$\begin{aligned}
 g_k(K, m, N) &= \frac{mC_N^m}{(1-q)q} \sum_{r=N-K+1}^N (r+K-N) \\
 &\quad \times \left(\frac{(1-q)}{C_N^m} C_{r-1}^{r-k} C_{N-r}^{m-r+k} - \frac{(1-q)}{C_N^{m-1}} C_{r-1}^{r-k} C_{N-r}^{m-1-r+k} \right) \\
 &\quad + \frac{mC_N^m}{(1-q)q} \sum_{r=N-K+1}^N (r+K-N) \left(q\delta_{r,k} + \frac{(1-q)}{C_N^{m-1}} C_{r-1}^{r-k} C_{N-r}^{m-1-r+k} \right) \\
 &\quad \times \frac{(1-q)k}{N-m+1} - \frac{mC_N^m}{(1-q)q} \sum_{r=N-K+1}^N (r+K-N) \\
 &\quad \times \left(q\delta_{r,k+1} + \frac{(1-q)}{C_N^m} C_{r-1}^{r-k-1} C_{N-r}^{m-r+k} \right) \frac{(1-q)k}{N-m+1} \\
 &= \frac{1}{q} \sum_{r=N-K+1}^N (r+K-N) \left(mC_{r-1}^{r-k} C_{N-r}^{m-r-k} \right. \\
 &\quad \left. - (N-m+q)C_{r-1}^{r-k} C_{N-r}^{m-1-r+k} \right) \\
 &\quad + \frac{k}{q}(1-q) \sum_{r=N-K+1}^N (r+K-N) \left(C_{r-1}^{r-k} C_{N-r}^{m-1-r+k} \right)
 \end{aligned}$$

$$\begin{aligned}
 & - C_{r-1}^{r-k-1} C_{N-r}^{m-r+k} \Big) - C_{k-N+K}^0 C_N^{m-1} \\
 & = k \left(\sum_{r=N-K+1}^N (r+K-N) (C_{r-1}^{r-k-1} C_{N-r}^{m-r+k} \right. \\
 & \quad \left. - C_{r-1}^{r-k} C_{N-r}^{m-1-r+k}) - C_{k-N+K}^0 C_N^{m-1} \right).
 \end{aligned}$$

Appendix 3: Proof of $g(K, m, N) \geq 0$

We define

$$h_k(K, m, N) \equiv \sum_{r=N-K+1}^N (r+K-N) \left(C_{r-1}^k C_{N-r}^{m-r+k} - C_{r-1}^{k-1} C_{N-r}^{m-1-r+k} \right),$$

so that $g_k(K, m, N) = k(h_k(K, m, N) - C_{k-N+K}^0 C_N^{m-1})$, and consider $\Delta_K h_k(K, m, N)$:

$$\begin{aligned}
 \Delta_K h_k(K, m, N) & \equiv h_k(K+1, m, N) - h_k(K, m, N) \\
 & = \sum_{r=N-K}^N \left(C_{r-1}^k C_{N-r}^{m-r+k} - C_{r-1}^{k-1} C_{N-r}^{m-1-r+k} \right).
 \end{aligned}$$

For $z = 0$, $\Delta_K h_k(K, m, N)$ can be written as

$$\begin{aligned}
 \Delta_K h_k(K, m, N) & = \sum_{r=N-K}^N \left(C_{r-1}^k C_{N-r-z}^{m-r+k} - C_{r-1}^{k-1} C_{N-r-z}^{m-1-r+k} \right) \\
 & \quad + C_{N-K-1}^k \sum_{x=0}^{z-1} C_{K-x}^{N-m-k}.
 \end{aligned}$$

Suppose that the above equality holds for some $z \geq 0$. Then, it also holds for $(z+1)$ as:

$$\begin{aligned}
 \Delta_k h_k & = \sum_{r=N-K}^N \left(C_{r-1}^k C_{N-r-z}^{m-r+k} - C_{r-1}^{k-1} C_{N-r-z}^{m-1-r+k} \right) + C_{N-K-1}^k \sum_{x=0}^{z-1} C_{K-x}^{N-m-k} \\
 & = \sum_{r=N-K-1}^{N-1} \left(C_r^k C_{N-r-(z+1)}^{m-r+k} - C_r^{k-1} C_{N-r-(z+1)}^{m-1-r+k} \right) + C_{N-K-1}^k \sum_{x=0}^{z-1} C_{K-x}^{N-m-k}
 \end{aligned}$$

$$\begin{aligned}
 &= \sum_{r=N-K-1}^N \left((C_{r-1}^k + C_{r-1}^{k-1}) C_{N-r-(z+1)}^{m-r+k} - (C_{r-1}^{k-1} + C_{r-1}^{k-2}) C_{N-r-(z+1)}^{m-1-r+k} \right) \\
 &\quad + C_{N-K-1}^k \sum_{x=0}^{z-1} C_{K-x}^{N-m-k} \\
 &= \sum_{r=N-K-1}^N \left(C_{r-1}^k C_{N-r-(z+1)}^{m-r+k} - C_{r-1}^{k-1} C_{N-r-(z+1)}^{m-1-r+k} + C_{r-1}^{k-1} C_{N-r-z}^{m-r+k} \right. \\
 &\quad \left. - C_r^{k-1} C_{N-r-z-1}^{m-1-r+k} \right) + C_{N-K-1}^k \sum_{x=0}^{z-1} C_{K-x}^{N-m-k} \\
 \Delta_k h_k &= \sum_{r=N-K-1}^N \left(C_{r-1}^k C_{N-r-(z+1)}^{m-r+k} - C_{r-1}^{k-1} C_{N-r-(z+1)}^{m-1-r+k} \right) + \sum_{r=N-K-2}^{N-1} \\
 &\quad C_r^{k-1} C_{N-r-z-1}^{m-r-1+k} - \sum_{r=N-K-1}^N C_r^{k-1} C_{N-r-z-1}^{m-1-r+k} + C_{N-K-1}^k \sum_{x=0}^{z-1} C_{K-x}^{N-m-k} \\
 &= \sum_{r=N-K-1}^N \left(C_{r-1}^k C_{N-r-(z+1)}^{m-r+k} - C_{r-1}^{k-1} C_{N-r-(z+1)}^{m-1-r+k} \right) \\
 &\quad + C_{N-K-2}^{k-1} C_{K+1-z}^{m-N+K+1+k} + C_{N-K-1}^k \sum_{x=0}^{z-1} C_{K-x}^{N-m-k} \\
 &= \sum_{r=N-K}^N \left(C_{r-1}^k C_{N-r-(z+1)}^{m-r+k} - C_{r-1}^{k-1} C_{N-r-(z+1)}^{m-1-r+k} \right) + C_{N-K-1}^k \sum_{x=0}^{(z+1)-1} C_{K-x}^{N-m-k}
 \end{aligned}$$

By induction, it holds for $z = K$. Hence, for all $k = 1, \dots, N - m$:

$$\begin{aligned}
 \Delta_K h_k &= C_{N-K-1}^k \sum_{x=0}^K C_{K-x}^{N-m-k} = C_{N-K-1}^k \left(\sum_{x=0}^{K-1} C_{K-x}^{N-m-k} + C_1^{N-m-k+1} \right) \\
 &= C_{N-K-1}^k \left(\sum_{x=0}^{K-2} C_{K-x}^{N-m-k} + C_2^{N-m-k+1} \right) = \dots = C_{N-K-1}^k \\
 &\quad \times \left(\sum_{x=0}^{K-K} C_{K-x}^{N-m-k} + C_K^{N-m-k+1} \right) = C_{N-K-1}^k C_{K+1}^{N-m-k+1}
 \end{aligned}$$

Now $g_k(K, m, N)$ can be written as follows:

$$\begin{aligned}
 g_k(K, m, N) &= k \left(h_k(K, m, N) - C_{k-N+K}^0 C_N^{m-1} \right) \\
 &= k \left(\sum_{z=1}^{K-1} \Delta_k h_k(z, m, N) + h_k(1, m, N) - C_{k-N+K}^0 C_N^{m-1} \right) \\
 &= k \left(\sum_{z=0}^{K-1} C_{N-z-1}^k C_{z+1}^{N-m-k+1} - C_{k-N+K}^0 C_N^{m-1} \right)
 \end{aligned}$$

It is easy to see that $g_k(K, m, N) \geq 0$ for all $K < N - k$ as $C_{k-N+K}^0 = 0$ in this case. Moreover, $g_k(K, m, N) > 0$ for $k = N - m$.

In the rest of this proof we show that $g_k(K, m, N) \geq 0$ for all $K \geq N - k$. It is easy to see that $g_k(K, m, N)$ is independent of K for $K \geq N - k$:

$$\begin{aligned}
 (g_k(K, m, N))_{K \geq N-k} &= k \left(\sum_{z=0}^{K-1} C_{N-z-1}^k C_{z+1}^{N-m-k+1} - C_N^{m-1} \right) \\
 &= k \left(\sum_{z=N-k-m}^{N-k-1} C_{N-z-1}^k C_{z+1}^{N-m-k+1} - C_N^{m-1} \right).
 \end{aligned}$$

Hence, for all $K \geq N - k$, $g_k(K, m, N) = kt(k, m, N)$, where

$$t(k, m, N) = \sum_{z=N-k-m}^{N-k-1} C_{N-z-1}^k C_{z+1}^{N-m-k+1} - C_N^{m-1}.$$

We will show that $t(k, m, N)$ is independent of k . To this end, we consider the difference $\Delta_k t(k, m, N) \equiv t(k + 1, m, N) - t(k, m, N)$:

$$\begin{aligned}
 \Delta_k t(k, m, N) &\equiv t(k + 1, m, N) - t(k, m, N) \\
 &= \sum_{z=N-k-m-1}^{N-k-2} C_{N-z-1}^{k+1} C_{z+1}^{N-m-k-1+1} - \sum_{z=N-k-m}^{N-k-1} C_{N-z-1}^k C_{z+1}^{N-m-k+1} \\
 &= \sum_{z=N-k-m}^{N-k-1} \left(C_{N-z}^{k+1} C_z^{N-m-k} - C_{N-z-1}^k C_{z+1}^{N-m-k+1} \right)
 \end{aligned}$$

For $x = 1$, $\Delta_k t(k, m, N)$ can be written as

$$\Delta_k t(k, m, N) = \sum_{z=N-k-m}^{N-k-x} \left(C_{N-z+1-x}^{k+1} C_z^{N-m-k} - C_{N-z-x}^k C_{z+1}^{N-m-k+1} \right)$$

Suppose that the above equality holds for some $x \geq 1$. Then, it also holds for $(x + 1)$ as:

$$\begin{aligned} \Delta_k t(k, m, N) &= \sum_{z=N-k-m}^{N-k-x} \left(C_{N-z+1-x}^{k+1} C_z^{N-m-k} - C_{N-z-x}^k C_{z+1}^{N-m-k+1} \right) \\ &= \sum_{z=N-k-m}^{N-k-x} \left(C_{N-z+1-(x+1)}^{k+1} C_z^{N-m-k} - C_{N-z+1-(x+1)}^k C_z^{N-m-k+1} \right) \\ &= \sum_{z=N-k-m}^{N-k-(x+1)} C_{N-z+1-(x+1)}^{k+1} C_z^{N-m-k} - \sum_{z=N-k-m+1}^{N-k-x} C_{N-z+1-(x+1)}^k C_z^{N-m-k+1} \\ &= \sum_{z=N-k-m}^{N-k-(x+1)} \left(C_{N-z+1-(x+1)}^{k+1} C_z^{N-m-k} - C_{N-z+(x+1)-(x+1)}^k C_{z+1}^{N-m-k+1} \right) \end{aligned}$$

By induction, it holds for $x = m$ and, therefore,

$$\begin{aligned} \Delta_k t(k, m, N) &= \sum_{z=N-k-m}^{N-k-m} \left(C_{N-z+1-m}^{k+1} - C_z^{N-m-k} C_{N-z-m}^k C_{z+1}^{N-m-k+1} \right) \\ &= C_{k+1}^{k+1} - C_k^k = 0. \end{aligned}$$

Thus, $t(k, m, N) = t(N - m, m, N)$ which is

$$\begin{aligned} t(N - m, m, N) &= \sum_{z=0}^{m-1} C_{N-z-1}^{N-m} C_{z+1}^1 - C_N^{m-1} \\ &= \sum_{z=1}^m C_{N-z}^{m-z} C_z^1 - \sum_{z=1}^m C_{N-z}^{m-z} = \sum_{z=2}^m C_{N-z}^{m-z} C_{z-1}^1 > 0. \end{aligned}$$

Hence, $t(k, m, N) = t(N - m, m, N) > 0$, $g_k(K, m, N) = kt(k, m, N) > 0$ for all $K \geq N - k$, and, consequently, $g(K, m, N) \geq 0$ with $g_k(K, m, N) > 0$ for some k .

References

Ficco, S., et al.: A theory of procedurally rational choice: optimization without evaluation. Tinbergen Institute discussion paper # 06-001/1 (2006)
 Kalai, G., Rubinstein, A., Spiegel, R.: Rationalizing choice functions by multiple rationale. *Econometrica* **70**(6), 2481–2488 (2002). doi:10.1111/1468-0262.00382
 Koopmans, T.C.: On the flexibility of future preferences. In: Shelly, M.W., Bryan, G.L. (eds.) *Human Judgments and Optimality*. New York, 246, 254 (1964)
 Krause, A.: Ranking opportunity sets in a simple intertemporal framework. *Econ Theory* **35**(1), 147–154 (2008). doi:10.1007/s00199-007-0243-6

- Kreps, D.M.: A representation theorem for preference for flexibility. *Econometrica* **47**, 565–577 (1979). doi:[10.2307/1910406](https://doi.org/10.2307/1910406)
- Lipman, B.L.: Information processing and bounded rationality: a survey. *Can J Econ Can Econ Assoc* **28**(1), 42–67 (1995)
- Manzini, P., Mariotti, M.: Sequentially rationalizable choice. *Am Econ Rev* **97**(5), 1824–1839 (2007). doi:[10.1257/aer.97.5.1824](https://doi.org/10.1257/aer.97.5.1824)
- Miller, G.A.: The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol Rev* **63**, 81–97 (1956). doi:[10.1037/h0043158](https://doi.org/10.1037/h0043158)
- Pattanaik, P.K., Xu, Y.: On ranking opportunity sets in terms of freedom of choice. *Rech Econ Louvain* **56**, 383–390 (1990)
- Puppe, C.: Freedom of choice and rational decisions. *Soc Choice Welf* **12**, 137–153 (1995). doi:[10.1007/BF00179829](https://doi.org/10.1007/BF00179829)
- Puppe, C.: An axiomatic approach to preferences for freedom of choice. *J Econ Theory* **68**, 174–199 (1996). doi:[10.1006/jeth.1996.0009](https://doi.org/10.1006/jeth.1996.0009)
- Simon, H.: A behavioral model of rational choices. *Q J Econ* **69**, 99–118 (1955). doi:[10.2307/1884852](https://doi.org/10.2307/1884852)
- Tversky, A.: Elimination by aspects: a theory of choice. *Psychol Rev* **79**, 281–299 (1972). doi:[10.1037/h0032955](https://doi.org/10.1037/h0032955)