Attraction-Effect Heuristics*

Nejat Anbarci† Kang Rong‡

October 25, 2016

Abstract

The Attraction Effect refers to an inferior option a's ability to increase the attractiveness of another alternative b after a is added to a choice set, such that a is exclusively dominated by b in the choice set. We consider three decision-making heuristics of which procedures incorporate the AE at different degrees. In the "Reference-independent Deterministic Choice" ("RIDC")'s procedure, the decision maker simply selects the alternative with the largest lower contour set (LCS). In the "Exogenous Reference-dependent Random Choice" ("Exogenous-RDRC")'s procedure, the decision maker draws a reference alternative, x, among dominated alternatives randomly and only considers alternatives that Pareto-dominate x, i.e., dominate x in all attributes. Then the decision maker randomly selects with equal probabilities one of the Pareto alternatives that Pareto-dominate x as her choice outcome. In the "Endogenous Reference-dependent Random Choice" ("Endogenous-RDRC")'s procedure, the decision maker selects any dominated alternative a as the reference alternative with a probability proportional to the number of Pareto alternatives in the upper contour set (UCS) of a, and then randomly selects with equal probabilities one of the Pareto alternatives that Pareto-dominate a as her choice. These heuristics' outcomes differ from each other in terms of the extent that they incorporate the Attraction Effect: RIDC incorporates it most strongly and the Exogenous-RDRC incorporates it the least. In addition, we provide axiomatic characterizations of the outcome sets of these heuristics to help uncover their most salient properties. We are also able to link the outcome sets of RIDC and Endogenous-RDRC to the Nash product as more alternatives are selected uniformly from a convex and compact set.

Keywords: The attraction effect, lower and upper contour sets, multiple attributes, incomplete preferences, reference alternative, random choice, the Nash product.

---

*We would like to thank Ivan Balbuzanov, Zhijun Chen, Chris Dubelaar, Nick Feltovich, Erkut Ozbay, Birendra Rai, Alex Teytelboym, Jun Xiao, and the participants at conference and seminar presentations for their comments and suggestions.

†Department of Economics, Deakin University, Burwood, Australia. Email: nejat.anbarci@deakin.edu.au

‡Shanghai University of Finance and Economics, Shanghai, China. Email: rong.kang@mail.shufe.edu.cn

JEL Classification Numbers: D11, D81.

1 Introduction

The *Attraction Effect* (AE) refers to an inferior option a's ability to increase the attractiveness of another alternative b after a is added to a choice set, such that a is exclusively dominated by b in the choice set. Since its introduction to the literature in the early 1980s by Huber, Payne and Puto (1982), the AE or the 'Asymmetric Dominance Effect' has become immensely popular in marketing and psychology literatures - as well as in other fields - that consider environments with a relatively small number of alternatives which have at least two attributes. (See, among others, Ariely and Wallsten, 1995, Doyle, O'Connor, Reynolds and Bottomley, 1999, Herne, 1997, Huber and Puto, 1983, Mourali, Bockenholt and Laroche, 2007, O'Curry and Pitts, 1995, Schwarzkopf, 2003, Schwartz and Chapman, 1999, among others.) Lately the AE has become prominent in economics as well (see the Relevant Literature section).

To illustrate the AE in detail, suppose a decision maker were equally likely to choose between two PCs, i.e., two personal computers, A and B of which main attributes - such as 'price' and 'processor speed' - are easily comparable, where A has a better price and worse speed than B. Now, consider a third option C such that it is clearly inferior to B (thus, C being more expensive and slower than B) but not inferior to A (thus, C being more expensive than A while having faster speed than A), i.e., C is Pareto-dominated by B but not by A. Then, while the inferior item C is not likely to chosen by the decision maker, it is significantly more likely that the decision maker will choose item B than A in the presence of C. Thus, dominating C makes B more attractive or prominent than it is in the absence of C and consequently more preferable than A as well (we will refer to this setup as *Example 1*). Suppose that one adds two more options, D and D', such that D is inferior to A only and D' is inferior to B only (by that token one could continue adding more such alternatives such that overall half of these alternatives are exclusively Pareto-dominated by A and the remaining half exclusively Pareto-dominated by B); likewise, suppose one adds several other options to the choice set consisting of A, B and C, such that these new alternatives are jointly dominated by both A and B. Presumably these symmetric additions would still leave B as the option more preferable than A, although perhaps not as much as it was the case with the choice set consisting of only A, B and C. In this paper, we will refer to options such as B, whose attractiveness gets a boost from the presence of the lergest number of other options that are inferior to it (i.e., to options with the largest 'lower contour set' (LCS) in a choice set) as *AE alternatives*.[1]

[1]In the same vein, Higgins' (1997) 'regulatory focus theory' (with Mourali et al.'s, 2007, empirical findings) provides insights. In the regulatory focus theory, the '*promotion* focus' is concerned with capturing

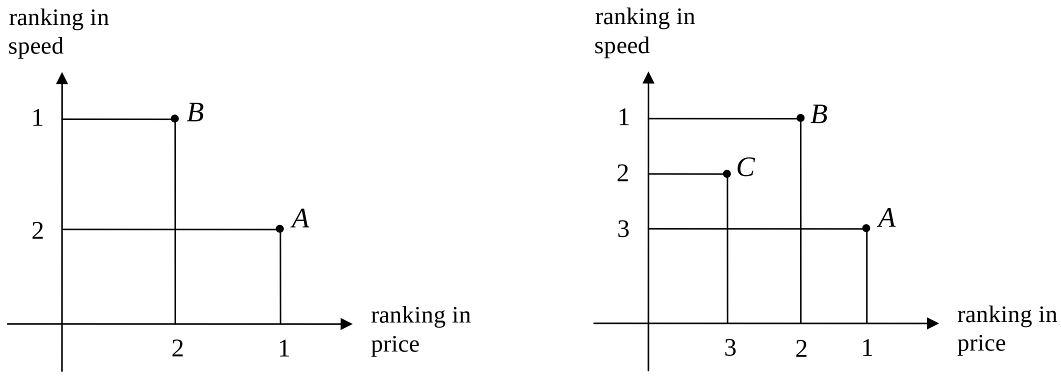

Figure 1: Example 1.

Despite its empirical prominence, the AE is considered an 'anomaly' by the well-studied and stylized *standard paradigm of rational choice*.[2] In other words, a prominent far-reaching phenomenon such as the AE cannot possibly be incorporated by the standard paradigm's rational choice procedure, since the AE violates this paradigm's key rationality criterion, namely WARP.[3] This paradigm, however, is known to make enormous demands on the decision maker's cognitive capacity to process information. Even the simplest decisions - expressed in the standard form of a decision tree - would easily overwhelm a typical individual's cognitive capabilities.

Not surprisingly, as has been pointed out long ago, in real life "relatively few decisions are made using analytical processes such as generating a variety of options and contrasting their strengths and weaknesses" (Klein, 1989, p. 47). As found by Payne, Bettman and Johnson (1993), when presented with decisions in the form of decision trees, matrices of alternatives and attributes, people - with their limited cognitive capacity to cope with trade-offs - instead adopt heuristics, i.e., lexicographic and other cognitive shortcuts or

opportunities and achieving gains. That is, the promotion focus, eager to 'capturing opportunities and achieving gains', would view the presence of a dominant alternative as an opportunity to be captured and thus be sensitive to the dominance heuristic of the AE.

[2]Within its tractable structure, the standard paradigm of rational choice considers a decision maker with a well-defined deterministic and complete ranking of all feasible alternatives, regardless of the particular choice problem the decision maker faces.

Nevertheless, strong and persistent experimental and field evidence (including evidence on the AE) increasingly kept pointing to various choice behavior that are inconsistent with the premises of this standard paradigm, leading to an increasing scrutiny of this paradigm in time (see Camerer, 1995, and Rabin, 1998, for surveys - we too will briefly elaborate on some such behavior shortly).

[3]In this standard rational choice paradigm, the decision maker always ends up choosing the option that she ranks the highest among any collection of feasible alternatives whenever that option is available, as the 'Weak Axiom of Revealed Preference'(WARP) implies.

Note that the AE violates WARP: in Example 1 both A and B are selected from the feasible set $\{A, B\}$, but B is selected alone from the feasible set $\{A, B, C\}$.

simple but efficient choice rules.[4] Heuristics can simplify complex decision problems very efficiently by obviating the need for trade-offs.[5] People resort to heuristics to avoid the responsibility for making trade-offs as many trade-offs are very cumbersome, involving multi-attribute, interdimensional comparisons - 'balancing proverbial apples and oranges' - that people do not have the cognitive and emotional equipment to perform easily (Tetlock, 2000).[6] There is also strong neuroimaging (i.e., fMRI) evidence provided by Hedgcock and Rao (2009) regarding the presence of 'trade-off aversion' generating negative emotion: "trade-offs represent difficult choices, ... they are a threat to the goal of selecting an option ..., and therefore trade-offs generate negative emotion. ... In other words, consumers prefer making comparisons between options that do not represent trade-offs" (p. 3); the threat of negative emotions in the presence of trade-off aversion in turn leads to heuristics that avoids trade-offs: "[t]he desire to avoid a choice task that generates negative emotion ... yield[s] cognitive processes that emphasize heuristics" (p. 28). In addition, there is strong evidence that *trade-offs between multi-attribute choices* deplete cognitive resources much more than trade-offs not involving multi-attribute choices (see Wang, Novemsky, Dhar and Baumeister, 2010).

The goal of this paper is to identify and examine an alternative framework, which makes much lesser demands on a typical decision maker's cognitive capacity (e.g., via trade-off aversion) that is also conducive to heuristics incorporating the AE - i.e., that can select the AE alternatives - at different degrees. Almost as importantly, another goal of this paper is to provide axiomatic characterizations to help uncover this framework's practical and heuristic choice procedures' most salient properties.

To that end, within such a framework we propose and analyze three heuristics of

[4]Tversky's (1969) 'lexicographic semiorder' constitutes one of the early heuristics that became very prominent. In that heuristic, a preference is generated by the sequential application of numerical criteria by deeming an alternative x better than another alternative y *if* Attribute 1 that distinguishes between x and y ranks x cardinally higher than y *by an amount exceeding a fixed threshold*; if the decision maker does not rank x higher than y in terms of Attribute 1 by an amount exceeding a fixed threshold, then their cardinal ranking by Attribute 2 would have to be considered. Likewise, if the issue could not be settled by Attribute 2 either, then Attribute 3 would be taken into consideration, and so on. Later, Manzini and Mariotti (2012) generalized Tversky's (1969) 'lexicographic semiorder' idea to a full-fledged model of boundedly rational choice.

[5]A popular such heuristic is the $1/N$ rule. In deciding how to allocate financial resources among N options, some individuals rely on the $1/N$ heuristic that allocates financial resources equally across all alternatives (Benartzi and Thaler, 2001). DeMiguel, Garlappi and Uppal (2009) compared the $1/N$ heuristic to 14 optimizing models, including Markowitz's mean-variance portfolio (a Nobel Prize–winning model), in seven investment problems. The $1/N$ heuristic came out first on certainty equivalent returns.

[6]As the moral philosopher Raz (1986, p. 22) put it, there are other deterrents of contemplating trade-offs at times: "It is impoverishing to compare the value of a marriage with an increase in salary. It diminishes one's potentiality as a human being to put a value on one's friendship in terms of improved living conditions." Even to think about such trade-offs may compromise one's standing as a moral being.

which procedures incorporate the AE alternatives such as B at different degrees: (1) the procedure of the "Reference-independent Deterministic Choice" heuristic ("RIDC" hereafter), which will always uniquely select the AE alternative(s) as the outcome by default; (2) the procedure of the "Endogenous Reference-dependent Random Choice" heuristic ("Endogenous-RDRC" hereafter), which will not necessarily select the AE alternative(s) uniquely but with the highest probability nevertheless; (3) the procedure of the "Exogenous Reference-dependent Random Choice" heuristic ("Exogenous-RDRC" hereafter) which will always choose the AE alternative(s) with very high probabilities but not necessarily with the *highest* probabilities, i.e., as will be seen, Exogenous-RDRC may favor alternatives with slightly smaller LCSs but with a larger set of exclusively Pareto-dominated alternatives and a smaller set of jointly Pareto-dominated alternatives than the AE alternative(s).[7]

It will be very clear that our heuritics' procedures, despite their different outcome sets (and as will be seen, despite their very different axioms in these axiomatizations), tend to have some common elements as well. Overall, they are simplistic, rule-of-thumb procedures entailing various short-cuts. They involve rather effortless/costless or random selections among Pareto (i.e., undominated) alternatives. The Endogenous-RDRC's and Exogenous-RDRC's procedures - i.e., the ones that choose the AE alternatives with high probabilities and other Pareto alternatives with low probabilities - will involve (i) a reference alternative x, which is encountered randomly from among dominated alternatives (exogenously in Exogenous-RDRC and endogenously in Endogenous-RDRC), and (ii) a random selection among Pareto alternatives dominating x, while RIDC - i.e., the procedure that never selects any alternatives apart from the AE alternatives - will not even need a reference alternative and will require randomness only to break ties between multiple AE alternatives, if any, to choose one of them. We will then provide axiomatic characterizations of the outcome sets of these procedures. To our knowledge our paper is not just the first one that provides cooperative foundations to heuristics incorporating the AE but also the first one that provides cooperative foundations to stochastic outcomes (i.e., those of the Exogenous-RDRC and Endogenous-RDRC) where the decision maker is not assumed to have stochastic preferences or make random choice errors. Finally, we will also establish a strong inherent link between the RIDC's outcome and the Nash product of a convex and compact domain as well as a weaker such link between the Endogenous-RDRC's outcome and the Nash product; it turns out that there is no such link between the Exogenous-RDRC's outcome and the Nash product.[8]

[7]In Example 1 above, all three heuristics' procedures will always select B uniquely. However, we will provide an example in Section 1.1 where these heuristics' outcomes will differ from each other.

[8]The Nash product, which is a very prominent concept in game theory, maximizes the LCS and thus the geometric average of a choice over a convex and compact set (see Nash, 1950, 1953).

1.1 Towards an Alternative Framework

As the AE violates WARP and thus cannot possibly be incorporated by the standard paradigm of rational choice - which, as alluded to before, easily overwhelms a typical individual's cognitive capabilities -, clearly a new framework is needed to embed choice heuristics that can incorporate a widespread phenomenon such as the AE. In our quest for such an alternative framework, we will start searching for some ingredients among the various choice behavior that violate the tenets of the standard paradigm's rational choice behavior.

Limited Completeness of Preferences in a Multi-attribute Context: One of the main tenets of the standard paradigm is the assumption that individuals have *complete preferences* (or well-defined rankings) over all feasible options, regardless of the trade-offs they encounter comparing them. The complete-preferences assumption has been deemed unrealistic since the seminal contributions of Aumann (1962) and Bewley (1986). As everyday decision making often involves choosing among options that differ in multiple-attribute dimensions and could pose major trade-offs between diffferent attributes, clearly this assumption would be even harder to justify in such a multi-attribute context, which naturally gives rise to phenomena such as the AE. First of all, decision makers may consider some attributes "noncompensatory," i.e., one attribute not being able to compensate for a deficiency in another - or one characteristic not to be traded off against each other. Even with compensatory attributes, there is no universally agreed-upon or intrinsic way to obtain valid and accurate evaluations of attribute weights for a sophisticated decision maker.[9] Forming complete preferences among available options with multiple attributes - with potentially formidable trade-offs between these attributes - would be a more daunting task for a typical consumer, especially when she also has limited experience with the options available to her (more so if individuals were further required to cope with trade-offs by weighing probabilities, time delays, and outcomes experienced by other persons in a consistent fashion). Nevertheless, there would be much less harm in assuming that the decision maker could rank a set of alternatives according to any particular single attribute, e.g., at least she would be able to recognize which PC has a better processing speed and which PC has a worse such speed, etc.

As Piccione and Spiegler (2012, p. 97) put it, although "standard models of market competition assume that consumers rank all the alternatives they are aware of, ... in reality, consumers are often unable to compare alternatives." In all of our procedures, as alluded to by Aumann (1962, p. 446), the decision maker will be "willing and able to arrive at preference decisions only for certain pairs of [alternatives]," i.e., only for alternatives

[9]Several approaches have been proposed in the marketing literature to handle the problem of aggregating a large numbers of attributes, each with its own specific drawbacks (see for instance Green and Srinivasan, 1990, Louviere, 1984, Oppewal, 1994).

that can be Pareto ranked, "while for others [she will] be unwilling or unable to arrive at a decision." It will turn out that such a limited completeness property (motivated by trade-off aversion), where the decision maker could rank a set of alternatives according to any particular single attribute, and the 'Pareto-dominance relationship' in the choice set are sufficient to give rise to RIDC, while Exogeneous-RDRC and Endogenous-RDRC will require the next two features as well.

*Reference-dependence via Limited Attention (*i.e., *Reference-dependent Consideration Sets)*: In addition, there is prevalent evidence since Kahneman and Tversky (1979) which suggests that choices are also often reference dependent: the 'presence of certain types of options' affects the choice behavior of many consumers. In most such cases a specific alternative, which can easily be identified, can serve the role of the reference point or *reference alternative*, e.g., the status quo choice, the endowment, the default option, and so on. Further, reference dependence may emerge even in choice situations where no alternative can serve the role of a status quo choice or a default option: the reference alternative could be an option that the decision maker somehow knows well without owning it, e.g., her spouse or kids or a very close friend may own one, or even one recently seen can serve as a reference alternative. The presence of a reference alternative in the feasible set of alternatives, S, in turn can help a decision maker restrict her attention to only a small fraction of items that are available to her instead of considering all options before making a decision (limited attention), in order to prevent her cognitive capacity from being overloaded.[10] In marketing and economics literatures, such a restricted set is called a "consideration set," i.e., the set of alternatives to which the decision maker pays attention in her choice process.[11]

Note that a consideration set can be very easily constructed with the help of a reference alternative x by restricting the set of alternatives only to x and the ones that Pareto-dominate x, i.e., to x's upper contour set (UCS) in S, which will be denoted by $U(x, S)$. As such, the decision maker would be able to construct a 'reference-dependent consideration set' (S, x) which is $x \cup U(x, S)$.[12] In the context of reference dependence and our proce-

[10]E.g., as a typical supermarket's cereal isle carries about 400+ items, it would be very unrealistic for a typical individual to contemplate all of this overwhelming number of alternatives avalilable to her.

[11]This concept dates back to Wright and Barbour (1977). For a theoretical model on how one can deduce both the decision maker's preferences and the alternatives to which she pays attention (and inattention) from the observed behavior, see Masatlioglu, Nakajima and Ozbay (2012).

[12]The fact that the decision maker prefers to use a (concrete) alternative as the reference alternative to make comparisons and form a consideration set is well known to the marketing literature. Klein and Oglethorpe (1987): "A novice buyer of a particular product should initially have little knowledge about the distribution of attribute levels in the market and some uncertainty about preferences for those levels. ... A concrete alternative should be an easier reference point to use: for instance a well known brand or the one that is first encountered. ... For example, a novice PC buyer may acquire information about the IBM PC first because it's the best known exemplar of the product category. Information about attributes of other

dures, while RIDC will not need a particular reference alternative, Exogeneous-RDRC and Endogenous-RDRC will always need a reference-dependent consideration set, (S, x), where the decision maker encounters her reference alternative randomly from dominated or Pareto alternatives in the feasible set S according to some probability function.[13]

Random Choice: Finally, as individuals routinely violate the key WARP property in experimental and field settings in a manner inconsistent with any fully *deterministic* model of rational choice, we will elaborate on the robust empirical finding of stochastic or *random choice*: 'when subjects are asked to choose from the same set of options many times, they often make different choices'. For instance, to investigate the origin of stochastic choice Agranov and Ortoleva (2016) have recently conducted an experiment.[14] They report that subjects mostly exhibit 'deliberate' random choice by selecting different lotteries in the repetition of the same question, whether they are explicitly informed about the repetition (71%) or not (90%); further, 29% of subjects choose *to pay a cost to flip a coin*. Not surprisingly, however, in Agranov and Ortoleva (2016) random choice is present almost exclusively in the presence of 'hard' questions (i.e., in the presence of difficult risk/return trade-offs), in which none of the available options is 'clearly' better than the other.[15] Indeed, when asked in the questionnaire following the experiment why they indeed chose different answers when the same question was asked multiple times, participants' typical answer was that they did so "because they did not know which option was best, and thus did not want to commit to a specific choice."[16] As will be elaborated later, random choice by the

---

personal computers is then compared to the levels possessed by an IBM PC."

[13]For instance, a recent influential model of market competition of Piccione and Spiegler (2012, p. 99) too assumes that a reference alternative is assigned to the consumer randomly, which the authors interpret "as a default option arising from previous consumption decisions."

[14]Apart from the presence of such recent empirical evidence on random choice, influential theoretical models of random choice too have been proposed in economics and psychology (see, among others, Cerreia-Vioglio, Dillenberger, Ortoleva and RiellaCerreia-Vioglio, 2015, Fudenberg, Iijima and Strzalecki, 2015, Gul, Natenzon and Pesendorfer, 2015, Machina, 1985, Marley, 1997, Swait and Marley, 2013).

[15]Dwenger, Kubler and Weizsacker (2013) too find that subjects prefer avoid making hard decisions but instead choose to delegate their choice to an external device both in the laboratory and in field data, indicating an explicit preference for randomization. They discuss the extent that their experimental data is consistent with a theory of 'responsibility aversion', a version of the regret theory by Loomes and Sugden (1982).

Such random choice could presumably be also due to the "diversification bias" (Simonson, 1989) as well. However, Agranov and Ortoleva show that "the tendency to diversify takes place virtually only for 'hard' questions" (p. 7).

[16]This is reminiscent of the principle of insufficient reason of Bernoulli (1713). Laplace wrote that uniformity should be assumed "[when] we have no reason to believe any particular case should happen in preference to any other" (see Dembski and Marks II, 2009). This principle asserts that where we do not have sufficient reason to regard one possible case as more probable than another, we may treat them as equally probable.

Binmore, Stewart and Voorhoeve (2012, pp. 215-6) examined the predictive power of alternative principles decision maker will be a significant feature of our Exogenous-RDRC and Endogenous-RDRC heuristics' procedures (while RIDC will involve the decision maker's random choice at the minimal possible extent).[17]

1.2 Informal Elaboration of Our Heuristics' Procedures and Axiomatic Characterizations

Exogenous-RDRC: In Exogenous-RDRC, the decision maker encounters her reference alternative x *randomly* from 'dominated' alternatives in S according to some probability function and then arrives at her reference-dependent consideration set (S, x) which only includes x and x's UCS, $U(x, S)$ (i.e., $x \cup U(x, S)$) as described before.[18] If there are no dominated alternatives in S, then one of the alternatives is selected randomly as the outcome. If there are dominated alternatives in S, then the decision maker further refines her consideration set (S, x) by discarding all alternatives that are themselves dominated by any other alternative in $U(x, S)$. Subsequently the decision maker uses a random selection stage only considering the Pareto alternatives in $U(x, S)$. In particular, she selects (i) y as the choice outcome if it is the only Pareto alternative in $U(x, S)$, and (ii) one of the Pareto alternatives in $U(x, S)$ as the choice outcome randomly with equal probabilities if there are multiple such alternatives.[19]

Endogenous-RDRC: In Endogenous-RDRC, the decision maker selects any dominated alternative a as the reference alternative with a probability proportional to the number of Pareto alternatives in $U(a, S)$, in which case one of those Pareto alternatives is selected randomly as the choice outcome. Once the decision maker has a reference alternative in Endogenous-RDRC, then the rest of that heuristic's procedure proceeds as in Exogenous-RDRC described above.

of choice under uncertainty, including the objective maximin and Hurwicz criteria, the sure-thing principle, and the principle of insufficient reason. They state that "[c]ontrary to our expectations, the principle of insufficient reason performed substantially better than rival theories in our experiment."

[17]Indeed, in a recent important model of market competition they have developed, Chiaveanu and Zhou (2013, p. 2451) too went on to assume that if consumers cannot make their minds up among products, "they choose one product randomly" - observe that this is also in line with the findings of Agranov and Ortoleva (2016), regarding individuals' behavior when facing hard choices.

[18]In an earlier version of this paper, we have also considered a case where the reference alternative could be seleceted from undominated alternatives as well. As the latter case has not added anything significant to the anlysis but only made the notation the mathematical structure more complicated, we have decided to proceed without that case.

[19]In the latter case, the decision maker can presumably consult with friends, family and acquaintances as well who may know something about those undominated options, but they too may have different opinions due to their own particular tastes or inadequate/disparate experience, which at the end may not differ much from a random choice among Pareto alternatives.

RIDC: RIDC, on the other hand, blatantly highlights the link between the AE and LCSs of alternatives. It features a decision maker who selects the alternative(s) with the largest LCS as the choice outcome(s). If there is only one such alternative with the largest LCS, it will be her choice outcome for sure. If there are multiple such alternatives, then she will select one of them as the choice outcome randomly with equal probabilities.

As mentioned before, RIDC always uniquely chooses the AE alternative(s) by default; Endogenous-RDRC does not necessarily choose the AE alternative(s) uniquely but chooses them with the highest probability; and Exogenous-RDRC always chooses the AE alternative(s) with high probabilities but not necessarily with the highest probabilities, as it distinguishes between exclusively Pareto-dominated alternatives vs. jointly Pareto-dominated alternatives in LCSs.

The Extent That Our Heuristics' Procedures Incorporate the AE - An Example: To illustrate the differences between these three procedures, suppose that the decision maker considers three Pareto-dominant alternatives X, Y, and Z and three Pareto-dominated alternatives, e, j, j', such that j and j' are jointly Pareto-dominated by X and Y, while e is exclusively Pareto-dominated by Z. Suppose that these alternatives' attribute rankings can be expressed as $X_{14}, Y_{23}, j_{35}, j'_{46}, Z_{51}, e_{62}$, where a smaller number for each attribute indicates a better (preferred) attribute ranking for the decision maker (we will refer to this setup as *Example* 2). That is, X is the best option for the decision maker in terms of Attribute 1 while Z is her best option in terms of Attribute 2. Note that X and Y have the same LCS consisting of both j, j', while Z has the LCS consisting of only e'. Thus, RIDC will select each of X and Y as the outcome with probability $\frac{1}{2}$. For Endogenous-RDRC's outcome probabilities, we first need to calculate UCS sizes of e, j, j'. Observe that they are $1, 2, 2$ respectively. So, e will become the reference alternative with probability $\frac{1}{5}$, j with probability $\frac{2}{5}$, and j' with $\frac{2}{5}$; thus, j and j' will have twice as large probability to become a reference alternative than e will, since the UCS of each of j and j' has twice as many Pareto alternatives as the UCS of e does. Therefore, each of X and Y will be selected as the outcome with probability $\frac{2}{5}$, and Z will become the outcome with probability $\frac{1}{5}$. In Exogenous-RDRC, each of e, j, j' will become the reference alternative with probability $\frac{1}{3}$, and therefore each of X, Y, Z will be selected as the outcome with probability $\frac{1}{3}$. This is because, while X and Y have twice as large LCS as that of Z, the lone alternative in Z's LCS is exclusively dominated by Z, whereas each alternative in the LCS of each of X and Y is jointly dominated by each X and Y. In other words, when e becomes the reference alternative, Z will be selected as the outcome for sure, since Z is the only Pareto alternative in $U(e, S)$, while when either j or j' becomes the reference alternative, either of X and Y will be selected as the outcome with probability $\frac{1}{2}$ each only, since $U(j, S)$ and $U(j', S)$ both include X and Y as Pareto alternatives.

Intuition on Our Heuristics' Incorporation of the AE at Different Levels: Thus, in all of our heuristics' procedures, the dominated alternatives can never be selected as outcomes,

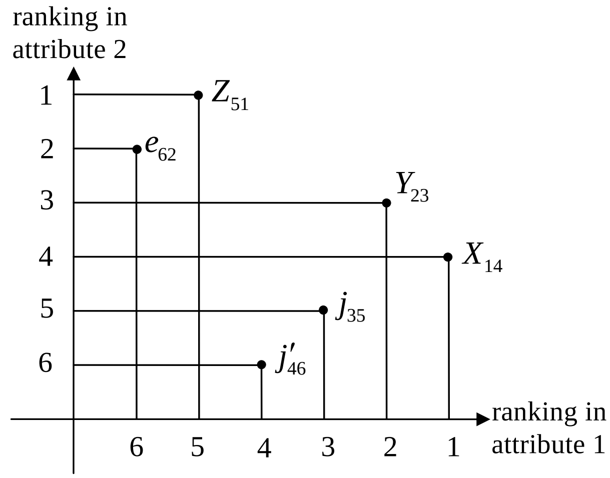

Figure 2: Example 2.

but they will be the "king makers" in that they will have great influence over the choice probabilities of Pareto alternatives that Pareto-dominate them. In RIDC, for each Pareto alternative y, each dominated alternative x will matter as much as any other dominated alternative that is dominated by y. In Endogenous-RDRC, each dominated alternative x will matter as much as the number of the Pareto alternatives in x's UCS, although - unlike Exogenous-RDRC - it does not matter whether the dominated alternatives are jointly dominated or exclusively dominated by some Pareto alternatives. For Exogenous-RDRC, on the other hand, the *exclusively-dominated alternatives* are the "true" king makers, while *jointly-dominated alternatives* are less so, especially if they are dominated jointly by a larger number of Pareto alternatives. In particular, the impact of a dominated alternative x on the choice probability of a Pareto alternative will be at its minimum when x is dominated by all Pareto alternatives. Further, in Exogenous-RDRC, the probability of each dominated alternative to become the reference alternative is the same (and thus exogenous) while in Endogenous-RDRC the probability of each dominated alternative x to become the reference alternative depends on the number of Pareto alternatives in x's UCS (and thus is endogenous).

Axiomatic Characterizations of our Heuristics' Outcome Sets: Apart from satisfying Weak Pareto Optimality, each of the axiomatic characterizations of Exogenous-RDRC and Endogenous-RDRC satisfies axioms that consider exclusive domination of an alternative x by y - i.e., when y is the only Pareto alternative dominates x - versus joint domination of x - i.e., when multiple Pareto alternatives such as y and z dominate x - where the relative attribute rankings of y and z do not change between these two setups. In Exogenous-RDRC's characterization, these axioms jointly deem that y's choice probability should increase exactly as much as z's choice probability decreases, while the choice probabilities of other Pareto alternatives should remain the same. In that of Endogenous-RDRC, these axioms jointly deem that y's choice probability need not increase exactly as much as z's choice probability decreases and the choice probabilities of other Pareto alternatives need not remain the same. When one considers exclusive vs. joint domination in the context of RIDC, however, - as RIDC does not distinguish between these types of domination -, y's choice probability need not even increase and z's choice probability need not even decrease, and the choice probabilities of other Pareto alternatives need not remain the same. Therefore, the crucial axiom in the characterization of RIDC only focuses on the circumstances that leave the choice outcome from one set, S, to a larger one, S', unchanged (in addition, RIDC's axiomatic characterization also requires a strong Symmetry axiom; although RIDC's axiomatic characterizations does not require Weak Pareto Optimality (WPO), it satisfies WPO nevertheless).

It is worth noting that the fact the outcomes of these three fairly different heuristics' procedures incorporating the AE differ significantly and the axioms in their axiomatizations differ from each other almost entirely, in a sense, implies a strong degree of theoretical robustness for the AE phenomenon.

The next section provides the Preliminaries. Sections 3 and 4 are on Exogenous-RDRC, Endogenous-RDRC and RIDC, while Sections 5, 6, and 7 provide the axiomatic characterizations of their outcome sets, respectively. Section 8 is on the Nash product convergence of RIDC's and Endogenous-RDRC's outcome sets. Relevant Literature is in Section 9. Section 10 concludes with some general remarks.

2 Preliminaries

Let $(S, \{\succ_I\}_{I\in\Lambda})$ or $(S, \succ_1, ..., \succ_N)$ be a *finite decision problem with multiple attributes*, where (i) S denotes the *feasible set* with a finite number of $\#S > 1$ alternatives, with $\#(\cdot)$ denoting the cardinality of any set $(\cdot)$, (ii) Λ denotes the *set of attributes* of the alternatives in S with $\#\Lambda = N$, and (iii) $\succ_I$ denotes the decision maker's preferences or *rankings* of alternatives in S in terms of Attribute I where $I = 1, 2, ..., N$.

Although $\succ_I$ could be expressed as $>_I$ for some measurable attributes such as price, storage, processor speed, it cannot be expressed as $>_I$ for unmeasurable attributes such as colors, odors, shapes, etc. We assume that $\succ_I$ is complete and transitive. That is, (i) the decision maker can rank all items in S according to each attribute $I \in \Lambda$ such that it is either $y \succ_I z$ or $y \prec_I z$ for all $y, z \in S$, and (ii) $z \succ_I y$ and $y \succ_I x$ then $z \succ_I x$ for all $x, y, z \in S$. For tractability, we rule out the indifference case $y \sim_I z$ for any two items $y, z \in S$.

Although in our analysis initially we will focus on the case of $\#\Lambda = 2$ attributes, later we will also consider the case $\#\Lambda > 2$. Thus, first let $\Lambda = \{I, J\}$. Let $s_{ij}, s_{i'j'} \in S$ such that $i < i'$ means that the decision maker prefers s_{ij} over $s_{i'j'}$ in terms of Attribute I and $j < j'$

means that the decision maker prefers s_{ij} over $s_{i'j'}$ in terms of Attribute J. Thus, $i = 1$ if s_{ij} is the decision maker's most preferred alternative in terms of Attribute I and $j = 1$ if s_{ij} is the decision maker's most preferred alternative in terms of Attribute J; likewise, $i = \#S$ if s_{ij} is the decision maker's least preferred alternative in terms of Attribute I and $j = \#S$ if s_{ij} is the decision maker's least preferred alternative in terms of Attribute J.

For any two alternatives $y, z \in S$, we will use $y \succ z$ to denote $y \succ_I z$ and $y \succ_J z$. As alluded to before, the decision maker's preferences, $\succ$, over alternatives in S are not complete (although, as mentioned above, her preferences over alternatives in terms of each attribute is complete). We assume that she can only compare and rank any two alternatives if and only if they are Pareto-rankable, that is, whenever $y \succ z$ or $y \prec z$. Likewise, the decision maker's preferences over alternatives in S are transitive if and only if they are Pareto-rankable: if $z \succ y$ and $y \succ x$ then $z \succ x$.

Let $U(x, S, \succ_1, \succ_2)$ and $L(x, S, \succ_1, \succ_2)$ denote the upper and lower contour sets of x for the decision maker such that $U(x, S, \succ_1, \succ_2) = \{y \in S : y \succ x\}$ and $L(x, S, \succ_1, \succ_2) = \{y \in S : x \succ y\}$. Let $\partial(S, \succ_1, \succ_2)$ denote the *Pareto Set* of S, where $\partial(S, \succ_1, \succ_2) = \{y \in S | \text{there is no } y' \in S \text{ such that } y' \succ y\}$. Let $\eta(S, \succ_1, \succ_2)$ denote the *Set of Dominated Alternatives* in S where $\eta(S, \succ_1, \succ_2) = S \backslash \partial(S, \succ_1, \succ_2)$. We will often simply write $L(x, S, \succ_1, \succ_2)$, $U(x, S, \succ_1, \succ_2)$, $\partial(S, \succ_1, \succ_2)$ and $\eta(S, \succ_1, \succ_2)$ as $L(x, S)$, $U(x, S)$, ∂S and ηS respectively, whenever no confusion will arise from doing so (e.g., especially when $(\succ_1, \succ_2)$ is fixed).

Let $\mathbb{D}_2$ be the set of all finite decision problems with two attributes $(S, \succ_1, \succ_2)$. A *choice rule* is a function $q\colon S \times \mathbb{D}_2 \to [0, 1]$ with $q(y, S, \succ_1, \succ_2) \geq 0\ \forall y \in S$ and $\sum_{y \in S} q(y, S, \succ_1, \succ_2) = 1$. Thus, the value $q(y, S, \succ_1, \succ_2)$ is the probability of choosing alternative y when the finite random reference decision problem is $(S, \succ_1, \succ_2)$. (We sometimes write $q(y, S, \succ_1, \succ_2)$ as $q(y, S)$ when no confusion will arise from doing so.)

3 Exogenous-RDRC

Let $x \in S$ denote a randomly selected *reference alternative* from S. In particular, we assume only dominated alternatives can be selected as the reference alternative (hence, each Pareto alternative is selected as the reference with probability zero). Thus, $U(x, S) \neq \emptyset$. Further, each dominated alternative is selected as the reference alternative with the same probability, i.e., with probability $\frac{1}{\#\eta S}$.[20]

The decision maker's choice procedure according to the *Exogenous Reference-dependent Random Choice* (Exogenous-RDRC) will explicitly operate as follows:

(1) Suppose there are no dominated alternatives. Then each of the Pareto alternatives

[20]As mentioned before, in an earlier version we also considered a case where undominated alternatives too could be selected as a reference alternative, which did not lead to any additional gain of insight but to some further complication mathematically.

is selected as the choice outcome randomly with equal probabilities.

Suppose there is at least one dominated alternative. Then:

(2) An alternative x in S is selected as the reference alternative. In particular, any Pareto alternative will be selected as the choice outcome with the same probability $\frac{1}{\#\eta S}$.

(3) Suppose there is a unique alternative y in $U(x, S)$. Then the decision maker will select y as the choice outcome. Suppose there are multiple alternatives in $U(x, S)$, and there is only one, y, in $\partial U(x, S)$. Then she will certainly select y as the choice outcome.

Thus the decision maker will bypass any x' such that $y \succ x' \succ x$, and select y, since there is no other alternative in $\partial U(x, S)$ that could possibly attract the decision maker. Thus, the decision maker will never move from x to another dominated x' alternative in $U(x, S)$ since such a move would then require another switch to an alternative y in $\partial U(x, S)$ which would necessarily dominate the intermediate inferior alternative x' (even if such a second move is not costly for the decision maker, the initial move to x' from x would at least be redundant, nevertheless).

(4) Suppose, however, that there are multiple alternatives in $\partial U(x, S)$. Then, the decision maker will select any of the alternatives in $\partial U(x, S)$ as the choice outcome randomly with equal probabilities, since the decision maker cannot possibly rank those alternatives in $\partial U(x, S)$.

Recall Example 2 in Section 1. Observe that in the Exogenous-RDRC's procedure while a dominated alternative x has no chance to be the choice, a Pareto alternative has to have a non-empty LCS to become a choice outcome. Further, ceteris paribus, if a Pareto alternative's LCS gets largers it gets a higher chance to become a choice alternative. In addition, ceteris paribus, if a Pareto alternative's LCS composition changes such that it has a larger number of exclusively-dominated alternatives vs. jointly dominated ones, then it gets even a higher chance to become a choice alternative.

For any $y \in \partial S$, let $J_i(y, S) \subseteq L(y, S)$ be the set of alternatives that are dominated by y and are dominated by exactly i Pareto alternatives (including y).

Theorem 1 *Let q^{ex} be the choice rule induced by Exogenous-RDRC. For any $y \in \partial S$, $q^{ex}(y, S) = \frac{1}{\#\eta S} \sum_{i=1}^{\#\partial S} \frac{\#J_i(y,S)}{i}$.*[21]

Consider the following example, to which we will refer as the $AA' - CC'$ example in the remainder of the paper. It involves a symmetric set of alternatives of which attribute rankings can be expressed as $A_{15}, a_{26}, C_{34}, C'_{43}, A'_{51}, a'_{62}$, where again a lower number for each attribute indicates a better (preferred) attribute ranking for the decision maker. Then observe that the choice outcome of Exogenous-RDRC is either A or A' with equal probabilities.

[21] Proofs of our results are in the Appendix unless stated otherwise.

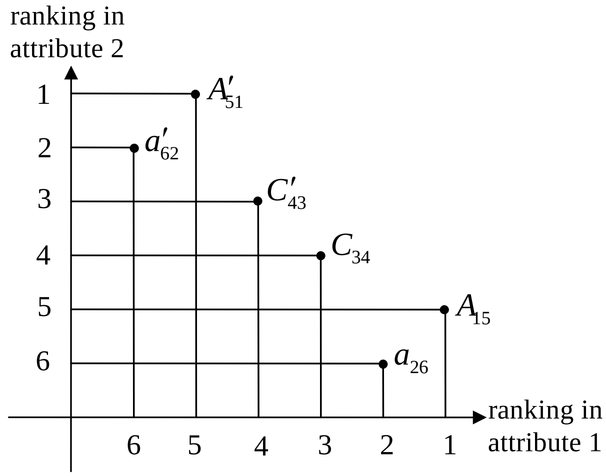

Figure 3: $AA' - CC'$ Example.

4 Endogenous-RDRC and RIDC

We now study an endogenous variant of Exogenous-RDRC, the so-called *Endogenous Reference-dependent Random Choice (Endogenous-RDRC)*. Its procedure is as follows. An alternative a is selected as the reference alternative according to the following rules: (i) as is the case in Exogenous-RDRC, only dominated alternatives can be selected as the reference alternative (thus, each Pareto alternative is selected as the reference with probability zero),[22] and (ii) for any $a \in \eta S$, a is selected with a probability that is proportional to $\#\partial U(x, S)$, i.e., proportional to the cardinality of the Pareto set of the set of alternatives that dominate a. The remainder of the Endogenous-RDRC's procedure is the same as that of the Exogenous-RDRC.

Let $I(S, \succ_1, \succ_2) = \sum_{y \in \partial(S, \succ_1, \succ_2)} \#L(y, S, \succ_1, \succ_2)$; in short, $I(S) = \sum_{y \in \partial S} \#L(y, S)$, when $(\succ_1, \succ_2)$ is fixed. Letting $\#\partial(S, \succ_1, \succ_2) = m$ and $\#\eta(S, \succ_1, \succ_2) = n$, we must have $n \leq I(S, \succ_1, \succ_2) \leq mn$. Roughly speaking, if we fix the number of Pareto alternatives and the number of dominated alternatives m and n respectively, then note that, as $I(S, \succ_1, \succ_2)$ increases, the dominated alternatives become more jointly dominated by Pareto alternatives (an extreme case is that all dominated alternatives are dominated by all Pareto alternatives, in which we have $I(S, \succ_1, \succ_2) = mn$). We thus call $I(S, \succ_1, \succ_2)$ the *degree of joint dominance* of $(S, \succ_1, \succ_2)$.

We then obtain the following result.

Theorem **2** *Let q^{en} be the choice rule induced by Endogenous-RDRC. For any $y \in \partial S$, $q^{en}(y, S) = \frac{\#L(y,S)}{I(S)}$.*

[22]If there are no dominated alternatives, then each of the Pareto alternatives is selected as the choice outcome randomly with equal probability.

Theorem 2 shows that in Exogenous-RDRC, the Pareto alternative that has the largest lower contour set will be chosen with the largest probability. As will be elaborated on later, there is a strong link between Endogenous-RDRC's outcome(s) and those of the Nash product.

Finally, we study an essentially trivial heuristic, namely the so-called *Reference-independent Deterministic Choice (RIDC)*. Its procedure is as follows. No alternative is designated as the decision maker's reference alternative. The decision maker directly selects the alternative y with the largest $L(y, S)$ as the choice outcome if there is a unique such alternative. If there are multiple alternatives $y_1, ..., y_n \in S$ such that $\#L(y_i, S) \geq \#L(z, S)$ with $i = 1, ..., n$ and any $z \in S$, then the decision maker will select any y_i randomly with equal probabilities as the choice outcome.

Theorem 3 *Let q^r be the choice rule induced by RIDC. Then, the alternatives that have positive probability under q^r must have the largest lower contour sets.*

Proof of the above theorem is obvious and thus omitted. Note that in RIDC only the alternatives with the largest LCSs are selected with positive probability. As such (and as will be elaborated later), there is a very direct strong link between the RIDC's outcome and the Nash product (i.e., much stronger link than the one between the Endogenous-RDRC's outcome and the Nash product).

Exogenous-RDRC, Endogenous-RDRC and RIDC coincide on the AE alternatives in the $AA' - CC'$ example above. However, it is easy to see that their outcomes need not coincide given an arbitrary S. Consider the following example: $s_{17}s_{28}s_{36}s_{43}s_{54}s_{61}s_{72}s_{85}$ (see Figure 4). Then observe that RIDC assigns $q(s_{17}) = q(s_{28}) = q(s_{36}) = q(s_{54}) = q(s_{72}) = q(s_{85}) = 0$ but $q(s_{43}) = q(s_{61}) = 0.5$, Endogenous-RDRC assigns $q(s_{28}) = q(s_{36}) = q(s_{54}) = q(s_{72}) = q(s_{85}) = 0$ but $q(s_{17}) = 0.2, q(s_{43}) = q(s_{61}) = 0.4$, while Exogenous-RDRC assigns $q(s_{28}) = q(s_{36}) = q(s_{54}) = q(s_{72}) = q(s_{85}) = 0$ but $q(s_{17}) = 0.25, q(s_{43}) = q(s_{61}) = 0.375$.

5 Axiomatic Characterization of the Exogenous-RDRC's Outcome Set: The 2−attributes Case

Consider the following standard property.

Weak Pareto Optimality (**WPO**): For any $(S, \succ_1, \succ_2) \in \mathbb{D}_2$, $q(y, S, \succ_1, \succ_2) = 0$ for $\forall y \in \eta(S, \succ_1, \succ_2)$.

Thus, WPO states that any dominated alternative cannot possibly be a choice outcome.

We will next define exclusive domination. Let $E(y, S, \succ_1, \succ_2) = \{x \in S | x \in L(y, S, \succ_1, \succ_2)$ where $y \in \partial(S, \succ_1, \succ_2)$ and $x \notin L(z, S, \succ_1, \succ_2)$ where $z \in \partial(S, \succ_1, \succ_2)$ and $z \neq y\}$. That is, $E(y, S, \succ_1, \succ_2)$ is the set of alternatives in S which are dominated only by the efficient alternative y in S and not dominated by another efficient alternative in S.

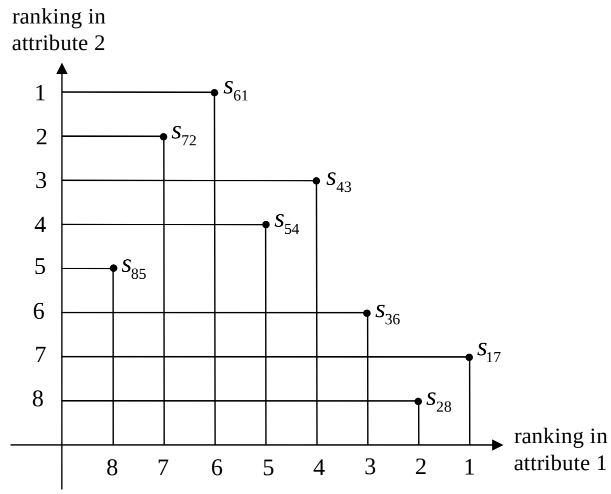

Figure 4: Example.

Next, we will define *joint domination*. Let $J(y, z, S, \succ_1, \succ_2) = \{x \in S | x \in L(y, S, \succ_1, \succ_2)$, $x \in L(z, S, \succ_1, \succ_2)$, and $x \notin L(w, S, \succ_1, \succ_2)$ for any $w \neq y, z$ and $w \in \partial(S, \succ_1, \succ_2)\ \}$. That is, $J(y, z, S, \succ_1, \succ_2)$ is the set of alternatives in S which are dominated by both y and z but not by any other Pareto alternative. More generally, we define $J(y_1, ..., y_n, S, \succ_1, \succ_2)$ as the set of alternatives in S which are dominated by $y_1, ..., y_n$ but not by any other Pareto alternative.

We mentioned above that dominated alternatives, although they cannot be the choice of the decision maker themselves, are king-makers. The *exclusively-dominated* alternatives are the "true" king-makers, while jointly dominated alternatives are less so, especially if they are dominated jointly by many Pareto alternatives. Observe that the impact of a dominated alternative x on the choice probability of a Pareto alternative y is at its minimum when x is dominated by all Pareto alternatives.

Let $\Delta q(y, S, \succ'_1, \succ'_2 | \succ_1, \succ_2) = q(y, S, \succ'_1, \succ'_2) - q(y, S, \succ_1, \succ_2)$ which denotes the difference in the choice probabilities of y when the decision problem is $(S, \succ'_1, \succ'_2)$ vs. when the decision problem is $(S, \succ_1, \succ_2)$.

For any $y \in S$ and $A \subseteq S$ let $R_i(y|A, \succ_1, \succ_2)$ be the *relative ranking* of alternative y in the set A according to Attribute i.

Suppose that for two decision problems $(S, \succ_1, \succ_2), (S, \succ'_1, \succ'_2) \in \mathbb{D}_2$, there is an $a \in \eta(S, \succ_1, \succ_2)$ and $y_1, ..., y_n \in \partial(S, \succ_1, \succ_2)$ such that $a \in J(y_1, ..., y_n, S, \succ_1, \succ_2)$, $a \in J(y_1, ..., y_{n-1}, S, \succ'_1, \succ'_2)$, and $R_i(y|S \backslash a, \succ_1, \succ_2) = R_i(y|S \backslash a, \succ'_1, \succ'_2)$ for any $y \in S \backslash a$ and any $i \in \{1, 2\}$. We call the above setup the *JD setup* (i.e., from More Joint Domination in $(S, \succ_1, \succ_2)$ to a Less One in $(S, \succ'_1, \succ'_2)$). We define the following three axioms, which concern the change of the choice probability of undominated alternatives in the JD setup.

Winners' Symmetric Gain (**WSG**): Consider the JD setup. Then $\Delta q(y_i, S, \succ'_1, \succ'_2 | \succ_1, \succ_2) = \Delta q(y_j, S, \succ'_1, \succ'_2 | \succ_1, \succ_2) \geq 0$ for any $y_i, y_j \in \{y_1, ..., y_{n-1}\}$.

Irrelevance of Uninvolved Alternatives (**IUA**): Consider the JD setup. Then $\Delta q(y, S, \succ_1', \succ_2' | \succ_1, \succ_2) = 0$ for any $y \in \partial(S, \succ_1, \succ_2)$ with $y \notin \{y_1, ..., y_n\}$.

Loser's Loss (**LL**): Consider the JD setup. Then $\Delta q(y, S, \succ_1', \succ_2' | \succ_1, \succ_2) = -\frac{1}{n \# \eta S}$ for $y = y_n$.

Note that WSG and LL (together with WPO) imply that $\Delta q(y, S, \succ_1', \succ_2' | \succ_1, \succ_2) = \frac{1}{\# \eta S [n(n-1)]}$ for any $y \in \{y_1, ..., y_{n-1}\}$.

Next, for our next property, we need the following definition. $(S, \succ_1, \succ_2)$ is *symmetric* (or simply, S is *symmetric*) if for any $s \in S$, there exists a $s' \in S$ such that $R_1(s|S, \succ_1, \succ_2) = R_2(s'|S, \succ_1, \succ_2)$ and $R_2(s|S, \succ_1, \succ_2) = R_1(s'|S, \succ_1, \succ_2)$.

Symmetry (**SYM**): Suppose (i) $(S, \succ_1, \succ_2) \in \mathbb{D}_2$ is symmetric and (ii) $L(y, S, \succ_1, \succ_2) = L(z, S, \succ_1, \succ_2)$ for any $y, z \in \partial(S, \succ_1, \succ_2)$. Then $q(y, S, \succ_1, \succ_2) = q(z, S, \succ_1, \succ_2)$ for ant $y, z \in \partial(S, \succ_1, \succ_2)$.

Thus, SYM requires that in a symmetric S where *all dominated alternatives are dominated by all Pareto alternatives*, all undominated alternatives are chosen with equal probabilities (e.g., it holds for sure if $S = \{x, x', y, y', y''\}$ and $\partial S = \{y, y', y''\}$ (where y is ranked first in attribute 1 and y' is ranked first in attribute 2) and x, x' are dominated by all y, y', y''). Note, however, that when S is symmetric in different ways (e.g., x is dominated by only y and x' is dominated by only y'), SYM does not have such a requirement.

Together with WPO, SYM would imply that in a symmetric S where all dominated alternatives are dominated by all Pareto alternatives, each Pareto alternative will be the decision maker's choice with probability $\frac{1}{\# \partial S}$, while no dominated alternative can ever be the decision maker's choice.

The following lemma shows that if $(S, \succ_1, \succ_2)$ is such that all dominated alternatives are dominated by all Pareto alternatives (where S can be asymmetric), then the Pareto set of S must be symmetric.

Lemma 1 *Suppose $(S, \succ_1, \succ_2)$ is such that $L(y, S, \succ_1, \succ_2) = L(z, S, \succ_1, \succ_2)$ for any $y, z \in \partial(S, \succ_1, \succ_2)$, then $\partial(S, \succ_1, \succ_2)$ is symmetric.*

The next lemma shows that WPO, WSG, IUA, and LL together imply SYM.

Lemma 2 *WPO, WSG, IUA, and LL imply SYM.*

The following lemma shows that for any decision problem $(S, \succ_1, \succ_2)$ where $a \in S$ is dominated by l Pareto alternatives with $l < \# \partial(S, \succ_1, \succ_2)$, then we can always "transform" the problem to a problem in which a is dominated by exactly one more Pareto alternative and the relative rankings of all other alternatives remain unchanged.

Lemma 3 *Suppose that $a \in J(y_1, ..., y_l, S, \succ_1, \succ_2)$ for some $\{y_1, ..., y_l\} \subset \partial(S, \succ_1, \succ_2)$ (strict relation), then there exists an attribute ranking $(\succ_1', \succ_2')$ such that $a \in J(y_1, ..., y_l, y^*, S, \succ_1, \succ_2)$ for some $y^* \in \partial(S, \succ_1, \succ_2) \backslash \{y_1, ..., y_l\}$ and $R_i(y|S \backslash a, \succ_1', \succ_2') = R_i(y|S \backslash a, \succ_1, \succ_2)$ for any $y \in S \backslash a$ and any $i = 1, 2$.*

The next axiom we will consider is an independence axiom, the *Independence of Innocuous Shuffle (IIS)* axiom. It requires that the choice probabilities should remain the same if the relative attribute rankings of dominated alternatives change but each dominated alternative is still dominated by the same Pareto alternative(s).

Independence of Innocuous Shuffle (**IIS**): For any two $(S, \succ_1, \succ_2), (S, \succ_1', \succ_2') \in \mathbb{D}_2$, suppose $\partial(S, \succ_1, \succ_2) = \partial(S, \succ_1', \succ_2')$, and $L(y, S, \succ_1, \succ_2) = L(y, S, \succ_1', \succ_2')$ and $R_i(y|\partial(S, \succ_1, \succ_2), \succ_1, \succ_2) = R_i(y|\partial(S, \succ_1', \succ_2'), \succ_1', \succ_2')$ for any $y \in \partial(S, \succ_1, \succ_2) = \partial(S, \succ_1', \succ_2')$ and any $i \in \{1, 2\}$. Then $q(y, S, \succ_1, \succ_2) = q(y, S, \succ_1', \succ_2')$ for any $y \in S$.

The next lemma shows that WPO, WSG, IUA, and LL together imply IIS.

Lemma 4 *If a choice rule satisfies WPO and WSG, IUA, and LL, then it satisfies IIS.*

We then have the following result:

Theorem 4 *The choice rule of the Exogenous-RDRC is the unique one that satisfies WPO, WSG, IUA, and LL.*

We extend the above axiomatization of Exogenous-RDRC's outcome set to the N-attributes case, where $N > 2$, in the Appendix.

We also establish the independence of the axioms used in the above theorem in the Appendix.

6 Axiomatic Characterization of the Endogenous-RDRC's Outcome Set: The 2−attributes Case

Recall that, for any given $(S, \succ_1, \succ_2)$, we have defined $I(S, \succ_1, \succ_2) = \sum_{y \in \partial(S, \succ_1, \succ_2)} \#L(y, S, \succ_1, \succ_2)$ as the *degree of joint dominance* of $(S, \succ_1, \succ_2)$. Again, recall that, letting $\#\partial(S, \succ_1, \succ_2) = m$ and $\#\eta(S, \succ_1, \succ_2) = n$, we must have $n \leq I(S, \succ_1, \succ_2) \leq mn$.

The following two axioms define how the choice probability of a Pareto alternative changes in a JD setup. Recall that a JD setup refers to the situation where there are two decision problems $(S, \succ_1, \succ_2)$ and $(S, \succ_1', \succ_2')$ and there is an $a \in \eta(S, \succ_1, \succ_2)$ and $y_1, ..., y_n \in \partial(S, \succ_1, \succ_2)$ such that $a \in J(y_1, ..., y_n, S, \succ_1, \succ_2)$, $a \in J(y_1, ..., y_{n-1}, S, \succ_1', \succ_2')$, and $R_i(y|S \backslash a, \succ_1, \succ_2) = R_i(y|S \backslash a, \succ_1', \succ_2')$ for any $y \in S \backslash a$ and any $i \in \{1, 2\}$.

Non-losers' Proportional Gain (**NPG**): Consider the JD setup. Then for any $y \in \partial(S, \succ_1, \succ_2) \backslash y_n$, $\Delta q(y, S, \succ_1', \succ_2' | \succ_1, \succ_2) > 0$ if and only if $\#LCS_y > 0$. In addition, $\frac{\Delta q(y, S, \succ_1', \succ_2' | \succ_1, \succ_2)}{\Delta q(y', S, \succ_1', \succ_2' | \succ_1, \succ_2)} = \frac{\#LCS_y}{\#LCS_{y'}}$ for any $y, y' \in \partial(S, \succ_1, \succ_2) \backslash y_n$ with $\#LCS_y > 0$ and $\#LCS_{y'} > 0$.

Loser's Loss (**LL 2**): Consider the JD setup. Then $\Delta q(y_n, S, \succ_1', \succ_2' | \succ_1, \succ_2) = -\frac{I(S, \succ_1, \succ_2) - \#LCS_{y_n}}{I(S, \succ_1, \succ_2)(I(S, \succ_1, \succ_2) - 1)}$.[23]

[23]In NPG and LL2, LCS_y refers to the lower contour set of y under $(S, \succ_1, \succ_2)$.

Note that NPG and LL 2 (together with WPO) imply that $\Delta q(y, S, \succ_1', \succ_2' | \succ_1, \succ_2) = \frac{\#LCS_y}{I(S,\succ_1,\succ_2)(I(S,\succ_1,\succ_2)-1)}$ for any $y \neq y_n$ and $y \in \partial(S, \succ_1, \succ_2)$.

We have the following result.

Theorem 5 *The choice rule of the Endogenous-RDRC is the unique one that satisfies WPO, NPG and LL 2.*

The idea of the proof of Theorem 5 is similar to Theorem 4 and is omitted.

We extend the above axiomatization of the Endogenous-RDRC's outcome set to the N-attributes case, where $N > 2$, in the Appendix.

We also establish the independence of the axioms used in the above theorem in the Appendix.

7 Axiomatic Characterization of the RIDC's Outcome Set: The 2−attributes Case

This section provides an axiomatic characterization for the RIDC outcome set.

Strong Symmetry (**S-SYM**): For any $(S, \succ_1, \succ_2) \in \mathbb{D}_2$, suppose (i) $(S, \succ_1, \succ_2)$ is symmetric and (ii) $L(y, S, \succ_1, \succ_2) = L(z, S, \succ_1, \succ_2)$ for any $y, z \in \partial(S, \succ_1, \succ_2)$. Then, $q(y, S, \succ_1, \succ_2) = \frac{1}{\#\partial(S,\succ_1,\succ_2)}$ for $\forall y \in \partial(S, \succ_1, \succ_2)$.

Thus, S-SYM requires that in a decision problem, if all dominated alternatives are dominated by all Pareto alternatives, then all Pareto alternatives will be chosen with positive probabilities. In addition, all these positive choice probabilities must be equal to $\frac{1}{\#\partial S}$.

The second and last axiom is 'Independence of Irrelevant Expansions'.

Independence of Irrelevant Expansions (**IIE**): For any two $(S, \succ_1, \succ_2), (S, \succ_1', \succ_2') \in \mathbb{D}_2$, suppose $S \subseteq S'$ and $\partial(S, \succ_1, \succ_2) \subseteq \partial(S', \succ_1', \succ_2')$ where $R_i(y|\partial(S, \succ_1, \succ_2), \succ_1, \succ_2) = R_i(y|\partial(S', \succ_1', \succ_2'), \succ_1', \succ_2')$ for any $y \in \partial(S, \succ_1, \succ_2)$ and any $i \in \{1, 2\}$. If for any $y \in S$ such that $q(y, S, \succ_1, \succ_2) > 0$, we have (i) $\#L(y, S', \succ_1', \succ_2') > \#L(z, S', \succ_1', \succ_2')$ for any $z \in \partial(S, \succ_1, \succ_2)$ with $q(z, S, \succ_1, \succ_2) = 0$ and any $z \in \partial(S', \succ_1', \succ_2') \backslash \partial(S, \succ_1, \succ_2)$, and (ii) $\#L(y, S', \succ_1', \succ_2') = \#L(z, S', \succ_1', \succ_2')$ for any $z \in \partial(S, \succ_1, \succ_2)$ with $q(z, S, \succ_1, \succ_2) > 0$, then $q(y, S', \succ_1', \succ_2') = q(y, S, \succ_1, \succ_2)$ for any $y \in S$.

Suppose S in contained in S' and each Pareto alternative y in S is Pareto in S' as well; also suppose that (i) the size of y's LCS is larger than the size of z's LCS in S' whenever y was a choice outcome (i.e., selected by the decision maker with positive probability) and z was not a choice outcome in S, and (ii) the size of y's LCS is equal to that of z's LCS in S' whenever y was a choice outcome and z was also a choice outcome in S. Then the IIE requires that the choice probabilities should remain the same in S'.

Note that in IIE, a "valid" expansion of a decision problem $(S, \succ_1, \succ_2)$ to another decision problem $(S', \succ_1', \succ_2')$ depends on the choice probabilities in the former decision problem.

Figure 5 illustrates two decision problems. If the choice probabilities for the problem on the left are such that $q(x) = q(y) = q(z) = 1/3$, then the problem on the right is a valid expansion of the problem on the left (and by IIE, the problem on the right should have the same choice probabilities for x, y, and z as the decision problem on the left). However, if the choice probabilities for the problem on the left are such that $q(x) = 1$ and $q(y) = q(z) = 0$, then we cannot apply IIE to get the choice probabilities for the problem on the right.

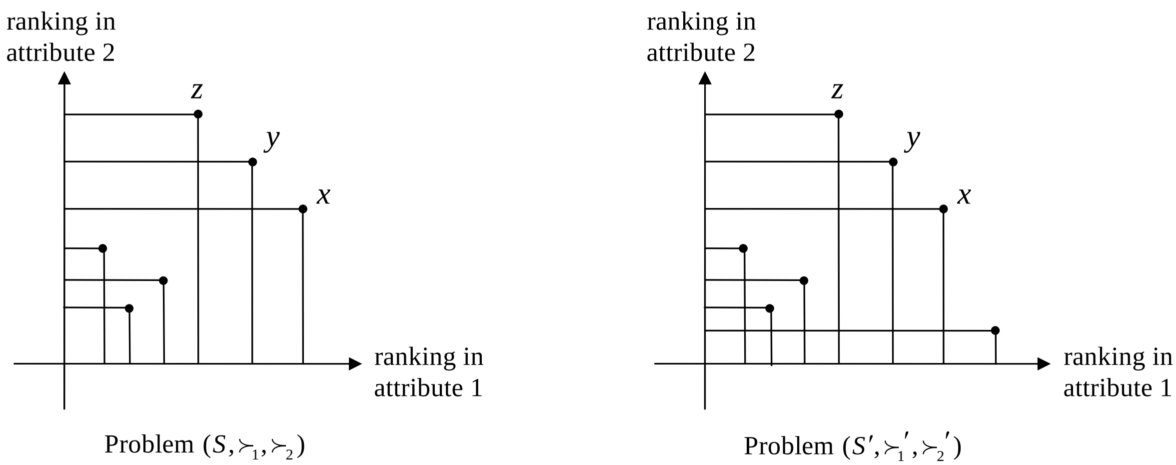

Figure 5: Expansion of a decision problem

THEOREM **6** *The choice rule of the RIDC is the unique one rule that satisfies S-SYM and IIE.*

Remark: Note that, although RIDC's axiomatic characterization does not require Weak Pareto optimality (WPO), RIDC satisfies it nevertheless.

We extend the above axiomatization of the Endogenous-RDRC's outcome set to the N-attributes case, where $N > 2$, in the Appendix.

We also establish the independence of the axioms used in the above theorem in the Appendix.

8 Nash Product Convergence of the Outcome Sets of RIDC and Endogenous-RDRC

Consider a comprehensive, convex and compact set $\overline{S} \in R^2_+$. Let $\bar{s}_1 = \max\{s_1 | (s_1, s_2) \in \overline{S}\}$ and $\bar{s}_2 = \max\{s_2 | (s_1, s_2) \in \overline{S}\}$. For any given "cell" parameter n, we can partition $[0, \bar{s}_1] \times [0, \bar{s}_2]$ to $n \times n$ cells such that the length of each cell is $\bar{s}_1/n$ and the height of each cell is $\bar{s}_2/n$. We normalize $\overline{S}$ such that $\bar{s}_1 = \bar{s}_2$ to make this setup even more tractable. Let C^n be the set of the n^2 cells. Let S^n be a (finite) set of alternatives that are uniformly distributed in $\overline{S}$ (with parameter n), in the sense that for each cell $c \in C^n$ with $c \cap \overline{S} \neq \emptyset$, there is one and only one alternative in S^n which belongs to c. The ranking of alternatives in S^n in each attribute can be easily obtained (for example, if $s = (s_1, s_2) \in S^n, s' = (s'_1, s'_2) \in S^n$ and $s_1 < s'_1$, then s' has higher ranking than s in attribute 1). We can use the grid method in Anbarci (1993) to ensure that no two alternatives have the same ranking in any attribute. As n goes to infinity, the feasible set S^n approaches $\overline{S}$.

Let $s^{NS} \in \overline{S}$ be the point that has the maximum Nash product (i.e., $s^{NS} = argmax_{(s_1,s_2)\in\overline{S}} s_1 s_2$). It is easy to see that the RIDC's outcome of S^n and the most probable outcome of Endogenous-RDRC of S^n will converge to s^{NS} as n goes to infinity.

For any $s \in \overline{S}$, let $U(s) = \{t | t \geq s \text{ and } t \in \overline{S}\}$ be the UCS of s, and $L(s) = \{t | t \leq s \text{ and } t \in \overline{S}\}$ be the LCS of s. Let $\partial(U(s))$ be the Pareto frontier of $U(s)$, and $l(\partial(U(s)))$ be the length of $\partial(U(s))$. It is obvious that, the most probable outcome of Exogenous-RDRC procedure will converge to any $s^* = argmax_{s\in\partial S} \int_{x\in L(s)} \frac{1}{l(\partial(U(x)))} dx$. Note, however, that each s^* may be different from s^{NS} if the Pareto frontier of $\overline{S}$ is not linear. For example, if $\overline{S} = \{(s_1, s_2) \in R^2_+ | s_1^2 + s_2^2 \leq 1\}$, then it can be verified that $s^* = \{(0.76, 0.65), (0.65, 0.76)\}$.

9 Relevant Literature

In this section we will focus on work that has not been mentioned in the Introduction. Although is no specific work in the literature that focuses on different heuristics' procedures (and their axiomatic characterizations) that can incorporate the prominent AE phenomenon at different degrees, there are papers that are close to our work in various aspects.

de Clippel and Eliaz (2012) proposed to view the resolution of the decision maker's choice problem as a cooperative solution to an *intra*personal bargaining problem among two "*selves*" of an individual, each self representing a different rationale (or a different attribute in our setup) for choice. To that end, they provided an axiomatic characterization of a solution concept, namely that of the 'fallback solution', which is very relevant for our work.[24] In some setups, the fallback solution can account for the AE as well as another important phenomenon, namely the *Compromise Effect* (CE), which arises when a decision maker tends to choose an intermediate option in her choice set.[25] When both the AE

[24]The fallback position's outcome set coincides with the subgame-perfect Nash equilibrium outcomes of the 'Voting by Alternating Offers and Vetoes' (VAOV) procedure of Anbarci (1993) and 'Honoring Past Concessions' (HPC) procedure of Anbarci (2006).

In both the VAOV and HPC procedures, two players take turns making offers until an alternative is accepted, and if no offer is accepted, the last remaining alternative is the outcome. In addition, in VAOV an offer rejected by a player is taken out of consideration, while in HPC all offers are on the table until one is accepted; at any stage, a Player i can either (1) accept the last offer of Player j or (2) accept any of the previous offers made by Player j.

[25]The CE dates back to Simonson (1989). Also see Chernev (2004), Dhar, Menon and Maach (2004), Doyle,

and CE are present with distinct alternatives as in the next figure, however, the fallback solution clearly favors the CE.[26]

In a companion paper (Anbarci and Rong, 2016) we provide a heuristic (and its axiomatic characterization) which favors the CE; in that sense, Anbarci and Rong (2016) is the closest work to ours. Anbarci and Rong (2016) first considers the case with two attributes, $A1$ and $A2$, such that the decision maker does not have a preference between them. In their heuristic's first procedure, the decision maker first eliminates the worst alternative in terms of $A1$ as well as the worst alternative in terms of $A2$. Then, among the remaining alternatives, she repeats the same kind of elimination, until at least one and at most two Pareto alternatives remain. If only one Pareto alternative remains, it will be her unambiguous choice outcome. In case two such alternatives remain, then she will pick one of them randomly as the choice outcome (their heuristic's second procedure considers the case where the DM does have a preference between the attributes).

Kamenica (2008) studies an application of the CE in which there is a market with rational consumers who are either informed or uninformed. Uninformed consumers exhibit a behavior that conforms with the CE (and choice overload). When many consumers are uninformed, the firm may try to manipulate consumers' beliefs by introducing premium loss leaders (expensive goods of overly high quality that increase the demand for other goods). Consequently, in this model, the CE emerges as equilibrium behavior in a specific market environment.

Going back to the literature on the AE, the framework of Ok et al. (2015) too can incorporate the AE. Their analysis, however, is of revealed preference that starts with a decision maker's observed choice behavior and intends to retrieve not only her underlying preferences but also her reference point (if any). In their model, for every feasible set, their model specifies whether there is a reference point, or not, and, if there is one, which one it is. When there is no reference point, then the agent acts fully rationally by choosing the elements that maximize her utility in that set. If, on the other hand, there is a reference point, then the agent instead only considers the available options that dominate the reference point; amongst those alternatives too, the decision maker is fully rational,

O'Connor, Reynolds and Bottomley (1999), Kivetz, Netzer and Srinivasan (2004), Mourali, Bockenholt and Laroche (2007), Sheng, Parker and Nakamoto (2005), Simonson and Tversky (1992), among others.

To illustrate the CE, suppose a decision maker were equally likely to choose between A and B, where A has a better price and worse speed than B. When a new PC, C, is added, such that C has a better price and worse speed than B but has a worse price and better speed than A, then most subjects tend to choose C.

[26] To illustrate the intuition behind the distinction between the AE and CE alternatives, consider the $AA' - CC'$ example. Then C and C' arise as the *pure CE alternatives* (while A and A' arise as the AE alternatives with the largest LCSs). A pure compromise alternative is a *median undominated alternative* with an empty LCS. In a nutshell, informally, the fallback solution's outcome set involves the alternative(s) with the closest rankings *among the undominated alternatives*. In the $AA' - CC'$ example, the fallback solution therefore selects the pure CE alternatives C and C instead of the AE alternatives A and A'.

choosing the alternatives maximizing her utility.

Related to Ok, Ortoleva and Riella (2015), its companion applied work, Ok, Ortoleva and Riella (2014), provides an application in the context of monopolistic vertical product differentiation. That paper considers the standard screening problem of a monopolist who chooses price-quality bundles to offer to consumers whose quality valuations are their private information. The model, however, also allows a fraction of the customers to be reference dependent. They find that, the AE would be exploited by the monopolist in equilibrium to overcome the incentive-compatibility constraints under some parametric restrictions.

Natenzon (2015) has recently introduced a random choice model in which a relatively uninformed decision maker gradually learns her utility of available alternatives. More specifically, based on noisy signals, the decision maker updates a symmetric prior using Bayes' rule which allows her to choose the alternative with highest posterior expected utility. Natenzon then shows how this process can incorporate the AE. As a key departure from random utility models, however, Natenzon's decision maker sometimes picks an alternative that does not have the highest signal realization.

Some recent random choice models also consider random consideration-set rules (Manzini and Mariotti, 2014) and random preferences over attributes (Gul et al., 2014). Manzini and Mariotti (2014) axiomatized a model where a boundedly-rational decision maker maximizes a preference relation but also makes random choice errors due to imperfect attention. In Gul et al. (2014)'s model of random choice of the Luce form, object attributes are obtained endogenously from the observed choices of a decision maker with stochastic preferences over attributes. Their model seeks to address the similarity effect, where 'options that share attributes compete more for the decision maker's attention than those that do not share attributes' (Tversky, 1972). These standard random choice models cannot account for the AE, however, mainly due to their adherence to the regularity (monotonicity) property which requires that the choice probabilities of any of the existing alternatives cannot increase whenever a new alternative is added to them.

10 Concluding Remarks

The differences in the outcomes of our three heuristics' procedures incorporating the AE and the vast variety of non-overlapping axioms in their axiomatizations, in a sense, imply a strong degree of theoretical robustness for the AE phenomenon. In addition, as noted before, to our knowledge, our paper is the first one that provides cooperative foundations to heuristics incorporating the AE as well as the first one that provides cooperative foundations to stochastic outcomes where the decision maker is not assumed to have stochastic preferences (a la Gul et al., 2014) or make random choice errors (a la

Manzini and Mariotti, 2014).

Next, we would like to allude to the common features between the AE and CE phenomena as well as their main differences. Most importantly, there are clear links between 'trade-off aversion', which was mentioned in the Introduction, and both of the AE and CE phenomena. It is not difficult to see that our heuristics' procedures exhibit trade-off aversion. Also note that, recalling the CE's brief discussion in the previous section, trade-off avoidance is a driver of the CE as well (as both de Clippel and Eliaz's fallback solution and Anbarci and Rong's, 2016, solution are linked to heuristics of which procedures involve an elimination contest between attributes, totally bypassing any trade-offs among multi-attribute options).

As mentioned in the Introduction, our heuristics' procedures studied in this paper simply involve rather effortless/costless or random selections among Pareto alternatives typically relying on a reference alternative, while procedures of heuristics that incorporate the CE involve elimination contests between alternatives according to their rankings in terms of different attributes/rationales without exhibiting any reliance on any reference alternatives or any random selections among alternatives.

Adopting a limited (i.e., depletable) resource approach to decision-making, it has been recently noted in the psychology and marketing literatures that even mundane binary one-attribute choices can be somehow cognitively depleting (Vohs et al., 2008). Further, Wang et al. (2010) have also established that the larger the trade-offs, the greater is this depletion effect. As such, findings on trade-off aversion (mentioned in the Introduction and above) also relate to Shugan (1980) who was first to come up with a formal theory of a 'cost of thinking'.[27] The essence of Shugan's theory is follows: 'a conjunctive comparison is less costly than a compensatory comparison'. In other words, (i) comparing any two alternatives which Pareto dominate each other would be a much simpler task for the decision maker than (ii) comparing any two alternatives that do not Pareto-dominate each other, i.e., comparing two alternatives that pose trade-offs. Thus, as the difference between the cost of (ii) and that of (i) increases for a decision maker, the AE and CE should tend to become more prominent outcomes for her.

[27]Shugan rightfully maintained that "[p]reference only partially influences choice by determining benefits. However, the determination has costs - rife information, numerous alternatives, time pressure, the consumer's limited information processing capabilities, and the general effort exerted to solve the problem. Generally, the net utility of finding the best product from one set of products may be different from the net utility of finding it as best from another set of products. That is, there may be a cost associated with the act of making a decision" (1980, p. 100).

11 Appendix

11.1 Proofs of the Results in the Main Text

Proof of Theorem 1:

Note that $J_i(y,S) \subseteq L(y,S)$ is the set of alternatives that are dominated by exactly i Pareto alternatives (including y). We thus have $\cup_{i=1}^{\#\partial S} J_i(y,S) = L(y,S)$ and $J_i(y,S)$ and $J_j(y,S)$ must be disjoint for any $i \neq j$ and $i, j \in \{1, ..., \#\partial S\}$ ($J_i(y,S)$ may be an empty set for some $i \in \{1, ..., \#\partial S\}$). Conditional on the event that a given alternative in $J_i(y,S)$ is chosen as the reference alternative, the probability that the decision maker will eventually choose y as the outcome is $\frac{1}{i}$. Also, note that each dominated alternative will be chosen with probability $\frac{1}{\#\eta S}$. So, the total probability that y is the outcome is $\frac{1}{\#\eta S} \sum_{i=1}^{\#\partial S} \#J_i(y,S)\frac{1}{i}$.

Proof of Theorem 2:

Proof: Let $p(x)$ be the probability that a dominated alternative x is selected as the reference alternative. Since $p(x)$ is proportional to $\#\partial U(x,S)$, we have $p(x) = c\#\partial U(x,S)$ for all $x \in \eta S$ for some constant c. So, for any $y \in \partial S$, we have $q^{en}(y) = \sum_{x \in L(y,S)} p(x) \frac{1}{\#\partial U(x,S)} = \sum_{x \in L(y,S)} c\#\partial U(x,S) \frac{1}{\#\partial U(x,S)} = c\#L(y,S)$ (i.e., the choice probability of y is proportional to the cardinality of y's LCS). Since $\sum_{y \in \partial S} q^{en}(y) = 1$, we have $c = \frac{1}{\sum_{y \in \partial S} \#L(y,S)} = \frac{1}{I(S)}$. This implies that $q^{en}(y) = \frac{\#L(y,S)}{I(S)}$.

Proof of Lemma 1:

Suppose there are n alternatives in $\partial(S, \succ_1, \succ_2)$. Since $(S, \succ_1, \succ_2)$ is such that all dominated alternatives in S are dominated by all Pareto alternatives in S, we must have $R_1(s|S, \succ_1, \succ_2) > n$ and $R_2(s|S, \succ_1, \succ_2) > n$ for any $s \in \eta(S, \succ_1, \succ_2)$, and $R_1(s|S, \succ_1, \succ_2) \leq n$ and $R_2(s|S, \succ_1, \succ_2) \leq n$ for any $s \in \partial(S, \succ_1, \succ_2)$. Let s^* be the alternative that has the highest ranking in terms of the first attribute. We will show that s^* must be ranked n-th highest in the second attribute. If not, suppose $R_2(s^*|S, \succ_1, \succ_2) < n$. Let s' be the alternative in S such that the ranking of the alternative in terms of the second attribute is n. We must have $s' \in \partial(S, \succ_1, \succ_2)$ (because for any $s \in \eta(S, \succ_1, \succ_2)$, we have $R_1(s|S, \succ_1, \succ_2) > n$ and $R_2(s|S, \succ_1, \succ_2) > n$). However, since s' ranks lower than s^* in both attributes, s' must be a dominated alternative. This is a contradiction with the fact that $s' \in \partial(S, \succ_1, \succ_2)$. So, it must be true that $R_2(s^*|S, \succ_1, \succ_2) = n$.

Similarly, we can show that the alternative that ranks second in terms of the first attribute must rank $n-1$-th in terms of the second attribute. In general, we can show that the alternative that ranks i-th in terms of the first attribute must rank $n-i+1$-th in terms of the second attribute for any $i \in \{1, ..., n\}$. This implies that the set of Pareto alternatives, $\partial(S, \succ_1, \succ_2)$, must be symmetric.

Proof of Lemma 2:

Suppose $(S, \succ_1, \succ_2) \in \mathbb{D}_2$ is symmetric and $L(y, S, \succ_1, \succ_2) = L(z, S, \succ_1, \succ_2)$ for any $y, z \in \partial(S, \succ_1, \succ_2)$. We will show that $q(y, S, \succ_1, \succ_2) = q(z, S, \succ_1, \succ_2)$ for any $y, z \in \partial(S, \succ_1, \succ_2)$.

For simplicity, we assume that $L(y, S, \succ_1, \succ_2)$ contains only one alternative (the proof for the general case where $L(y, S, \succ_1, \succ_2)$ contains more than one alternative is similar). Let $\#\partial(S, \succ_1, \succ_2) = m$. To illustrate the idea of the proof, we next consider the case where $m = 2$ and $m = 3$ (the proof for the more general case where $m > 3$ is similar and is omitted).

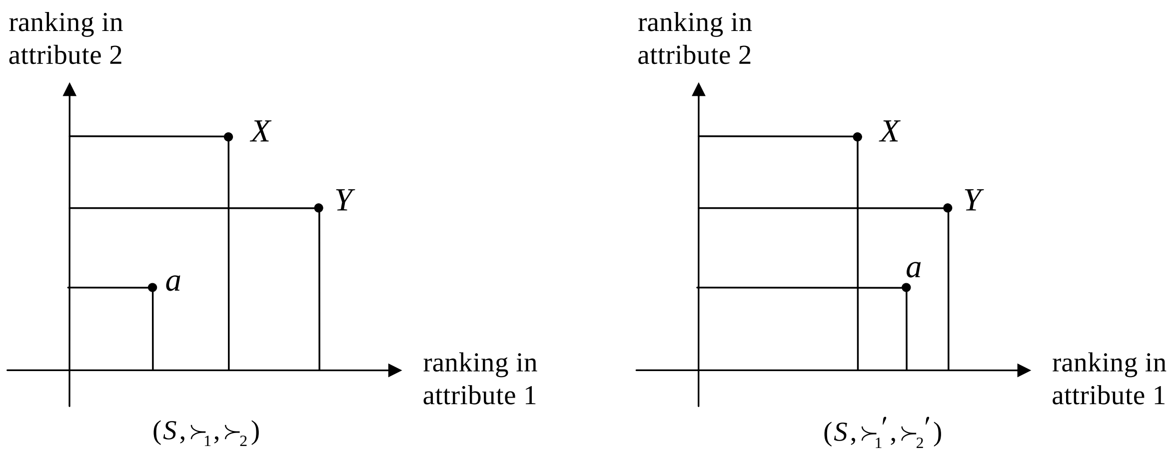

Figure 6

Assume that $m = 2$. Refer to Figure 6. The change from $(S, \succ_1, \succ_2)$ to $(S, \succ'_1, \succ'_2)$ forms a JD setup (in particular, a, which was initially dominated by both X and Y, becomes dominated by Y only). Using LL, we have $\Delta q(X, S, \succ'_1, \succ'_2 | \succ_1, \succ_2) = -\frac{1}{n \# \eta S} = -\frac{1}{2}$. Using WPO, we have $\Delta q(Y, S, \succ'_1, \succ'_2 | \succ_1, \succ_2) = -\Delta q(X, S, \succ'_1, \succ'_2 | \succ_1, \succ_2) = \frac{1}{2}$. The above two equations then imply that $q(X, S, \succ_1, \succ_2) \geq \dfrac{1}{2}$ and $q(Y, S, \succ_1, \succ_2) \leq \dfrac{1}{2}$ (otherwise, we have $q(X, S, \succ'_1, \succ'_2) < 0$ and $q(Y, S, \succ'_1, \succ'_2) > 1$). Now, we consider the JD setup illustrated in Figure 7, then using LL and WPO, we have $\Delta q(X, S, \succ'_1, \succ'_2 | \succ_1, \succ_2) = \frac{1}{2}$ and $\Delta q(Y, S, \succ'_1, \succ'_2 | \succ_1, \succ_2) = -\frac{1}{2}$. This implies that $q(X, S, \succ_1, \succ_2) \leq \dfrac{1}{2}$ and $q(Y, S, \succ_1, \succ_2) \geq \dfrac{1}{2}$. We thus have $q(X, S, \succ_1, \succ_2) = q(Y, S, \succ_1, \succ_2) = \dfrac{1}{2}$.

Assume that $m = 3$. Refer to Figure 8. The change from $(S, \succ_1, \succ_2)$ to $(S, \succ'_1, \succ'_2)$ forms a JD setup (in particular, a—which is initially dominated by X, Y and Z—becomes dominated by Y and Z only). Using LL, we have $\Delta q(X, S, \succ'_1, \succ'_2 | \succ_1, \succ_2) = -\frac{1}{n \# \eta S} = -\frac{1}{3}$. Using WPO and WSG, we have $\Delta q(Y, S, \succ'_1, \succ'_2 | \succ_1, \succ_2) = \Delta q(Z, S, \succ'_1, \succ'_2 | \succ_1, \succ_2) = \frac{1}{6}$. In addition, note that $q(Y, S, \succ'_1, \succ'_2) = q(Z, S, \succ'_1, \succ'_2) = \dfrac{1}{2}$ (the proof is similar to the case of $m = 2$)[28]—

[28] In particular, consider the JD setup where a is initially jointly dominated by Y and Z, but then becomes

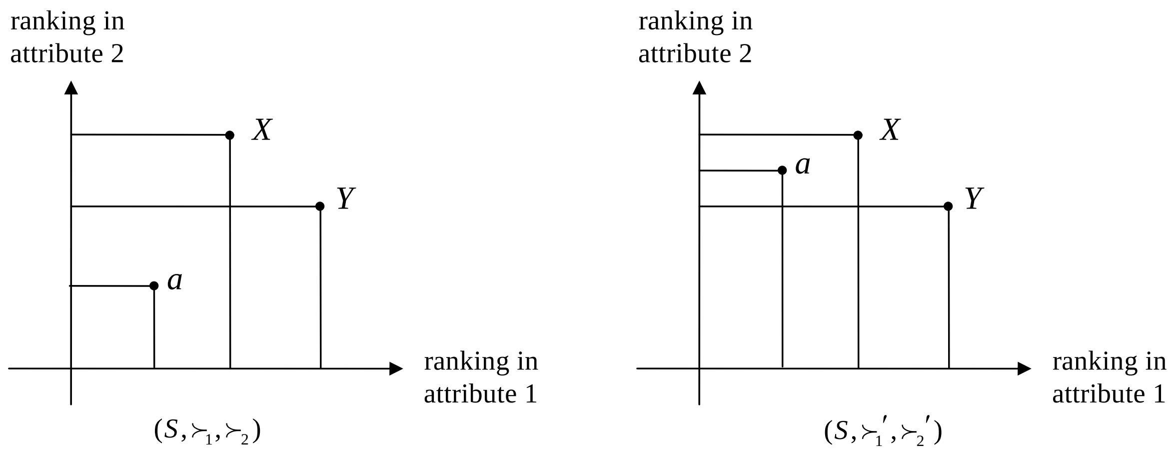

Figure 7

which also implies that $q(X, S, \succ_1', \succ_2') = 0$. We thus have $q(X, S, \succ_1, \succ_2) = q(Y, S, \succ_1, \succ_2) = q(Z, S, \succ_1, \succ_2) = \frac{1}{3}$.

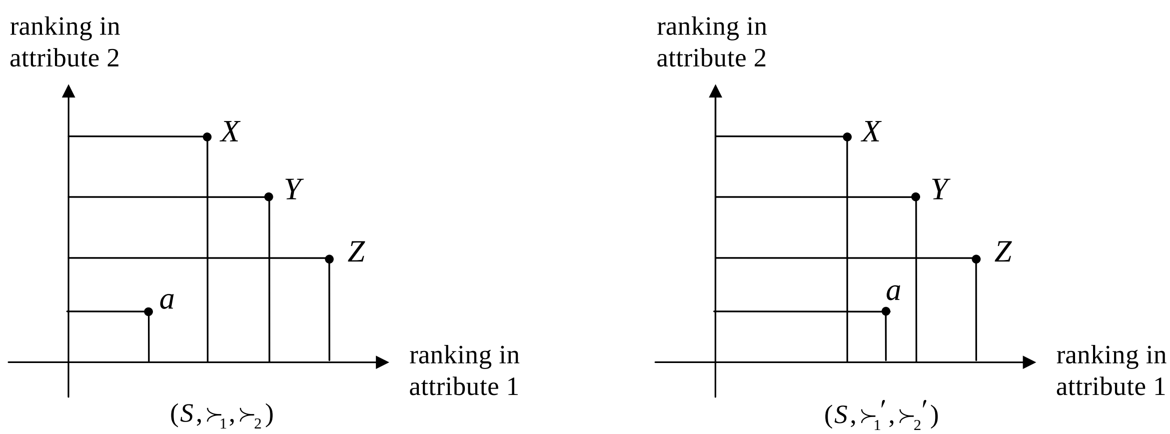

Figure 8

Proof of Lemma 3:

Suppose that $a \in J(y_1, ..., y_l, S, \succ_1, \succ_2)$ for some $\{y_1, ..., y_l\} \subset \partial(S, \succ_1, \succ_2)$. We re-label $\{y_1, ..., y_l\}$ such that $R_1(y_1|S, \succ_1, \succ_2) < R_1(y_2|S, \succ_1, \succ_2) < ... < R_1(y_l|S, \succ_1, \succ_2)$. Note that for

exclusively dominated by Z, then using WPO, LL, and IUA, we can show that $q(Y, S, \succ_1', \succ_2') \geq \dfrac{1}{2}$ and $q(Z, S, \succ_1', \succ_2') \leq \dfrac{1}{2}$. Consider the JD setup where where a is initially jointly dominated by Y and Z, but then becomes exclusively dominated by Y, then using WPO, LL, and IUA, we can show that $q(Z, S, \succ_1', \succ_2') \geq \dfrac{1}{2}$ and $q(Y, S, \succ_1', \succ_2') \leq \dfrac{1}{2}$. Thus, we have $q(Z, S, \succ_1', \succ_2') = \dfrac{1}{2}$ and $q(Y, S, \succ_1', \succ_2') = \dfrac{1}{2}$.

any two Pareto alternatives, y and y', we have $R_1(y|S,\succ_1,\succ_2) > R_1(y'|S,\succ_1,\succ_2)$ if and only if $R_2(y|S,\succ_1,\succ_2) < R_2(y'|S,\succ_1,\succ_2)$. Thus, we must have $R_2(y_1|S,\succ_1,\succ_2) > R_2(y_2|S,\succ_1,\succ_2) > ... > R_2(y_l|S,\succ_1,\succ_2)$ since $\{y_1,...,y_l\}$ are all Pareto alternatives. Since a is dominated by $y_1,...,y_l$, we must have $R_1(a|S,\succ_1,\succ_2) > R_1(y_l|S,\succ_1,\succ_2)$ and $R_2(a|S,\succ_1,\succ_2) > R_2(y_1|S,\succ_1,\succ_2)$.

Let $\hat{S} = \{y_1,...,y_l\}$ and $\hat{S}^c = \partial(S,\succ_1,\succ_2)\backslash\hat{S}$. Since $\{y_1,...,y_l\} \subset \partial(S,\succ_1,\succ_2)$, $\hat{S}^c$ must be a non-empty set. For any alternative $y \in \hat{S}^c$, we must have either $R_1(y|S,\succ_1,\succ_2) < R_1(y_1|S,\succ_1,\succ_2)$ or $R_1(y|S,\succ_1,\succ_2) > R_1(y_l|S,\succ_1,\succ_2)$ (suppose not, then there exists some $y \in \hat{S}^c$ such that $R_1(y_1|S,\succ_1,\succ_2) < R_1(y|S,\succ_1,\succ_2) < R_1(y_l|S,\succ_1,\succ_2)$, which implies that $R_2(y_1|S,\succ_1,\succ_2) > R_1(y|S,\succ_1,\succ_2) > R_1(y_l|S,\succ_1,\succ_2)$; this will be a contradiction with the fact that a is not dominated by y - noting that we have $R_1(a|S,\succ_1,\succ_2) > R_1(y_l|S,\succ_1,\succ_2)$ and $R_2(a|S,\succ_1,\succ_2) > R_2(y_1|S,\succ_1,\succ_2)$). So, either the set $\hat{S}^c_1 = \{y \in \hat{S}^c | R_1(y|S,\succ_1,\succ_2) > R_1(y_l|S,\succ_1,\succ_2)\}$ or the set $\hat{S}^c_2 = \{y \in \hat{S}^c | R_1(y|S,\succ_1,\succ_2) < R_1(y_1|S,\succ_1,\succ_2)\}$ is non-empty. Suppose the first set is non-empty (the proof for the case where the second set is non-empty is similar). Let $y^* \in \hat{S}^c_1$ be the alternative such that the ranking of y^* is the highest among all alternatives in $\hat{S}^c_1$ (i.e., $R_1(y|S,\succ_1,\succ_2)$ is the lowest for $y = y^*$ among all alternatives in $\hat{S}^c_1$). Define a new attribute ranking $(\succ'_1,\succ'_2)$ such that $R_i(y|S\backslash a,\succ'_1,\succ'_2) = R_i(y|S\backslash a,\succ_1,\succ_2)$ for any $y \in S\backslash a$ and any $i \in \{1,2\}$, and $R_1(a|S,\succ'_1,\succ'_2) = R_1(y^*|S,\succ'_1,\succ'_2) + 1$ and $R_2(a|S,\succ'_1,\succ'_2) = R_2(y^*|S,\succ'_1,\succ'_2)$. Then a will be dominated by $y_1,...,y_l$ and y^*, and not dominated by any other Pareto alternatives.

Proof of Lemma 4:

We provide a sketch of proof.[29] Suppose a choice rule q satisfies WPO, WSG, IUA, and LL. We will show that q satisfies IIS. It is sufficient to show that for any two decision problems $(S,\succ_1,\succ_2)$ and $(S,\succ'_1,\succ'_2)$ where $\partial(S,\succ_1,\succ_2) = \partial(S,\succ'_1,\succ'_2)$ and $R_i(y|\partial(S,\succ_1,\succ_2),\succ_1,\succ_2) = R_i(y|\partial(S,\succ'_1,\succ'_2),\succ'_1,\succ'_2)$ for any $y \in \partial(S,\succ_1,\succ_2) = \partial(S,\succ'_1,\succ'_2)$ and any $i \in \{1,2\}$, and $J(y_1,...,y_l,S,\succ_1,\succ_2) = J(y_1,...,y_l,S,\succ'_1,\succ'_2)$ for any $\{y_1,...,y_l\} \subseteq \partial(S,\succ_1,\succ_2) = \partial(S,\succ'_1,\succ'_2)$, then $q(y,S,\succ_1,\succ_2) = q(y,S,\succ'_1,\succ'_2)$ for any $y \in S$. It is sufficient to show the following: for any two decision problems $(S,\succ_1,\succ_2)$ and $(S,\succ'_1,\succ'_2)$ where $\partial(S,\succ_1,\succ_2) = \partial(S,\succ'_1,\succ'_2)$ and $R_i(y|\partial(S,\succ_1,\succ_2),\succ_1,\succ_2) = R_i(y|\partial(S,\succ'_1,\succ'_2),\succ'_1,\succ'_2)$ for any $y \in \partial(S,\succ_1,\succ_2) = \partial(S,\succ'_1,\succ'_2)$ and any $i \in \{1,2\}$, and $J(y_1,...,y_l,S,\succ_1,\succ_2) = J(y_1,...,y_l,S,\succ'_1,\succ'_2)$ for *some* $\{y_1,...,y_l\} \subseteq \partial(S,\succ_1,\succ_2) = \partial(S,\succ'_1,\succ'_2)$ and $R_i(y|S\backslash J(y_1,...,y_l,S,\succ_1,\succ_2),\succ_1,\succ_2) = R_i(y|S\backslash J(y_1,...,y_l,S,\succ_1,\succ_2),\succ'_1,\succ'_2)$ for any $y \notin J(y_1,...,y_l,S,\succ_1,\succ_2)$, then $q(y,S,\succ_1,\succ_2) = q(y,S,\succ'_1,\succ'_2)$ for any $y \in S$.[30]

The idea of the proof is as follows. We first consider the case where $\{y_1,...,y_l\} \subset$

[29] The full proof can be obtained from the authors.

[30] This is because for any two decision problems $(S,\succ_1,\succ_2)$ and $(S,\succ'_1,\succ'_2)$ where $J(y_1,...,y_l,S,\succ_1,\succ_2) = J(y_1,...,y_l,S,\succ'_1,\succ'_2)$ for *any* $\{y_1,...,y_l\} \subseteq \partial(S,\succ_1,\succ_2) = \partial(S,\succ'_1,\succ'_2)$, we can always find a sequence of decision problems where the first problem is $(S,\succ_1,\succ_2)$ and the last problem is $(S,\succ'_1,\succ'_2)$ and any two consecutive problems in the sequence only differ in that $J(y_1,...,y_l,S,\succ_1,\succ_2) = J(y_1,...,y_l,S,\succ'_1,\succ'_2)$ for *some* $\{y_1,...,y_l\}$ in the Pareto set (and all other alternatives of S have the same relative rankings in the two consecutive problems).

$\partial(S,\succ_1,\succ_2)$. Pick any $y \in J(y_1,...,y_l,S,\succ_1,\succ_2) = J(y_1,...,y_l,S,\succ_1',\succ_2')$, then according to the proof of Lemma 3, we can transform the problem $(S,\succ_1,\succ_2)$ to a problem such that y is dominated by $y_1,...,y_l$ and some additional Pareto alternative y^*. Similarly, we can transform the problem $(S,\succ_1',\succ_2')$ to a problem such that y is dominated by $y_1,...,y_l$ and y^*. In addition, we can choose the transformations such that for each of the above two transformed problems, y is ranked highest in both attributes among all alternatives that are dominated by $y_1,...,y_l,y^*$. We then pick another alternative in $J(y_1,...,y_l,S,\succ_1,\succ_2)$ (if there is any), say y' (where $y' \neq y$), and make further transformations on the above two transformed problems such that in each problem, y' is dominated by $y_1,...,y_n,y^*$ and y' is ranked the second highest in both attributes among all alternatives that are dominated by $y_1,...,y_l,y^*$. We may continue the above procedure until for both $(S,\succ_1,\succ_2)$ and $(S,\succ_1',\succ_2')$, all alternatives that are originally dominated by $y_1,...,y_l$ are transformed such that they are all dominated by $y_1,...,y_l,y^*$. We use $(S,\hat{\succ}_1,\hat{\succ}_2)$ and $(S,\hat{\succ}_1',\hat{\succ}_2')$ to denote the two transformed problems respectively. Noting that the rankings of *all* alternatives of S are the same in the two transformed problems $(S,\hat{\succ}_1,\hat{\succ}_2)$ and $(S,\hat{\succ}_1',\hat{\succ}_2')$, the choice probabilities in the two problems must be the same. According to WSG, IUA, and LL, the change of the choice probabilities of the above two transformations (i.e., the change of choice probabilities when the decision problem changes from $(S,\succ_1,\succ_2)$ to $(S,\hat{\succ}_1,\hat{\succ}_2)$ and the change of the choice probabilities when the decision problem changes from $(S,\succ_1',\succ_2')$ to $(S,\hat{\succ}_1',\hat{\succ}_2')$) must be the same. This implies that the choice probabilities for the two problems $(S,\succ_1,\succ_2)$ and $(S,\succ_1',\succ_2')$ are the same.

Now, we consider the case where $\{y_1,...,y_l\} = \partial(S,\succ_1,\succ_2)$. We can transform both $(S,\succ_1,\succ_2)$ and $(S,\succ_1',\succ_2')$ to problems such that y is dominated by exactly one less Pareto alternatives.[31] The remaining proof is similar to the case where $\{y_1,...,y_l\} \subset \partial(S,\succ_1,\succ_2)$ and is omitted.

Proof of Theorem 4:

Our proof consists of two parts.

Part 1: Let q^* be a choice rule that satisfies the four axioms. Let q^r be the choice rule induced by the Exogenous-RDRC procedure. We will show that the choice rule q^* must be the same as the choice rule q^r. We first introduce some notation. For any $m \in \mathbb{N}$ and $n \in \mathbb{N} \cup \{0\}$, let $\mathbb{D}_{m,n} = \{(S,\succ_1,\succ_2) | \#\partial(S,\succ_1,\succ_2) = m, \#\eta(S,\succ_1,\succ_2) = n\}$, i.e., $\mathbb{D}_{m,n}$ is the collection of all finite decision problems where the number of Pareto alternatives is m and the number of dominated alternatives is n. For any given $(S,\succ_1,\succ_2)$, we define $I(S,\succ_1,\succ_2) = \sum_{y\in\partial(S,\succ_1,\succ_2)} \#L(y,S,\succ_1,\succ_2)$. We call $I(S,\succ_1,\succ_2)$ the *degree of joint dominance* of $(S,\succ_1,\succ_2)$. Note that for any $(S,\succ_1,\succ_2) \in \mathbb{D}_{m,n}$, we must have $n \leq I(S,\succ_1,\succ_2) \leq mn$. For any $n \leq k \leq mn$, let $\mathbb{D}^k_{m,n} = \{(S,\succ_1,\succ_2) | (S,\succ_1,\succ_2) \in \mathbb{D}_{m,n}$ and $I(S,\succ_1,\succ_2) = k\}$ be the set of finite

[31]The only case where this is not possible is that $\partial(S,\succ_1,\succ_2)$ only contains one undominated alternative. In this case, WPO implies that the only undominated alternative must be chosen with probability one.

decision problems where the number of Pareto alternatives is m, the number of dominated alternatives is n, *and* the degree of joint dominance is k. Note that $\mathbb{D}_{m,n} = \cup_{k=n}^{mn} \mathbb{D}_{m,n}^k$.

We fix an (arbitrary) $m \in \mathbb{N}$ and an (arbitrary) $n \in \mathbb{N} \cup \{0\}$. We next show that for any $(S, \succ_1, \succ_2) \in \mathbb{D}_{m,n}$, we have $q^*(y, S, \succ_1, \succ_2) = q^r(y, S, \succ_1, \succ_2)$ for any $y \in S$. The proof is by induction.

(i) We will show that for any $(S, \succ_1, \succ_2) \in \mathbb{D}_{m,n}^{mn}$, we have $q^*(y, S, \succ_1, \succ_2) = q^r(y, S, \succ_1, \succ_2)$ for any $y \in S$. Note that for any $(S, \succ_1, \succ_2) \in \mathbb{D}_{m,n}^{mn}$, we have $I(S, \succ_1, \succ_2) = mn$, i.e., all n dominated alternatives are jointly dominated by all of the m Pareto alternatives. We have the following two subcases:

(a) $(S, \succ_1, \succ_2)$ is symmetric. In this case, since q^* satisfies WPO and SYM (which is implied by WPO, WSG, IUA and LL according to Lemma 2), we must have $q^*(y, S, \succ_1, \succ_2) = 1/m$ for any $y \in \partial(S, \succ_1, \succ_2)$ and $q^*(y, \succ_1, \succ_2) = 0$ for any $y \in \eta(S, \succ_1, \succ_2)$. So, we have $q^*(y, S, \succ_1, \succ_2) = q^r(y, S, \succ_1, \succ_2)$ for any $y \in S$.

(b) If $(S, \succ_1, \succ_2)$ is not symmetric, then we can transform it to a symmetric problem (see Figure 9 for an illustration, in which $S = \{s^1, s^2, s^3, s^4, s^5, s^6\}$; note that the Pareto set of S is symmetric according to Lemma 1). Then, using Lemma 4 and SYM (which is implied by WPO, WSG, IUA and LL), we must have $q^*(y, S, \succ_1, \succ_2) = 1/m$ for any $y \in \partial(S, \succ_1, \succ_2)$. So, we have $q^*(y, S, \succ_1, \succ_2) = q^r(y, S, \succ_1, \succ_2)$ for any $y \in S$.

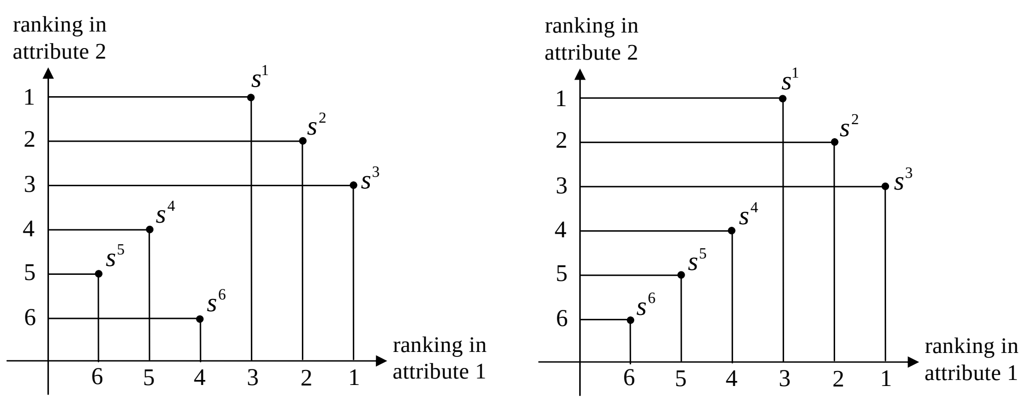

Figure 9: Transformation to a symmetric problem.

(ii) Suppose we have shown that for any $(S, \succ_1, \succ_2) \in \mathbb{D}_{m,n}^k$ where $n < k \leq mn$, we have $q^*(y, S, \succ_1, \succ_2) = q^r(y, S, \succ_1, \succ_2)$ for any $y \in S$, we next show that for any $(S, \succ_1, \succ_2) \in \mathbb{D}_{m,n}^{k-1}$, we have $q^*(y, S, \succ_1, \succ_2) = q^r(y, S, \succ_1, \succ_2)$ for any $y \in S$.

Pick any $(S, \succ_1, \succ_2) \in D_{m,n}^{k-1}$. Since $k - 1 < nm$, there must exist a dominated alternative such that the alternative is not dominated by all Pareto alternatives. Let $a^* \in \eta(S, \succ_1, \succ_2)$ be such an alternative. Let $\hat{S} = \{y_1, ..., y_l\} \subset \partial(S, \succ_1, \succ_2)$ be the collection of all Pareto alternatives that dominate a^*. That is $a^* \in J(y_1, ..., y_l, S, \succ_1, \succ_2)$. Let $\hat{S}^c = \partial(S, \succ_1, \succ_2) \backslash \hat{S}$, which is a non-empty set. Using Lemma 3, there must exist an attribute ranking profile $(\succ_1', \succ_2')$ and some alternative $y^* \in \hat{S}^c$ such that (i) $a^* \in J(y_1, ..., y_l, y^*, S, \succ_1', \succ_2')$, and (ii) $R_i(y|S\backslash a^*, \succ_1$

$,\succ_2) = R_i(y|S\backslash a^*, \succ_1', \succ_2')$ for any $y \in S\backslash a^*$ and any $i \in \{1,2\}$ (see Figure 10 for an illustration, where $S = \{s^1, s^2, s^3, s^4, s^5, s^6\}$, $y^* = s^1$, $a^* = s^5$ and $\hat{S} = \{s^2, s^3\}$). According to WSG, IUA and LL, we have $\Delta q^*(y_i, S, \succ_1', \succ_2'|\succ_1, \succ_2) = -\frac{1}{l(l+1)\#\eta S}$ for $i = 1, ..., l$, $\Delta q^*(y^*, S, \succ_1', \succ_2'|\succ_1, \succ_2) = \frac{1}{(l+1)\#\eta S}$, and $\Delta q^*(y, S, \succ_1', \succ_2'|\succ_1, \succ_2) = 0$ for any $y \in \partial(S, \succ_1, \succ_2)$ with $y \notin \{y_1, ..., y_l, y^*\}$. Also, it is easy to verify that $\Delta q^r(y_i, S, \succ_1', \succ_2'|\succ_1, \succ_2) = -\frac{1}{l(l+1)\#\eta S}$ for $i = 1, ..., l$, $\Delta q^r(y^*, S, \succ_1', \succ_2'|\succ_1, \succ_2) = \frac{1}{(l+1)\#\eta S}$, and $\Delta q^r(y, S, \succ_1', \succ_2'|\succ_1, \succ_2) = 0$ for any $y \in \partial(S, \succ_1, \succ_2)$ with $y \notin \{y_1, ..., y_l, y^*\}$. That is, $\Delta q^*(y, S, \succ_1', \succ_2'|\succ_1, \succ_2) = \Delta q^r(y, S, \succ_1', \succ_2'|\succ_1, \succ_2)$ for any $y \in S$.[32] In addition, it is obvious that $I(S, \succ_1', \succ_2') = k$, so $q^*(y, S, \succ_1', \succ_2') = q^r(y, S, \succ_1', \succ_2')$ for any $y \in S$ by the induction assumption. So, we must have $q^*(y, S, \succ_1, \succ_2) = q^r(y, S, \succ_1, \succ_2)$ for any $y \in S$.

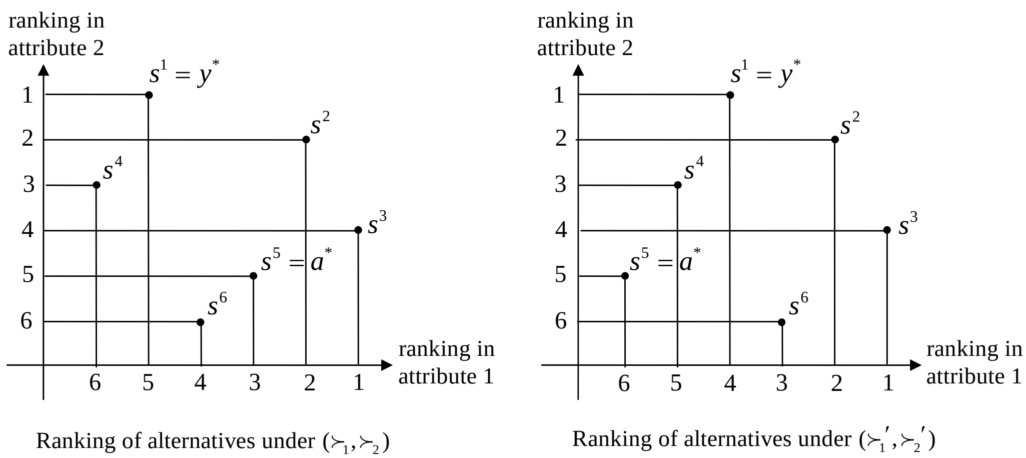

Figure 10: Change of attributes' rankings of a^*.

So, we have shown that for any given $m \in \mathbb{N}$ and any given $n \in \mathbb{N} \cup \{0\}$, we have $q^*(y, S, \succ_1, \succ_2) = q^r(y, S, \succ_1, \succ_2)$ for any $(S, \succ_1, \succ_2) \in \mathbb{D}_{m,n}$ and any $y \in S$. Since $\cup_{\{m \in \mathbb{N}, n \in \mathbb{N} \cup \{0\}\}} \mathbb{D}_{m,n} = \mathbb{D}_2$, we thus have shown that $q^*(y, S, \succ_1, \succ_2) = q^r(y, S, \succ_1, \succ_2)$ for any $(S, \succ_1, \succ_2) \in \mathbb{D}_2$.

Part 2: We next show that Exogenous-RDRC satisfies WPO, WSG, IUA, and LL. Exogenous-RDRC satisfies WPO because by construction, Exogenous-RDRC will only assign positive probabilities to Pareto alternatives.

We next show that Exogenous-RDRC satisfies WSG, IUA and LL. Suppose that there is an $a \in \eta(S, \succ_1, \succ_2)$ and $y_1, ..., y_n \in \partial(S, \succ_1, \succ_2)$ such that $a \in J(y_1, ..., y_n, S, \succ_1, \succ_2)$, $a \in J(y_1, ..., y_{n-1}, S, \succ_1', \succ_2')$, and $R_i(y|S\backslash a, \succ_1, \succ_2) = R_i(y|S\backslash a, \succ_1', \succ_2')$ for any $y \in S\backslash a$ and any $i \in \{1,2\}$. Since the relative rankings of all alternatives (except a) remain the same under $(\succ_1, \succ_2)$ and $(\succ_1', \succ_2')$, we have (i) $L(y, S, \succ_1, \succ_2) = L(y, S, \succ_1', \succ_2')$ for any $y \in \{y_1, ..., y_{n-1}\}$, (ii)

[32] It is easy to see that $\Delta q^*(y, S, \succ_1', \succ_2'|\succ_1, \succ_2) = \Delta q^r(y, S, \succ_1', \succ_2'|\succ_1, \succ_2)$ for any $y \in \partial(S, \succ_1, \succ_2)$. We have $\Delta q^*(y, S, \succ_1', \succ_2'|\succ_1, \succ_2) = \Delta q^r(y, S, \succ_1', \succ_2'|\succ_1, \succ_2)$ for any $y \in \eta(S, \succ_1, \succ_2)$, because (i) q^* satisfies WPO and thus assigns zero probability to any dominated alternative and (ii) q^r also assigns zero probability to any dominated alternative.

$L(y,S,\succ_1,\succ_2)=L(y,S,\succ'_1,\succ'_2)\cup\{a\}$ for $y=y_n$, (iii) for any $y\in\{y_1,...,y_n\}$ and any $x\in L(y,S,\succ_1,\succ_2)$ with $x\neq a$, we have $\#\partial(U(x,\succ_1,\succ_2),\succ_1,\succ_2)=\#\partial(U(x,\succ'_1,\succ'_2),\succ'_1,\succ'_2)$, and (iv) for $x=a$, we have $\#\partial(U(x,\succ_1,\succ_2),\succ_1,\succ_2)=n$ and $\#\partial(U(x,\succ'_1,\succ'_2),\succ'_1,\succ'_2)=n-1$. This implies that $\Delta q(y,S,\succ'_1,\succ'_2|\succ_1,\succ_2)=\frac{1}{n(n-1)\#\eta S}$ for any $y\in\{y_1,...,y_{n-1}\}$, $\Delta q(y,S,\succ'_1,\succ'_2|\succ_1,\succ_2)=-\frac{1}{n\#\eta S}$ for $y=y_n$, and $\Delta q(y,S,\succ'_1,\succ'_2|\succ_1,\succ_2)=0$ for any $y\notin\{y_1,...,y_n\}$. So, Exogenous-RDRC satisfies WSG, IUA and LL.

The proof in Part 1 together with the proof in Part 2 imply that Exogenous-RDRC is the unique choice rule that satisfies WPO, WSG, IUA, and LL.

Proof of Theorem 6:

The sufficient part is obvious. We next show the necessity part. Suppose q^* is a choice rule that satisfies S-SYM and IIE, we will show that $q^*(y,S,\succ_1,\succ_2)=q^L(y,S,\succ_1,\succ_2)$ for any $(S,\succ_1,\succ_2)\in\mathbb{D}_2$ and any $y\in S$.

Pick any $(S,\succ_1,\succ_2)\in\mathbb{D}_2$. Let $N(S,\succ_1,\succ_2)=\{y\in\partial(S,\succ_1,\succ_2)|\#L(y,S,\succ_1,\succ_2)\geq\#L(y',S,\succ_1,\succ_2)$ for any $y'\in S\}$ be the set of alternatives that have the largest number of alternatives in their lower contour sets. Pick any $y^*\in N(S,\succ_1,\succ_2)$. We will define a symmetric decision problem which contains $N(S,\succ_1,\succ_2)$ and all alternatives that are dominated by y^*. Note that we have $\#L(y^*,S,\succ_1,\succ_2)=\#L(y,S,\succ_1,\succ_2)$ for any $y\in N(S,\succ_1,\succ_2)$. Let $S'=L(y^*,S,\succ_1,\succ_2)\cup N(S,\succ_1,\succ_2)$. Let $(\succ'_1,\succ'_2)$ be such that (i) $\partial(S',\succ'_1,\succ'_2)=N(S,\succ_1,\succ_2)$, (ii) $R_i(y|\partial(S',\succ'_1,\succ'_2),\succ'_1,\succ'_2)=R_i(y|N(S,\succ_1,\succ_2),\succ_1,\succ_2)$ for $y\in\partial(S',\succ'_1,\succ'_2)=N(S,\succ_1,\succ_2)$ and $i=1,2$, (iii) $x\in L(y,S,\succ'_1,\succ'_2)$ for any $x\in L(y^*,S,\succ_1,\succ_2)$ and any $y\in\partial(S',\succ'_1,\succ'_2)$, and (iv) $L(y^*,S,\succ_1,\succ_2)$ is symmetric under $(\succ'_1,\succ'_2)$. That is, $(\succ'_1,\succ'_2)$ "transforms" the set $L(y^*,S,\succ_1,\succ_2)$ to a symmetric set of alternatives which are dominated by all alternatives in $N(S,\succ_1,\succ_2)$. Using Lemma 1, $\partial(S',\succ'_1,\succ'_2)$ is symmetric. This together with the fact that $L(y^*,S,\succ_1,\succ_2)$ is symmetric under $(\succ'_1,\succ'_2)$ imply that S' is symmetric under $(\succ'_1,\succ'_2)$. According to S-SYM, we must have that $q^*(y,S',\succ'_1,\succ'_2)=\frac{1}{\#\partial(S',\succ'_1,\succ'_2)}=\frac{1}{\#N(S,\succ_1,\succ_2)}$ for any $y\in\partial(S',\succ'_1,\succ'_2)$.

Now, we compare $(S,\succ_1,\succ_2)$ and $(S',\succ'_1,\succ'_2)$. We have $S'\subseteq S$ and $\partial(S',\succ'_1,\succ'_2)\subseteq\partial(S,\succ_1,\succ_2)$. In addition, for any $y\in\partial(S',\succ'_1,\succ'_2)=N(S,\succ_1,\succ_2)$,[33] we have $\#L(y,S,\succ_1,\succ_2)>\#L(y',S,\succ_1,\succ_2)$ for any $y'\in\partial(S,\succ_1,\succ_2)\backslash\partial(S',\succ'_1,\succ'_2)=\partial(S,\succ_1,\succ_2)\backslash N(S,\succ_1,\succ_2)$ and $\#L(y,S,\succ_1,\succ_2)=\#L(y',S,\succ_1,\succ_2)$ for any $y'\in\partial(S',\succ'_1,\succ'_2)=N(S,\succ_1,\succ_2)$. So, by IIE, we have $q^*(y,S,\succ_1,\succ_2)=q^*(y,S',\succ'_1,\succ'_2)=\frac{1}{\#N(S,\succ_1,\succ_2)}$ for any $y\in\partial(S',\succ'_1,\succ'_2)=N(S,\succ_1,\succ_2)$. This implies that $q^*(y,S,\succ_1,\succ_2)=q^L(y,S,\succ_1,\succ_2)$ for any $y\in S$.

[33]Note that $q^*(y,S',\succ'_1,\succ'_2)>0$ if and only if $y\in\partial(S',\succ'_1,\succ'_2)=N(S,\succ_1,\succ_2)$.

11.2 The $N-$attributes Cases

11.2.1 Axiomatic Characterization of the Exogenous-RDRC's Outcome Set: The N−attributes Case

We now consider the case where each alternative in the feasible set S has $\#\Lambda > 2$ attributes; let $\#\Lambda = N$. Recall that $(S, \succ_1, ..., \succ_N)$ denotes a finite random reference dependent decision problem. Let $\mathbb{D}_N$ denote the set of all finite random reference dependent decision problems where the number of attributes is $N > 2$. The notation $q(y, S, \succ_1, ..., \succ_N)$, $U(x, S, \succ_1, ..., \succ_N)$, $L(x, S, \succ_1, ..., \succ_N)$, $\partial(x, S, \succ_1, ..., \succ_N)$, $\eta(x, S, \succ_1, ..., \succ_N)$, $J(y_1, ..., y_n, S, \succ_1, \succ_2)$, $R_i(y|A, \succ_1, ..., \succ_N)$ and the Exogenous-RDRC's procedure can be defined in a way similar to their counterparts in the 2-attribute case.

The next four axioms are simply the $N > 2$ versions of the ones in the $N = 2$ case.

Weak Pareto Optimality (**WPO**): For any $(S, \succ_1, ..., \succ_N) \in \mathbb{D}_N$, $q(y, S, \succ_1, ..., \succ_N) = 0$ for $\forall y \in \eta(S, \succ_1, ..., \succ_N)$.

Winners' Symmetric Gain (**WSG**): Suppose $(S, \succ_1, ..., \succ_N)$ and $(S, \succ'_1, ..., \succ'_N)$ form a JD setup. Then $\Delta q(y_i, S, \succ'_1, ..., \succ'_N|\succ_1, ..., \succ_N) = \Delta q(y_j, S, \succ'_1, ..., \succ'_N|\succ_1, ..., \succ_N) \geq 0$ for any $y_i, y_j \in \{y_1, ..., y_{n-1}\}$.

Irrelevance of Uninvolved Alternatives (**IUA**): Suppose $(S, \succ_1, ..., \succ_N)$ and $(S, \succ'_1, ..., \succ'_N)$ form a JD setup. Then $\Delta q(y, S, \succ'_1, ..., \succ'_N|\succ_1, ..., \succ_N) = 0$ for any $y \in \partial(S, \succ_1, ..., \succ_N)$ with $y \notin \{y_1, ..., y_n\}$.

Loser's Loss (**LL**): Suppose $(S, \succ_1, ..., \succ_N)$ and $(S, \succ'_1, ..., \succ'_N)$ form a JD setup. Then $\Delta q(y, S, \succ'_1, ..., \succ'_N|\succ_1, ..., \succ_N) = -\frac{1}{n\#\eta S}$ for $y = y_n$.

Theorem 7 *Suppose there are $N > 2$ attributes. The choice rule of the Exogenous-RDRC is the unique one that satisfies WPO, WSG, IUA, and LL.*

Proof of Theorem 7:

The proof of Theorem 7 is similar to that of Theorem 4, except the proof of Lemma 3. In the following, we will present the lemma and prove it for the N-attribute case. The remaining proof of Theorem 7 is omitted.

Lemma 5 *Suppose that $a \in J(y_1, ..., y_l, S, \succ_1, ..., \succ_N)$ for some $\{y_1, ..., y_l\} \subset \partial(S, \succ_1, ..., \succ_N)$ (strict relation), then there exists an attribute ranking $(\succ'_1, ..., \succ'_N)$ such that $a \in J(y_1, ..., y_l, y^*, S, \succ_1, ..., \succ_N)$ for some $y^* \in \partial(S, \succ_1, ..., \succ_N)\backslash\{y_1, ..., y_l\}$ and $R_i(y|S\backslash a, \succ'_1, ..., \succ'_N) = R_i(y|S\backslash a, \succ_1, ..., \succ_N)$ for any $y \in S\backslash a$ and any $i = 1, ..., N$.*

Suppose that $a \in J(y_1, ..., y_l, S, \succ_1, ..., \succ_N)$ for some $\{y_1, ..., y_l\} \subset \partial(S, \succ_1, ..., \succ_N)$. Let $\hat{S} = \{y_1, ..., y_l\}$ and $\hat{S}^c = \partial(S, \succ_1, ..., \succ_N)\backslash\hat{S}$. For any $\{y_1, y_2, ..., y_k\} \subseteq \partial(S, \succ_1, ..., \succ_N)$, let $\overline{R}_i(y_1, ..., y_k|S, \succ_1, ..., \succ_N) = \max\{R_1(y_1|S, \succ_1, ..., \succ_N), R_1(y_2|S, \succ_1, ..., \succ_N), ..., R_1(y_k|S, \succ_1, ..., \succ_N)\}$

be the lowest ranking of $y_1, ..., y_k$ in attribute i (recalling that a larger ranking number means a lower, less-preferred ranking). Pick any alternative in $\hat{S}^c$ and denote it by y^*. We can define a new attribute ranking $(\succ'_1, ..., \succ'_N)$ such that $R_i(y|S\backslash a, \succ'_1, ..., \succ'_N) = R_i(y|S\backslash a, \succ_1, ..., \succ_N)$ for any $y \in S\backslash a$ and any $i \in \{1, ..., N\}$, and $R_i(a|S, \succ'_1, ..., \succ'_N) = \overline{R}_i(y_1, ..., y_l, y^*|S, \succ'_1, ..., \succ'_N) + 1$ for any $i \in \{1, ..., N\}$. That is, $(\succ'_1, ..., \succ'_N)$ is obtained by "moving" a such that for each attribute, the attribute ranking of a is just below the lowest attribute ranking of $\{y_1, y_2, ..., y_l, y^*\}$ where the relative rankings of all alternative other than a are unchanged. Note that a is a dominated alternative under both $(\succ_1, ..., \succ_N)$ and $(\succ'_1, ..., \succ'_N)$, so the above transformation will not change the set of Pareto alternatives. In addition, it is obvious that $a \in L(y|S, \succ'_1, ..., \succ'_N)$ for any $y \in \{y_1, y_2, ..., y_l, y^*\}$. If a is not dominated by any $y \in \hat{S}^c\backslash y^*$, then we have $a \in J(y_1, ..., y_l, y^*, S, \succ_1, ..., \succ_N)$ and the lemma is proved. If not, we proceed with the following process.

Suppose a is dominated by some $y^{**} \in \hat{S}^c\backslash y^*$. Since y^{**} dominates a, we must have that $R_i(y^{**}|S, \succ'_1, ..., \succ'_N) < \overline{R}_i(y_1, ..., y_l, y^*|S, \succ'_1, ..., \succ'_N)$ for any $i \in \{1, ..., N\}$. However, we cannot have $R_i(y^{**}|S, \succ'_1, ..., \succ'_N) < \overline{R}_i(y_1, ..., y_l|S, \succ'_1, ..., \succ'_N)$ for all $i \in \{1, ..., N\}$ because otherwise y^{**} will dominate a under $(S, \succ_1, ..., \succ_N)$, which will be a contradiction with the fact that $a \in J(y_1, ..., y_l, S, \succ_1, ..., \succ_N)$. The above two facts then imply that $\overline{R}_i(y_1, ..., y_l|S, \succ'_1, ..., \succ'_N) < R_i(y^{**}|S, \succ'_1, ..., \succ'_N) < R_i(y^*|S, \succ'_1, ..., \succ'_N)$ for some $i \in \{1, ..., N\}$. Then, we can define a new attributes' ranking, say $(\succ''_1, ..., \succ''_N)$, such that $R_i(y|S\backslash a, \succ''_1, ..., \succ''_N) = R_i(y|S\backslash a, \succ'_1, ..., \succ'_N)$ for any $y \in S\backslash a$ and any $i \in \{1, 2\}$, and $R_i(a|S, \succ''_1, ..., \succ''_N) = \overline{R}_i(y_1, ..., y_l, y^{**}|S, \succ'_1, ..., \succ'_N) + 1$ for any $i \in \{1, 2\}$. Obviously, under $(\succ''_1, ..., \succ''_N)$, a must be dominated by $y_1, ..., y_l$ and y^{**} and possibly by others. However, since $R_i(y^{**}|S, \succ'_1, ..., \succ'_N) < \overline{R}_i(y_1, ..., y_l, y^*|S, \succ'_1, ..., \succ'_N)$ for any $i \in \{1, 2\}$, we must have $R_i(a|S, \succ''_1, ..., \succ''_N) \leq R_i(a|S, \succ'_1, ..., \succ'_N)$ for any $i \in \{1, 2\}$. This implies that fewer alternatives will dominate a under $(\succ''_1, ..., \succ''_N)$ than under $(\succ'_1, ..., \succ'_N)$. That is, $U(a|S, \succ''_1, ..., \succ''_N) \subseteq U(a|S, \succ'_1, ..., \succ'_N)$. In addition, this relation must be strict because y^* dominates a under $(\succ'_1, ..., \succ'_N)$ but not under $(\succ''_1, ..., \succ''_N)$ (because $\overline{R}_i(y_1, ..., y_l|S, \succ'_1, ..., \succ'_N) < R_i(y^{**}|S, \succ'_1, ..., \succ'_N) < R_i(y^*|S, \succ'_1, ..., \succ'_N)$ for some $i \in \{1, 2\}$, as we have shown above). If a is not dominated by any $y \in \hat{S}^c\backslash y^{**}$, then we have $a \in J(y_1, ..., y_l, y^{**}, S, \succ_1, ..., \succ_N)$ and the lemma is proved. If not, we can repeat the above process. Note that each repetition of the above process will strictly shrink the size of the set of alternatives that dominates a, while keeping $y_1, ..., y_l$ and at least one other Pareto alternative in the set. So, eventually, the repetition of the above process will reach a stage where the size of the set of alternatives that dominates a is exactly $n + 1$. That is, we have shown that there exists an attribute ranking and a Pareto alternative in $\hat{S}^c$ such that under the attribute ranking, a is dominated by this alternative and $\{y_1, ..., y_l\}$ but not dominated by any other alternatives (and the relative rankings of alternatives other than a are the same as in the case where the attributes' ranking profile is $(\succ_1, ..., \succ_N)$).

11.2.2 Axiomatic Characterization of the Endogenous-RDRC's Outcome Set: The $N-$attributes Case

The axioms that characterize the Endogenous-RDRC for the 2-attribute case can be easily generalized to the $N-$attributes case. In particular, we define the following axioms.

Weak Pareto Optimality (**WPO**): For any $(S, \succ_1, ..., \succ_N) \in \mathbb{D}_N$, $q(y, S, \succ_1, ..., \succ_N) = 0$ for $\forall y \in \eta(S, \succ_1, ..., \succ_N)$.

Non-losers' Proportional Gain (**NPG**): Suppose $(S, \succ_1, ..., \succ_N)$ and $(S, \succ'_1, ..., \succ'_N)$ form a JD setup. Then for any $y \in \partial(S, \succ_1, ..., \succ_N)\backslash y_n$, $\Delta q(y, S, \succ'_1, ..., \succ'_N | \succ_1, ..., \succ_N) > 0$ if and only if $\#LCS_y > 0$. In addition, $\frac{\Delta q(y, S, \succ'_1, ..., \succ'_N | \succ_1, ..., \succ_N)}{\Delta q(y', S, \succ'_1, ..., \succ'_N | \succ_1, ..., \succ_N)} = \frac{\#LCS_y}{\#LCS_{y'}}$ for any $y, y' \in \partial(S, \succ_1, ..., \succ_N)\backslash y_n$ with $\#LCS_y > 0$ and $\#LCS_{y'} > 0$.

Loser's Loss (**LL 2**): Suppose $(S, \succ_1, ..., \succ_N)$ and $(S, \succ'_1, ..., \succ'_N)$ form a JD setup. Then $\Delta q(y_n, S, \succ'_1, ..., \succ'_N | \succ_1, ..., \succ_N) = -\frac{I(S, \succ_1, ..., \succ_N) - \#LCS_{y_n}}{I(S, \succ_1, ..., \succ_N)(I(S, \succ_1, ..., \succ_N) - 1)}$.

Theorem **8** *The choice rule of the Endogenous-RDRC is the unique one that satisfies WPO, NPG and LL 2.*

11.2.3 Axiomatic Characterization of the RIDC's Outcome Set: The $N-$attributes Case

Suppose each alternative in the feasible set S has $N > 2$ attributes. In order to generalize the RIDC's axiomatization to the $N > 2$ attribute case, we need even a stronger version of the S-SYM axiom, which is defined as follows.

Strong Pareto Symmetry (**SP-SYM**): For any $(S, \succ_1, ..., \succ_N) \in \mathbb{D}_N$, suppose $L(y, S, \succ_1, ..., \succ_N) = L(z, S, \succ_1, ..., \succ_N)$ for any $y, z \in \partial(S, \succ_1, ..., \succ_N)$. Then, $q(y, S, \succ_1, ..., \succ_N) = \frac{1}{\#\partial(S, \succ_1, ..., \succ_N)}$ for $\forall y \in \partial(S, \succ_1, ..., \succ_N)$.

The IIE axiom in the 2-attribute case can be easily extended to the case of $N > 2$ attributes.

Independence of Irrelevant Expansions (**IIE**): For any two $(S, \succ_1, ..., \succ_N), (S', \succ'_1, ..., \succ'_N) \in \mathbb{D}_N$, suppose $S \subseteq S'$ and $\partial(S, \succ_1, ..., \succ_N) \subseteq \partial(S', \succ'_1, ..., \succ'_N)$ where $R_i(y|\partial(S, \succ_1, ..., \succ_N), \succ_1, ..., \succ_N) = R_i(y|\partial(S', \succ'_1, ..., \succ'_N), \succ'_1, ..., \succ'_N)$ for any $y \in \partial(S, \succ_1, ..., \succ_N)$ and any $i \in \{1, ..., N\}$. If for any $y \in S$ such that $q(y, S, \succ_1, ..., \succ_N) > 0$, we have (i) $\#L(y, S', \succ'_1, ..., \succ'_N) > \#L(z, S', \succ'_1, ..., \succ'_N)$ for any $z \in \partial(S, \succ_1, ..., \succ_N)$ with $q(z, S, \succ_1, ..., \succ_N) = 0$ and any $z \in \partial(S', \succ'_1, ..., \succ'_N)\backslash\partial(S, \succ_1, ..., \succ_N)$, and (ii) $\#L(y, S', \succ'_1, ..., \succ'_N) = \#L(z, S', \succ'_1, ..., \succ'_N)$ for any $z \in \partial(S, \succ_1, ..., \succ_N)$ with $q(z, S, \succ_1, ..., \succ_N) > 0$, then $q(y, S', \succ'_1, ..., \succ'_N) = q(y, S, \succ_1, ..., \succ_N)$ for any $y \in S$.

We have the following result (its proof can be obtained by mimicking that of Theorem 6 and is thus omitted).

Theorem **9** *Suppose there are $N > 2$ attributes. The choice rule of RDRC is the unique one that satisfies SP-SYM and IIE.*

11.3 Independence of Axioms

11.3.1 Independence of the Axioms of the Exogenous-RDRC's Outcome Set

This section discusses the independence of axioms in the Exogenous-RDRC solution. We focus on the case of $N = 2$. The analysis in this section can be easily extended to the general case of $N > 2$.

We first consider a choice rule that satisfies WSG, IUA and WPO but violates LL. Let q be a choice rule such that $q(y) = \frac{1}{\#\partial(S,\succ_1,\succ_2)}$ for any $y \in \partial(S,\succ_1,\succ_2)$ and any $(S,\succ_1,\succ_2) \in \mathbb{D}_2$. It can be verified that q satisfies WSG, IUA and WPO but violates LL.

We now consider a choice rule that satisfies WSG, IUA and LL, but violates WPO. We define the following choice rule:

(i) if $(S,\succ_1,\succ_2)$ has one undominated alternative and one dominated alternative, then the choice probability of both the undominated alternative and the dominated alternative are 1/2.

(ii) in all other cases, the choice rule is the same as the choice rule of the Exogenous-RDRC.

It can be easily verified that the above choice rule satisfies WSG, IUA and LL, but violates WPO.

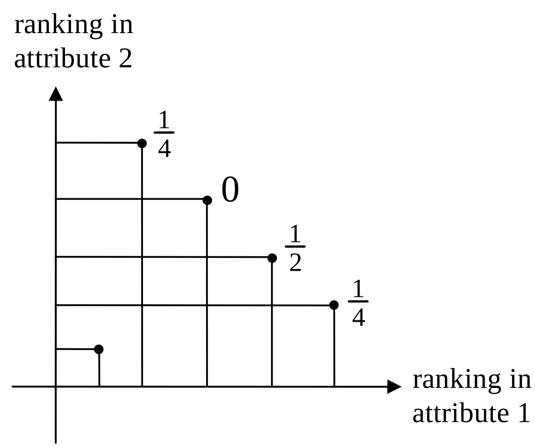

Figure 11

We next define a choice rule that satisfies IUA, LL, WPO, but violates WSG. In particular, consider the following choice rule:

(i) if $(S,\succ_1,\succ_2)$ is as illustrated in Figure 11 (i.e., there are four undominated alternatives and one dominated alternative which is dominated by all undominated alternatives), then the choice probabilities are those prescribed in Figure 11.

(ii) in all other cases, the choice rule is the same as the choice rule of the Exogenous-RDRC.

It can be easily verified that the above choice rule satisfies IUA, LL and WPO, but violates WSG.

Finally, we define a choice rule which satisfies WSG, LL, WPO, but violates IUA. In particular, consider the following choice rule:

(i) if $(S, \succ_1, \succ_2)$ is as illustrated in Figure 12 (i.e., there are three Pareto alternatives and two dominated alternative, where one dominated alternative is exclusively dominated by the Pareto alternative that has the highest ranking in attribute 1 and the other dominated alternative is exclusively dominated by the Pareto alternative that has the highest ranking in attribute 2), then the choice probabilities are those prescribed in Figure 12.

(ii) in all other cases, the choice rule is the same as the choice rule of the Exogenous-RDRC.

It can be verified that the above choice rule satisfies WSG, LL and WPO, but violates IUA.

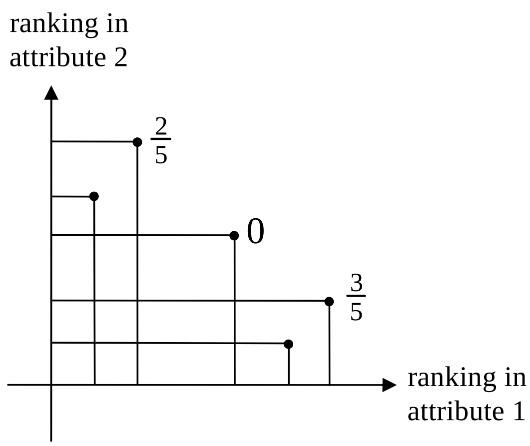

Figure 12

11.3.2 Independence of the Axioms of the Endogenous-RDRC's Outcome Set

For simplicity, we focus on the case of $N = 2$. We will define three choice rules, where the first two are similar to those constructed in the previous subsection.

First, consider the following choice rule:

(i) if $(S, \succ_1, \succ_2)$ has one undominated alternative and one dominated alternative, then the choice probability of both the undominated alternative and the dominated alternative are $1/2$.

(ii) in all other cases, the choice rule is the same as the choice rule of the Endogenous-RDRC.

It can be verified that the above choice rule satisfies NPG and LL 2, but violates WPO.

Second, consider the following choice rule:

(i) if $(S, \succ_1, \succ_2)$ is as illustrated in Figure 12, then the choice probabilities are those prescribed in Figure 12.

(ii) in all other cases, the choice rule is the same as the choice rule of the Endogenous-RDRC.

It can be verified that the above choice rule satisfies LL 2 and WPO, but violates NPG.

Finally, consider the following choice rule:

(i) if $(S, \succ_1, \succ_2)$ is as illustrated in Figure 13, then the choice probabilities are those prescribed in Figure 13.

(ii) in all other cases, the choice rule is the same as the choice rule of the Endogenous-RDRC.

It can be verified that the above choice rule satisfies NPG and WPO, but violates LL 2.

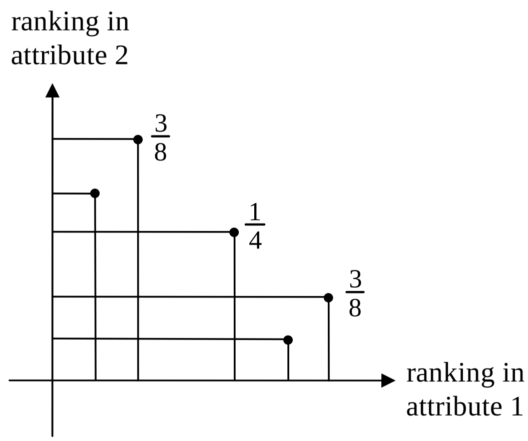

Figure 13

11.3.3 *Independence of the Axioms of the RIDC Outcome's Set*

This section discusses the independence of axioms that characterizes the RIDC's outcome set. For simplicity, we focus on the case where the number of attributes is 2.

First, consider the following choice rule. For any decision problem $(S, \succ_1, \succ_2)$, let $\{y_1, ..., y_n\} \subseteq S$ be the set of alternatives such that for any $y_i \in \{y_1, ..., y_n\}$, we have $\#L(y_i, S) \geq \#L(z, S)$ for any $z \in S$. If $\{y_1, ..., y_n\}$ contains only one alternative, then that alternative is chosen with the outcome with probability one. If $\{y_1, ..., y_n\}$ contains multiple alternatives, then the alternative in $\{y_1, ..., y_n\}$ with highest ranking in Attribute 1 is chosen as the outcome with probability one.

It can be shown that the above choice rule satisfies IIE, but does not satisfy S-SYM.

Next, consider a new choice rule. For any decision problem $(S, \succ_1, \succ_2)$, let $\{w_1, ..., w_n\} \subseteq \partial(S, \succ_1, \succ_2)$ be the set of alternatives such that for any $w_i \in \{w_1, ..., w_n\}$, we have $\#L(w_i, S) \leq \#L(z, S)$ for any $z \in \partial(S, \succ_1, \succ_2)$. If $\{w_1, ..., w_n\}$ contains only one alternative, then that alternative is chosen as the outcome with probability one. If $\{w_1, ..., w_n\}$ contains multiple alternatives, then each alternative in $\{w_1, ..., w_n\}$ will be chosen as the outcome with equal probabilities, and any alternative outside $\{w_1, ..., w_n\}$ will be chosen with probability zero.

It can be verified that the above choice rule satisfies S-SYM, but violates IIE. Also, note that another choice rule that satisfies S-SYM but violates IIE is the choice rule induced by Exogenous-RDRC.

12 References

Agranov, M. and P. Ortoleva (2016), "Stochastic Choice and Preferences for Randomization," forthcoming, *Journal of Political Economy*.

Anbarci, N. (1993), "Noncooperative Foundations of the Area Monotonic Solution," *Quarterly Journal of Economics*, 108, 245-258.

Anbarci, N. (2006), "Finite Alternating-Move Arbitration Schemes and the Equal Area Solution," *Theory and Decision*, 61, 21-50.

Anbarci, N. and K. Rong (2016), "A 'Successive-Elimination' Heuristic and the Compromise Effect," mimeo, Deakin University.

Ariely, D. and T.S. Wallsten (1995), "Seeking Subjective Dominance in Multidimensional Space: An Explanation of the Asymmetric Dominance Effect," *Organizational Behavior and Human Decision Processes*, 63, 223-232.

Aumann, R. (1962), "Utility Theory without the Completeness Axiom," *Econometrica*, 30, 445-462.

Bernoulli, J. (1713), "Ars Conjectandi (The Art of Conjecturing)."

Bewley, T. (1986), "Knightian Uncertainty Theory: Part I," Cowles Foundation Discussion Paper 807.

Benartzi, S., and R.H. Thaler (2001), "Naive Diversification Strategies in Defined Contribution Saving Plans," *American Economic Review*, 91, 79-98.

Binmore, K., L. Stewart, and A. Voorhoeve (2012), "How Much Ambiguity Aversion? Finding Indifferences between Ellsberg's Risky and Ambiguous Bets," *Journal of Risk and Uncertainty*, 45, 215-238.

Camerer, C. (1995), "Individual Decision Making," in *Handbook of Experimental Economics*, 2, ed. A.R.J. Kagel, Princeton University Press, Princeton.

Cerreia-Vioglio, S., D. Dillenberger, P. Ortoleva, and G. Riella (2015), "Deliberately Stochastic" mimeo, Columbia University.

Chernev, A. (2004). "Extremeness Aversion and Attribute Balance Effects in Choice," *Journal of Consumer Research*, 31, 249-263.

Chioveanu, I. and J. Zhou (2013), "Price Competition with Consumer Confusion," *Management Science*, 59, 2450-2469.

de Clippel, G. and K. Eliaz (2012), "Reason-based Choice: A Bargaining Rationale for the Attraction and Compromise Effects," *Theoretical Economics*, 7, 125-162.

Dembski, W.A. and R.J. Marks II (2009), "Bernoulli's Principle of Insufficient Reason and Conservation of Information in Computer Search," *Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics*.

DeMiguel V., L. Garlappi, and R. Uppal (2009), "Optimal versus Naive Diversification: How Inefficient is the 1/N Portfolio Strategy?" *Review of Financial Studies*, 22, 1915-53.

Dhar, R., A. Menon, and B. Maach (2004). "Toward Extending the Compromise Effect to Complex Buying Contexts," *Journal of Marketing Research*, 41, 258-261.

Doyle, J.R., D.J. O'Connor, G.M. Reynolds, and P.A. Bottomley (1999), "The Robustness of the Asymmetrically Dominated Effect: Buying Frames, Phantom Alternatives, and In-Store Purchases." *Psychology and Marketing,* 16, 225-243.

Dwenger, N., D. Kubler, and G. Weizsacker (2013), "Flipping a Coin: Theory and Evidence," mimeo, Humboldt-Universitat zu Berlin.

Fudenberg, D., R. Iijima, and T. Strzalecki (2015), "Stochastic Choice and Revealed Perturbed Utility," *Econometrica*, 83, 2371-2409.

Green, P.E. and V. Srinivasan (1990), "Conjoint Analysis in Marketing: New Developments With Implications for Research and Practice," *Journal of Marketing*, 54, 3-19.

Gul, F., P. Natenzon, and W. Pesendorfer (2014), "Random Choice as Behavioral Optimization," *Econometrica*, 82, 1873-1912.

Hedgcock, W. and A.R. Rao (2009), "Trade-off Aversion as an Explanation for the Attraction Effect: A Functional MagneticResonance Imaging Study," *Journal of Marketing Research*, 46, 1-13.

Herne, K. (1997), "Decoy Alternatives in Policy Choices: Asymmetric Domination and Compromise Effects," *European Journal of Political Economy*, 13, 575-589.

Higgins, E.T. (1997), "Beyond Pleasure and Pain," *American Psychologist*, 52, 1280-1300.

Huber, J., J.W. Payne, and C. Puto (1982), "Adding Asymmetrically Dominated Alternatives: Violations of Regularity and the Similarity Hypothesis," *Journal of Consumer Research*, 9, 90-98.

Kahneman, D. and A. Tversky (1979), "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 47, 263-291.

Kamenica, E. (2008), "Contextual Inference in Markets: On the Informational Content of Product Lines," *American Economic Review*, 98, 2127-2149.

Kivetz, R., O. Netzer, and V. Srinivasan (2004), "Alternative Models for Capturing the Compromise Effect," *Journal of Marketing Research*, 41, 237-257.

Klein, G.A. (1989), "Recognition-Primed Decisions," *Advances in Man-Machine Systems Research*, 5, 47-92.

Klein, N.H. and J.E. Oglethorpe (1987), "Cognitive Reference Points in Consumer Decision Making," in *Advances in Consumer Research*, 14, eds. M. Wallendorf and P. Anderson, Association for Consumer Research, Provo, UT, 183-187 - http://www.acrwebsite.org/search/view-conference-proceedings.aspx?Id=6682.

Loomes, G. and R. Sugden (1982), "Regret Theory: An Alternative Theory of Rational Choice under Uncertainty," *Economic Journal*, 92, 805-824.

Louviere, J.J. (1984), "Hierarchical Information Integration: A New Method for the Design and Analysis of Complex Multiattribute Judgement Problems," *Advances in Consumer Research*, 11, 148-55.

Luce, R.D. (1959), *Individual Choice Behavior: A Theoretical Analysis*, Wiley, NY.

Machina, M.J. (1985), "Stochastic Choice Functions Generated from Deterministic Preferences over Lotteries," *Economic Journal*, 95, 575-594.

Marley, A. (1997),"Probabilistic Choice as a Consequence of Nonlinear (Sub) Optimization," *Journal of Mathematical Psychology*, 41, 382-391.

Manzini, P. and M. Mariotti (2012), "Choice by Lexicographic Semiorders," *Theoretical Economics*, 7, 1-23.

Manzini, Paola and Marco Mariotti (2014), "Stochastic Choice and Consideration Sets," *Econometrica*, 82, 1153-1176.

Masatlioglu, Y., D. Nakajima, and E.Y. Ozbay (2012), "Revealed Attention," *American Economic Review*, 102, 2183-2205.

Mourali, M., U. Bockenholt, and M. Laroche (2007), "Compromise and Attraction Effects under Prevention and Promotion Motivations," *Journal of Consumer Research*, 34, 234-247.

Natenzon, P. (2015), "Random Choice and Learning," mimeo, Washington University at St. Louis.

O'Curry, Y.P. Sue, and R. Pitts (1995), "The Attraction Effect and Political Choice in Two Elections," *Journal of Consumer Psychology*, 4, 85-101.

Ok, E.A., P. Ortoleva, and G. Riella (2014), "Theory of Product Differentiation in the Presence of the Attraction Effect," mimeo, New York University.

Ok, E.A., P. Ortoleva, and G. Riella (2015), "Revealed (P)Reference Theory," *American Economic Review*, 105, 299-321.

Oppewal, H., J.J. Louviere, and J.P.H. Timmermans (1994), "Modeling Hierarchical Conjoint Processes with Integrated Choice Experiments," *Journal of Marketing Research*, 31, 92-105.

Payne, J.W., J.R. Bettman, and E.J. Johnson (1993), *The Adaptive Decision Maker*. Cambridge University Press, Cambridge.

Piccione, M. and R. Spiegler (2012), "Price Competition under Limited Comparability," *Quarterly Journal of Economics*, 127, 97-135.

Rabin, M. (1998), "Psychology and Economics," *Journal of Economic Literature*, 36, 11-46.

Raz, J. (1986), *The Morality of Freedom*. Clarendon Press and Oxford University Press, New York.

Schwartz, J.A., and G.B. Chapman (1999), "Are More Options Always Better? The Attraction Effect in Physicians' Decisions about Medications," *Medical Decision Making*, 19, 315-323.

Schwarzkopf, D.L. (2003), "The Effects of Attraction on Investment Decisions," *Journal of Behavioral Finance*, 96–108.

Sheng, S., A. Parker, and K. Nakamoto (2005), "Understanding the Mechanism and Determinants of Compromise Effects," *Psychology and Marketing*, 22, 591-609.

Shugan, S.M. (1980), "The Cost of Thinking," *Journal of Consumer Research*, 7, 99-111.

Simonson, I. (1989), "Choice Based on Reason: The Case of Attraction and Compromise Effects," *Journal of Consumer Research*, 16, 158-174.

Simonson, I. and A. Tversky (1992), "Choice in Context: Tradeoff Contrast and Extremeness Aversion," *Journal of Marketing Research*, 29, 281-95.

Swait, J.D. and A.A.J. Marley (2013), "Probabilistic Choice (Models) as a Result of Balancing Multiple Goals," *Journal of Mathematical Psychology*, 57, 1-14.

Tetlock, P.E. (2000), "Coping with Trade-offs: Psychological Constraints and Political Implications," in *Elements of Reason: Cognition, Choice, and the Bounds of Rationality,* eds. A. Lupia, M.D. McCubbins and S.L. Popkin, Cambridge University Press, Cambridge, 239-322.

Tversky, A. (1969), "Intransitivity of Preferences," *Psychological Review*, 76, 31-48.

Tversky, A. (1972), "Elimination by Aspects: A Theory of Choice," *Psychological Review*, 79, 281–299.

Vohs, K.D., R.F. Baumeister, B.J. Schmeichel, J.M. Twenge, N.M. Nelson, and D.M. Tice (2008), "Making Choices Impairs Subsequent Self-Control: A Limited Resource Account of Decision Making, Self-Regulation, and Active Initiative," *Journal of Personality and Social Psychology*, 94, 883-98.

Wang, J. N. Novemsky, R. Dhar, and R.F. Baumeister (2010), "Trade-Offs and Depletion in Choice," *Journal of Marketing Research*, 47, 910-919.

Wright, P. and F. Barbour (1977), "Phased Decision Strategies: Sequels to an Initial Screening," in *Multiple Criteria Decision Making*, eds. M.K. Starr and M. Zeleny, North-Holland Publishing Co., Amsterdam, 91-109.