Every choice function is pro-con rationalizable*

Serhat Dogan† Kemal Yildiz ‡

November 17, 2019

Abstract

We present a new choice model. An agent is endowed with two sets of orderings: pro-orderings and con-orderings. For each choice set, if an alternative is the top-ranked by a *pro-ordering* (*con-ordering*), then this is a *pro* (*con*) for choosing that alternative. The alternative with more pros than cons is chosen from each choice set. Each ordering may have a *weight* reflecting its salience. In this case, each alternative is chosen with a probability proportional to the total weight of its pros and cons. We observe that every nuance of the rich human choice behavior can be captured via this structured model. Our technique requires an extension of Ford-Fulkerson Theorem, which may be of independent interest. We analyze a restricted model, in which there are two unobservable orderings relevant for choice. As an application of our results, we show that every choice rule is plurality-rationalizable.

Keywords: Choice function, random choice, attraction effect, additivity, integer programming.

---

*We are grateful to Ariel Rubinstein, Efe Ok, Yusufcan Masatlioglu, Erkut Ozbay, Ed Green, Kfir Eliaz, Mustafa Celebi Pinar, Ran Spiegler, Burak Pac, Paola Manzini, Marco Mariotti, Michael Richter, Battal Dogan, Selman Erol, Faruk Gul, Wolfgang Pesendorfer, Leeat Yariv, David Dillenberger, Pietro Ortoleva, and seminar participants at Bilkent University, Princeton University, University of Maryland, University of Bristol, University of Glasgow, ADA University, New School of Economics, and Higher School of Economics for their helpful comments. We thank Rana Nelson for her linguistic review. Kemal Yildiz gratefully acknowledges the hospitality of New York University, Department of Economics during his visit in 2017-2018 academic year, and the support from the Scientific and Research Council of Turkey (TUBITAK) under grant number 1059B191601712.

†Bilkent University, dserhat@bilkent.edu.tr

‡Bilkent University, kemal.yildiz@bilkent.edu.tr

Contents

1 Introduction

Charles Darwin, the legendary naturalist, wrote "The day of days!" in his journal on November 11, 1838, when his cousin Emma Wedgwood accepted his marriage proposal. However, whether to marry at all had been a hard decision for Darwin. Just a few months prior, Darwin had scribbled a carefully considered list of *pros* –such as "constant companion", "charms of music", "female chit-chat"–and *cons* –such as "may be quarrelling", "fewer conversations with clever people", "no books"– regarding the potential impact of marriage on his life.[1] With this list of pros and cons, Darwin seems to follow a choice procedure ascribed to Benjamin Franklin.[2] Here we present Franklin (1887)'s choice procedure in his own words.

> *To get over this, my Way is, to divide half a Sheet of Paper by a Line into two Columns, writing over the one Pro, and over the other Con. I endeavour to estimate their respective Weights; and where I find two, one on each side, that seem equal, I strike them both out: If I find a Reason pro equal to some two Reasons con, I strike out the three. If I judge some two Reasons con equal to some three Reasons pro, I strike out the five; and thus proceeding I find at length where the Ballance lies. And tho' the Weight of Reasons cannot be taken with the Precision of Algebraic Quantities, yet when each is thus considered separately and comparatively, and the whole lies before me, I think I can judge better, and am less likely to take a rash Step; and in fact I have found great Advantage from this kind of Equation, in what may be called Moral or Prudential Algebra.*

We formulate the pro-con choice model in the deterministic choice setup by extending Franklin's prudential algebra to choice sets that possibly contain more than two alternatives. A *(deterministic) pro-con model* (pcM) is a pair $\langle \succ, \rhd \rangle$ such that $\succ = \{\succ_1, \cdots, \succ_m\}$ is a set of *pro-orderings* and $\rhd = \{\rhd_1, \cdots, \rhd_q\}$ is a set of *con-*

[1] See Glass (1988) for the full list.

[2] In 1772, Joseph Priestley wrote a letter to Benjamin Franklin asking for Franklin's advice on a decision he was trying to make. Franklin wrote back indicating that he could not tell him what to do, but he could tell him how to make his decision, and suggested his *prudential algebra*.

orderings. We require that an ordering can not both serve as a pro- and con-ordering. Since $\succ$ and $\triangleright$ are defined as sets of orderings rather than lists or profiles of orderings, each ordering can be used only once as a pro- or con-ordering.[3] Given an pcM $\langle \succ, \triangleright \rangle$, for each choice set S and alternative x, if x is the $\succ_i$-top-ranked alternative in S for some $\succ_i \in \succ$, then we interpret this as a 'pro' for choosing x from S. On the other hand, if x is the $\triangleright_i$-top-ranked alternative in S for some $\triangleright_i \in \triangleright$, then we interpret this as a 'con' for choosing x from S. Our central new concept is the following: A *choice function* [4] is *pro-con rational(izable)* if there is an pcM $\langle \succ, \triangleright \rangle$ such that for each choice set S, an alternative x is chosen from S if and only if pros for choosing x from S are more than the cons for choosing x from S.[5]

A *random pro-con model* (RpcM) is a triplet $\langle \succ, \triangleright, \lambda \rangle$, where $\succ$ and $\triangleright$ stand for the sets of pro-orderings and con-orderings, as before. The weight function λ assigns to each pro-ordering $\succ_i \in \succ$ and con-ordering $\triangleright_i \in \triangleright$, a value in the $[-1, 1]$ interval, which we interpret as a measure of the salience of each ordering.[6] The total weight of an alternative in a choice set is the total weight of the pro-and con-orderings at which it is top-ranked. To make a choice from each choice set, a pro-con–rational agent considers the alternatives with a positive total weight, and chooses each alternative from this consideration set with a probability proportional to its total weight.

RpcM offers a choice procedure that both carries the act of selecting an ordering to be maximized as in the *random utility model* (RUM), and elimination of the alternatives based on their attributes as in Tversky (1972)'s *elimination by aspects* (EBA).

[3]One concern is the number of orderings in $\succ$ and $\triangleright$. It follows from this requirement that if there are n alternatives in X, then at most $n!$-many orderings are used in a pro-con model.

[4]A *choice function* C singles out an alternative from each *choice set* S, which is a nonempty subset of the grand *alternative set* X.

[5]In extending Franklin's prudential algebra, one can consider a sequential pro-con model in which first the alternatives that fail to have more pros than cons in the given choice set are eliminated, and then the same elimination continues until an alternative is singled out. Our pro-con model is a specific sequential pro-con model in which all the alternatives but the chosen one are eliminated in the first step.

[6] In line with the experimental findings of Shafir (1993) indicating that the weight assigned to the pros is more than the weight assigned to the cons, we require the total weighted sum of pro-orderings and con-orderings be unity.

The most familiar stochastic choice model in economics is the RUM, which assumes that an agent is endowed with a probability measure μ over a set of orderings $\succ$ such that he randomly selects an ordering to be maximized from $\succ$ according to μ. An RUM $\langle \succ, \mu \rangle$ is an RpcM in which there is no set of con-orderings. As an alternative to the RUM, Tversky proposes EBA in which an agent views each alternative as a set of attributes. Then, at each stage, the agent selects an attribute with probability proportional to its weight and eliminates all the alternatives without the selected attribute. To clarify the connection, consider a con-ordering $\triangleright_i$ and an alternative x, suppose "*x has attribute i in choice set S*" unless x is the $\triangleright_i$-top-ranked alternative in S. Thus, for a given RpcM, each alternative without attribute i in choice set S is eliminated with a probability proportional to the weight of attribute i. A pro-con–rational agent's attitude to these attributes is twofold: If it is a pro-ordering, then he seeks maximization as in the RUM, if it is a con-ordering, then he is satisfied by elimination of the worst alternative as in the EBA.

It may of interest to view our model from the perspective of probabilistic social choice. Existing work in this literature show that the class of probabilistic group decision rules have considerable richness and appeal.[7] Along these lines, for a social-choice interpretation of the mixed-sign representation in an RpcM, consider a chair who stochastically aggregates different opinions in a committee to make a choice. It is typically assumed that as more committee members top rank an alternative, the choice probability of this alternative increases. However, there may be an antagonistic relationship between the chair and some committee members, so that the chair would be less likely to choose the alternative favored by them.

To highlight the similarity between the RpcM and the RUM, both models are *additive*, in the sense that the choice probability of an alternative is calculated by

[7] As a partial list one can consider Intriligator (1973), Barberá & Sonnenschein (1978), Pattanaik & Peleg (1986), and Intriligator (1982). These studies typically investigate the structure of coalitional power under probabilistic social decision rules. The closest to our work is Pattanaik & Peleg (1986) who axiomatically characterize the random dictatorship procedure, in which there is a probability measure μ on the members of the society N such that for each profile of individual preferences $\{\succ_i\}_{i \in N}$, the society chooses from each choice set according to the RUM $\langle \{\succ_i\}_{i \in N}, \mu \rangle$.

summing up the weights assigned to the orderings. The primitives of both the RpcM and RUM are *structurally invariant*, in the sense that the decision maker uses the same $\langle \succ, \mu \rangle$ and $\langle \succ, \triangleright, \lambda \rangle$ to make a choice from each choice set. These two features of RUM reflect themselves in its characterization.[8] Despite the similarity between the RpcM and the RUM, in Theorem 1, we show that every random choice function is pro-con rational by using an original extension of Ford Jr & Fulkerson (2015)'s seminal result in optimization theory. Then, by using the construction in Theorem 1's proof together with two key results from the integer-programming literature, we show that each (deterministic) choice function is pro-con rational.[9]

We believe that being inclusive does not take away from the relevance of a structured model, but opens up new directions to pursue. For example, in the classical "attraction effect" scenario it seems that there are only two relevant attributes for choice, such as price and quantity. The pro- and con-orderings used in our Example 1 correspond to these attributes. As a result, the choice probability of an alternative may increase when a *decoy* is added, since this alternative may no longer be the worst one according to a relevant attribute. Motivated by this example, in Section 3, we analyze the *dual pro-con model*, in which there are two unobservable orderings of the alternatives that are relevant for choice.

The remaining observations in the paper are as follows. For the uniqueness of representation, the RpcM has characteristics similar to the RUM, which we discuss in Sections 2.3 and 4.3. In Section 4.4, we observe that our Theorem 3 fails to hold in the context of multi-valued choice rules unless we allow multiple appearance of an ordering as a pro- or con-ordering. In Section 4.5, we illustrate that our results facilitate identification of other inclusive choice models, by showing that each choice

[8] Namely, the RCFs that render a random utility representation are those with nonnegative Block-Marschak polynomials. See Block & Marschak (1960), Falmagne (1978), McFadden (1978), and Barberá & Pattanaik (1986).

[9] This result does not directly follow from Theorem 1, since a pro-con model is not a direct adaptation of the random pro-con model, in that we require each ordering to have a fixed unit weight instead of having fractional weights. To best of our knowledge the use of integer programming techniques in this context is new.

function is *plurality-rationalizable*.[10]

1.1 Related literature

In the deterministic choice literature, previous choice models proposed by Kalai et al. (2002) and Bossert & Sprumont (2013) yield similar "anything goes" results. A choice function is *rationalizable by multiple rationales* (Kalai et al. (2002)) if there is a collection of preference relations such that for each choice set the choice is made by maximizing one of these preferences. Put differently, the decision maker selects an ordering to be maximized for each choice set. A choice function is *backwards-induction rationalizable* (Bossert & Sprumont (2013)) if there is an extensive-form game such that for each choice set the backwards-induction outcome of the restriction of the game to the choice set coincides with the choice. In this model, for each choice set, a new game is obtained by pruning the original tree of all branches leading to unavailable alternatives. In the stochastic choice setup, Manzini & Mariotti (2014) provide an anything-goes result for the *menu-dependent random consideration set rules*. In this model, an agent keeps a single preference relation and attaches to each alternative a choice-set-specific attention parameter. Then, from each choice he chooses an alternative with the probability that no more-preferable alternative grabs his attention. In contrast to these models, we believe that the pro-con model is more structured, and exhibits limited context dependency. In that, an agent following a pro-con model only restricts the pro-orderings and con-orderings to the given choice set to make a choice.

Our Theorem 1 is related to a result in a contemporary paper by Saito (2017), who offers characterizations of the mixed logit model. It follows from the results of this paper, which is proved by using a different approach, that each RCF can be expressed as an affine combination of two random utility functions.[11] We discuss the

[10] It turns out that a pro-con model can also be viewed as a collective decision making model based on plurality voting. The model and the result can thought as a generalization of an earlier model and a related result by McGarvey (1953).

[11] We are grateful to an anonymous referee for bringing this connection to our awareness.

technical differences in Remark 1.

2 Random pro-con choice

2.1 The model

Given a nonempty finite alternative set X, any nonempty subset S is called a **choice set**. Let Ω denote the collection of all choice sets. A **random choice function** (RCF) p is a mapping that assigns each choice set $S \in \Omega$, a probability measure over S. For each $S \in \Omega$ and $x \in S$, we denote by $p(x, S)$ the probability that alternative x is chosen from choice set S. An **ordering**, denoted generically by $\succ_i$ or $\rhd_i$, is a complete, transitive, and antisymmetric binary relation on X.

A **random pro-con model** (RpcM) is a triplet $\langle \succ, \rhd, \lambda \rangle$, where $\succ = \{\succ_1, \cdots, \succ_m\}$ and $\rhd = \{\rhd_1, \cdots, \rhd_q\}$ are sets of pro- and con-orderings on X. We require that if an ordering appears as a pro-ordering, then it can not appear as a con-ordering. Finally, the **weight function**, denoted by λ is such that for each $\succ_i \in \succ$ and $\rhd_i \in \rhd$, we have $\lambda(\succ_i) \in (0, 1]$, $\lambda(\rhd_i) \in [-1, 0)$, and the weighted sum of pro-orderings and con-orderings is one, i.e. $\sum_{\{\succ_i \in \succ\}} \lambda(\succ_i) + \sum_{\{\rhd_i \in \rhd\}} \lambda(\rhd_i) = 1$. The weight function λ acts like a probability measure over the set of orderings that can assign negative values. In measure theoretic language, the primitive of a random pro-con model is a *signed probability measure* defined over the set of orderings.

Given an RpcM $\langle \succ, \rhd, \lambda \rangle$, for each choice set S and alternative $x \in S$, if x is the $\succ_i$-top-ranked alternative in S for some $\succ_i \in \succ$, then we interpret this as a 'pro' for choosing x from S. On the other hand, if x is the $\rhd_i$-top-ranked alternative in S for some $\rhd_i \in \rhd$, then we interpret this as a 'con' for choosing x from S. We interpret the weight assigned to each pro-ordering or con-ordering as a measure of the strength of that ordering.

To define when an RCF is pro-con rational, let $Pros(x, S) = \{\succ_i \in \succ : \ x = max(S, \succ_i)\}$ and $Cons(x, S) = \{\rhd_i \in \rhd : \ x = max(S, \rhd_i)\}$. Next, we formally define when an RCF is pro-con rational. For a given RpcM $\langle \succ, \rhd, \lambda \rangle$, for each choice set

$S \in \Omega$ and $x \in S$, we denote the total weight of x in S by $\lambda(x, S)$, i.e. $\lambda(x, S) = \lambda(Pros(x, S)) + \lambda(Cons(x, S))$. For each choice set $S \in \Omega$, let S^+ be the set of alternatives in S that receives a positive total weight, i.e. $S^+ = \{x \in S \,:\, \lambda(x, S) > 0\}$.

Definition 1 *An RCF p is **pro-con rational** if there is an RpcM $\langle \succ, \triangleright, \lambda \rangle$ such that for each choice set $S \in \Omega$ and $x \in S$,*

$$p(x, S) = max\left\{ \; 0 \; , \; \frac{\lambda(x, S)}{\sum_{\{y \in S^+\}} \lambda(y, S)} \; \right\} \tag{1}$$

That is, to make a choice from each choice set S, a pro-con–rational agent considers the alternatives with a positive total weight, and chooses each alternative from this consideration set with a probability proportional to its total weight.[12]

2.2 Main result

In our main result, we show that every random choice function is pro-con rational. We present a detailed discussion of the result in the introduction. We present the proof in Section 6. As a notable technical contribution, we extend and use Ford-Fulkerson Theorem (Ford Jr & Fulkerson (2015)) from combinatorial matrix theory.[13] Next, we state the theorem and present an overview of the proof. Then, we discuss the technical connection to Saito (2017).

Theorem 1 *Every random choice function is pro-con rational.*

An overview of the proof: For a given RCF p, we show that there is a *signed weight function* λ, which assigns each ordering $\succ_i$, a value $\lambda(\succ_i) \in [-1, 1]$ such that λ *represents* p. That is, for each choice set S and $x \in S$, $p(x, S)$ is the sum of the weights over orderings at which x is the top-ranked alternative. We prove this by induction. To clarify the induction argument, for $k = 1$, let $\Omega_1 = \{X\}$ and let $\mathcal{P}^1$

[12] An equivalent formulation is as follows. An RCF p is pro-con rational if there is an RpcM $\langle \succ, \triangleright, \lambda \rangle$ such that for each choice set $S \in \Omega$ and $x \in S$, $p(x, S) = \lambda(Pros(x, S)) + \lambda(Cons(x, S))$, where $\lambda(Pros(x, S))$ and $\lambda(Cons(x, S))$ are the sum of the weights over $Pros(x, S)$ and $Cons(x, S)$. Proof of Theorem 1 clarifies this equivalency.

[13] It is also known as *max-flow min-cut theorem* in optimization theory.

consists of n-many equivalence classes such that each class contains all the orderings that top rank the same alternative, irrespective of whether these are chosen with positive probability. That is, for $X = \{x_1, \ldots, x_n\}$, we have $\mathcal{P}^1 = \{[\succ^{x_1}], \cdots, [\succ^{x_n}]\}$, where for each $i \in \{1, \ldots, n\}$ and ordering $\succ_i \in [\succ^{x_i}]$, $x_i = \max(X, \succ_i)$. Now for each $x_i \in X$, define $\lambda^1([\succ^{x_i}]) = p(x_i, X)$. It directly follows that λ^1 is a signed weight function over $\mathcal{P}^1$ that represents the restriction of the given RCF to Ω_1, denoted by p_1. By proceeding inductively, it remains to show that we can construct λ^{k+1} over $\mathcal{P}^{k+1}$ that represents p_{k+1}. In Step 1 of the proof we show that finding such a λ^{k+1} boils down to finding a solution to the system of equalities described by *row sums* (RS) and *column sums* (CS). Up to this point the proof structure is similar to the one followed by Falmagne (1978) and Barberá & Pattanaik (1986) for the characterization of RUM.

To understand (RS), while moving from the k^{th}-step to the $(k+1)^{th}$-step, each $[\succ^k]$ is decomposed into a collection $\{[\succ_j^{k+1}]\}_{j \in J}$ such that for each $[\succ_j^{k+1}]$ there exists an alternative x_j that is not linearly ordered by $[\succ^k]$, but placed at $[\succ_j^{k+1}]$ right on top of the alternatives that are not linearly ordered by $[\succ^k]$. Therefore, the sum of the weights assigned to $\{[\succ_j^{k+1}]\}_{j \in J}$ should be equal to the weight assigned to $[\succ^k]$. This gives us the set of equalities formulated in (RS). To understand (CS), let S be the set of alternatives that are not linearly ordered by $[\succ^k]$. Now, we should design λ^{k+1} such that for each $x_j \in S$, $p(x_j, S)$ should be equal to the sum of the weights assigned to orderings at which x_j is the top-ranked alternative in S. The set of equalities formulated in (CS) guarantees this. This follows from our Lemma 6, which we obtain by using the *Mobius inversion*.[14]

Our proof is based on two interwoven observations. To understand the first, let us turn back to the induction argument. It is easy to see that the signed weight function λ^2 over $\mathcal{P}^2$ that represents p_2 is determined uniquely. That is, there is a unique λ^2 that satisfies equalities (RS) and (CS) formed for $k = 2$. But, then for λ^3 (in general for each $k \geq 3$) to be defined over $\mathcal{P}^3$, the solution to the associated (RS) and (CS) for $k = 3$ is no longer unique. The difficulty is that although any λ^3 that satisfies equalities (RS) and (CS) for the $k = 3$ represents p_3, depending on the

[14] Fiorini (2004) is the first who makes the same observation.

choice of λ^3, the (RS) and (CS) formed for a future step, $k > 3$, may not have a solution. Therefore, to conclude the induction successfully, for each $k \geq 3$, we should be "forwarding looking" in choosing λ^k.

Our second critical observation is that finding a solution to the system described by (RS) and (CS) can be translated to the following basic problem: Let $R = [r_1, \ldots, r_m]$ and $C = [c_1, \ldots, c_n]$ be two real-valued vectors such that the sum of R equals to the sum of C. Now, for which R and C can we find an $m \times n$ matrix $A = [a_{ij}]$ such that A has row sum vector R and column sum vector C, with each entry $a_{ij} \in [-1, 1]$? Ford Jr & Fulkerson (2015) provide a full answer to this question when R and C are positive real valued.[15] However, there are two issues peculiar to our problem. First issue is that the row and column sums can be negative real valued. Indeed, we get nonnegative-valued rows and columns only if the Block-Marschak polynomials are nonnegative, that is, the given p is an RUM. Second issue is that, related to our previous observation, we need "forward looking" solutions.

In our Lemma 5, we provide an extension of Ford-Fulkerson Theorem that paves the way for our proof by solving the two issues. To get an intuition for Lemma 5, it is easy to see that the sum of the absolute values of the rows and columns should be bounded in order to extend the result to real-valued vectors. So, in Lemma 5, we require this sum be less than or equal to $2m$, where m is the number of rows. The choice of this specific bound has two implications. First, we can extend Ford-Fulkerson Theorem with real-valued rows and columns. This solves the first issue. Second, we guarantee that there is a solution that satisfy the bound formulated in item (iii) of Lemma 5. This solution turns out be the forwarding looking solution, which solves the second issue.

Rest of the proof is as follows. In Step 2, we show that (RS) equals (CS). In Step 3, by using a structural result presented in Lemma 7, we show that the row and column vectors associated with (RS) and (CS) satisfy the premises of our Lemma 5. This completes the construction of the desired signed weight function.

Remark 1 As discussed in Section 1.1, Saito (2017) independently shows that each

[15] Brualdi & Ryser (1991) provides a detailed account of similar results.

RCF can be expressed as an affine combination of two random utility functions by using different techniques. In our Theorem 1 we require the weights used in this affine combination be chosen from $[-1, 1]$ interval. This requirement is critical for obtaining Theorem 3. In that, for a deterministic pro-con model, we require pro- and con-orderings be sets of orderings rather than lists of orderings in which the same ordering can appear multiple times. To see the technical implication of this requirement for our proof, note that by following the construction in our proof and directly applying the Ford-Fulkerson Theorem, each RCF can be expressed as an affine combination of random utility functions. To show that these weights can be chosen from $[-1, 1]$ interval, we extend the Ford-Fulkerson Theorem (see Lemma 5) and follow a deliberate induction argument supported by other structural results, such as Lemma 7. We believe that this technique can be fruitful in solving similar random choice problems.

2.3 Uniqueness

An RCF may have different random utility representations even with disjoint sets of orderings. Falmagne (1978) argues that random utility representation is essentially unique. That is, the sum of the probabilities assigned to the orderings at which an alternative x is the k^{th}-top-ranked in a choice set is the same for all random utility representations of the given RCF. Similarly, the primitives of an RpcM are structurally invariant in the sense that the agent uses the same triplet $\langle \succ, \triangleright, \lambda \rangle$ to make a choice from each choice set. As an instance of this similarity, both models render a unique representation when there are only three alternatives.[16] As for the general case, our Proposition 1 provides a uniqueness result for the RpcM, which can be thought as the counterpart of Falmagne's result for the RUM.

For a given RpcM $\langle \succ, \triangleright, \lambda \rangle$, let for each $S \in \Omega$ and $x \in S$, $\lambda(x = B_k | S, \succ, \triangleright)$ be the sum of the weights assigned to the pro- and con-orderings at which x is the k^{th}-top-ranked alternative in S. In our next result, we show that for each RCF the sum of the

[16] This directly follows from the construction used to establish the base of induction in Theorem 1's proof.

weights assigned to the orderings at which x is the k^{th}-top-ranked alternative in S is the same for each pro-con representation of the given RCF. That is, $\lambda(x = B_k | S, \succ, \triangleright)$ is fixed for each RpcM $\langle \succ, \triangleright, \lambda \rangle$ that represents the given RCF.

Proposition 1 *If $\langle \succ, \triangleright, \lambda \rangle$ and $\langle \succ', \triangleright', \lambda' \rangle$ are random pro-con representations of the same RCF p, then for each $S \in \Omega$ and $x \in S$,*

$$\lambda(x = B_k | S, \succ, \triangleright) = \lambda'(x = B_k | S, \succ', \triangleright'). \tag{2}$$

Proof. Let $\langle \succ, \triangleright, \lambda \rangle$ and $\langle \succ', \triangleright', \lambda' \rangle$ be two RPMs that represent the same RCF p. Now, for each choice set $S \in \Omega$, both λ and λ' should satisfy the identity (CS) used in Step 1 of the proof of Theorem 1. That is, for each $S \in \Omega$ and $x \in S$ both λ and λ' generates the same $q(x, S)$ value. Therefore, if we can show that $\lambda(x = B_k | S, \succ, \triangleright)$ can be expressed in terms of $q(x, \cdot)$, then (2) follows. To see this, let $\langle \succ, \triangleright, \lambda \rangle$ be any RpcM that represents p. Next, for each $S \in \Omega$, $x \in S$, and $k \in \{1, \ldots, |S|\}$, consider a partition (S_1, S_2) of S such that $x \in S_2$ and $|S_1| = k - 1$. Let $\mathbb{P}(S, x, k)$ be the collection of all these partitions. Now, for each fixed $(S_1, S_2) \in \mathbb{P}(S, x, k)$, let $\lambda(x | S_1, S_2, \succ, \triangleright)$ be the sum of the weights of the orderings at which x is the top-ranked alternative in S_2 and the top-ranked alternative in S_1. Note that for each such ordering, x is the k^{th}-top-ranked alternative in S. Now, it follows that we have:

$$\lambda(x = B_k | S, \succ, \triangleright) = \sum_{\{(S_1, S_2) \in \mathbb{P}(S, x, k)\}} \lambda(x | S_1, S_2, \succ, \triangleright). \tag{3}$$

Since for each $T \in \Omega$ such that $S_2 \subset T$ and $T \subset X \setminus S_1$, by definition, $q(x, T)$ gives the total weight of the orderings at which x is the top-ranked alternative in S, it follows that

$$\sum_{\mathbb{P}(S,x,k)} \lambda(x | S_1, S_2, \succ, \triangleright) = \sum_{\mathbb{P}(S,x,k)} \sum_{S_2 \subset T \subset X \setminus S_1} q(x, T). \tag{4}$$

Finally, if we substitute (3) in (4), then we express $\lambda(x = B_k | S, \succ, \triangleright)$ only in terms of $q(x, \cdot)$, as desired. ■

3 Dual random pro-con choice

One natural concern is the number of orderings in $\succ$ and $\triangleright$. As noted before, if an ordering is used on the pro-side of a pro-con model and its inverse is used in the con-side of the same model, then we can remove both from the pro-con representation. It follows that the total number of orderings that can be used in a pro-con model is at most $n!$, where n is the number of alternatives in X. In this section, we focus on a particular choice problem in which there are only two unobserved orderings $(\succ_1, \succ_2)$ that are relevant for choice, such as price and quality. In this vein, our next example, which presents an *attraction effect* scenario, provides further motivation for analyzing such a model. First, we present this example illustrating that when we introduce an asymmetrically dominated alternative, the choice probability of the dominating alternative may go up. This choice behavior, known as the *attraction effect*, is incompatible with any RUM.[17]

Example 1 (Attraction Effect) Suppose $X = \{x, y, z\}$, where x and y are two competing alternatives such that none clearly dominates the other, and z is another alternative that is dominated by x but not y. Consider the following RpcM $\langle \succ, \triangleright, \lambda \rangle$, in which there is single pair of orderings used both as the pro- and con-orderings. We can interpret this ordering pair as two distinct criteria that order the alternatives.

(1)	(1)	$(-\frac{1}{2})$	$(-\frac{1}{2})$
$\succ_1$	$\succ_2$	$\succ_1^{-1}$	$\succ_2^{-1}$
x	y	y	z
z	x	z	x
y	z	x	y

Now, since for both criteria x is better than z, we get $p(x, \{x, z\}) = 1$. Since x and y fail to dominate each other, and y fail to dominate z, we get $p(y, \{x, y\}) = p(y, \{y, z\}) =$

[17] Experimental evidence for the attraction effect is first presented by Payne & Puto (1982) and Huber & Puto (1983). Following their work, evidence for the attraction effect has been observed in a wide variety of settings. For a list of these results, consult Rieskamp et al. (2006).

$1/2$. That is, z is a 'decoy' for x when y is available. Note that when only x and y are available, since x is the $\succ_2$-worst alternative, x is eliminated with a weight of $1/2$. However, when the decoy z is added to the choice set, then x is no longer the $\succ_2$-worst alternative, and we get $p(x, \{x, y, z\}) = 2/3$. That is, availability of decoy z increases the choice probability of x.

In our analysis, we provide a set of choice axioms, which guarantee that the observed choices can be generated via an RpcM in which the pro-orderings and the con-orderings are obtained from the same ordering pair, and the weight function satisfies two regularity conditions. A notable aspect of this characterization is that several axioms, such as *attraction* and *decoy invariance*, discipline the choice behavior when choice sets present attraction effect scenarios. This indicates the tight connection between the model and the attraction effect phenomena.[18]

Formally, an RCF p is **dual pro-con rational (dPC-rational)**, if there is an RpcM $\langle \succ, \rhd, \lambda \rangle$ that represents p, where $\succ = (\succ_1, \succ_2)$, $\rhd = (\succ_1^{-1}, \succ_2^{-1})$, and the respective weights $\lambda_1, \lambda_2, \delta_1, \delta_2$ have properties 1 and 2 below. That is, both the pro-orderings and the con-orderings are obtained from $\succ_1$ and $\succ_2$. We assume that the weight function has the following two structural properties: (P1) $\lambda_1 > \delta_2$ and $\lambda_2 > \delta_1$, and (P2) if $\lambda_1 - \delta_2 = \lambda_2 - \delta_1$, then $\lambda_1 = \lambda_2$. P1 tells that for each choice set, if two distinct alternatives are top ranked by $(\succ_1, \succ_2)$, then both of them must be chosen with positive probability. As for P2, suppose $\lambda_1 - \delta_2 = \lambda_2 - \delta_1$, which means that in every binary comparison each alternative is chosen evenly. Now, for each choice set, if two distinct alternatives are top ranked by $(\succ_1, \succ_2)$ and neither of them is bottom ranked by $(\succ_1, \succ_2)$, then P2 requires each of the alternatives be chosen evenly. We believe that P1 and P2 are mild technical requirements, yet an axiomatic characterization

[18]The dual pro-con model is related to the literature on dual-self models. Among these, De Clippel & Eliaz (2012) propose a deterministic choice model in which an agent seeks to reach a compromise between two inner selves that represent two attributes of the available alternatives. In contrast, they assume that two orderings that represent the two inner selves are observable. They characterize a model of reason-based choice obtained as a result of a cooperative solution to the bargaining problem between the two selves, which accounts for both the deterministic formulations of the attraction and the comprimise effects.

without these two properties seems rather long and troublesome.

In a recent and closely related paper, Manzini & Mariotti (2018) characterize the random utility model with two orderings. A *dual random utility model* (dRUM) is an RUM that uses only two orderings such that the probabilities of the orderings may depend on the choice set to which they are applied. That is, an RCF p is an dRUM if there exists a triple $(\succ_1, \succ_2, \alpha)$ such that $\succ_1$ and $\succ_2$ are orderings on X, and $\alpha : \Omega \rightarrow (0, 1)$ is a function that assigns probabilities $\alpha(S)$ to $\succ_1$ and $1 - \alpha(S)$ to $\succ_2$. For each choice set S and $x \in S$, $p(x, S)$ is the sum of the probabilities of $\succ_i$ such that $x = max(S, \succ_i)$ for $i \in \{1, 2\}$. Since we assume that P1 holds, dPC-rational model is a specific dRUM that puts additional structure on the set contingent probability function α. In particular, given $\lambda_1, \lambda_2, \delta_1, \delta_2$, for each choice set S, $\alpha(S)$ can take one of the following four values $\{\frac{\lambda_1-\delta_2}{1-\delta_1-\delta_2}, \frac{\lambda_1-\delta_2}{1-\delta_2}, \frac{\lambda_1}{1-\delta_1}, \frac{\lambda_1}{\lambda_1+\lambda_2}\}$. Moreover, each value is associated with choice sets that have specific structures.

Manzini & Mariotti (2018) show that an RCF is an dRUM if and only if it satisfies *modal monotonicity, modal contraction consistency* and *modal impact consistency*. Since a dPC-rational RCF is a specific dRUM, it satisfies these there axioms, which we refer to as the *modal axioms*. Next, we provide the definitions of modal axioms. One can consult Manzini & Mariotti (2018) for the interpretation of these three axioms.[19]

An alternative y **modally impacts** another alternative x in a choice set S if $p(x, S) = 1$ and $p(x, S \cup \{y\}) \in (0, 1)$ or $p(x, S) > 0$ and $p(x, S \cup \{y\}) = 0$. That is, y modally impacts x if adding y transforms the choice of x from possible (including certain) to impossible or from certain to merely possible.

Modal monotonicity: Let $x \in T \subset S$. (i) If $p(x, T) = 0$, then $p(x, S) = 0$. (ii) If $p(x, S) = 1$, then $p(x, T) = 1$.

Modal contraction consistency: If y modally impacts x in S and $T \subset S$, then y

[19] Eliaz et al. (2011) provide an axiomatic characterization of the *top-and-top choice rule* that chooses the $(\succ_1, \succ_2)$-best alternatives for a pair of unobserved orderings $(\succ_1, \succ_2)$ obtained from an agents' deterministic choices. We obtain another axiomatic characterization of the dPC-rational model by replacing the *modal axioms* with the axioms of Eliaz et al. (2011). The only caveat is to impose these axioms on p^+, which we define later in this section.

modally impacts x in T.

Modal impact consistency: If there is no $x \in S$ such that y modally impacts x in S, then $p(y, S \cup \{y\}) = 0$.

Next, we introduce our new axioms used in characterizing the dPC-rational model. But, first, we introduce few basic definitions used in formulating our axioms. For each $x, y \in X$, x **dominates** y, denoted by $x \ D \ y$, if $p(x, \{x, y\}) = 1$. If neither x dominates y nor y dominates x, then x is **independent** of y, denoted by $x \perp y$. For each $S \in \Omega$ if x is independent of every $y \in S$, then x is independent of S, denoted by $x \perp S$. As in an attraction effect scenario, for each $x, y, z \in X$, if neither y dominates x or z, nor x or z dominates y, but x dominates z, then z **is a decoy for** x **when** y **is available.** Our first axiom, *attraction*, requires that adding a decoy z for x when y is available should not decrease the choice probability of the target alternative x.

Axiom 1: (Attraction) For each $S \in \Omega$, $x, y \in S$, and $z \in X$, if z is a decoy for x when y is available, then $p(x, S \cup \{z\}) \geq p(x, S)$.

To introduce our second axiom, consider three distinct alternatives x, y, z such that x is both independent of y and z. Now, in the vein of WARP, if x is chosen with positive probability when compared to y and z together, then *binary WARP* requires the choice probability of x be the same when compared to y and z separately.

Axiom 2: (Binary WARP) For each distinct $x, y, z \in X$ such that $x \perp y$ and $x \perp z$, if $p(x, \{x, y, z\}) > 0$, then $p(x, \{x, y\}) = p(x, \{x, z\})$.

As a counterpart of the *revealed preference* in deterministic choice setup, we define for each $x, y \in X$, x **is stochastically preferred to** y, denoted by $x \ R \ y$ if $p(x, \{x, y\}) \in [1/2, 1)$. Our third axiom requires that in every binary comparison if no alternative dominates the other, then the alternative that is stochastically preferred to the other is chosen with a fixed probability. Put differently, there exists $\alpha \in (0, 1)$ such that in every binary comparison, choice probabilities take one of the values in $\{0, \alpha, 1 - \alpha, 1\}$.

Axiom 3: (Binary invariance) For each $x, x', y, y' \in X$, if $x \ R \ y$ and $x' \ R \ y'$, then $p(x, \{x, y\}) = p(x', \{x', y'\})$.

For the next axiom, consider two choice sets with three alternatives. If in both choice sets there is a decoy for the stochastically preferred alternative, then *decoy invariance* requires the choice probabilities of these alternatives be the same.

Axiom 4: (Decoy invariance) For each $x, y, z \in X$ such that z is a decoy for x when y is available and $x', y', z' \in X$ such that z' is a decoy for x' when y' is available, if $x \; R \; y$ and $x' \; R \; y'$, then $p(x, \{x, y, z\}) = p(x', \{x', y', z'\})$.

For the next axiom, again consider two choice sets with three alternatives. If both choice sets contain an alternative that is dominated by the rest, then *dominated alternative invariance* requires the choice probabilities of the stochastically preferred alternatives be the same.

Axiom 5: (Dominated alternative invariance) For each $x, y, z \in X$ such that both x and y dominate z and $x', y', z' \in X$ such that both x' and y' dominate z', if $x \; R \; y$ and $x' \; R \; y'$, then $p(x, \{x, y, z\}) = p(x', \{x', y', z'\})$.

To introduce the next axiom, we first define the **choice likelihood of** x **from** S as the ratio of the probability that alternative x is chosen from choice set S to the probability that any other alternative is chosen from S, i.e. $L(x, S) = \frac{p(x,S)}{1-p(x,S)}$. Now, to get an intuition for our next axiom, consider the choice sets $\{x, y\}$ and $\{x, y, w\}$ such that in the first one x is compared to y alone, whereas in the second there is additional alternative w that is a decoy for y when x is available. Suppose to both of these choice sets we add an alternative z, which is a decoy for x when y is available. *Attraction* axiom requires the choice probability of x does not decrease in both choice sets, i.e. $\frac{L(x,\{x,y,z\})}{L(x,\{x,y\})} \geq 1$ and $\frac{L(x,\{x,y,z,w\})}{L(x,\{x,y,w\})} \geq 1$. *Likelihood ratio invariance* quantifies *attraction* axiom by requiring the likelihood ratios be the same under both scenarios.

Axiom 6: (Likelihood ratio invariance) For each $x, y, z, w \in X$, if z is a decoy for x when y is available and w is a decoy for y when x is available, then

$$\frac{L(x, \{x, y, z\})}{L(x, \{x, y\})} = \frac{L(x, \{x, y, z, w\})}{L(x, \{x, y, w\})}.$$

Our last axiom, *expansion*, requires that if two choice sets S and S' are similar to each other in the sense that the same alternatives are chosen from both choice sets such that the independence of these alternatives from the choice sets coincide, then the choice probability of each alternative should be the same. Formally, for each $S \in \Omega$ the set of alternatives that are chosen with positive probability in S, namely support of p in S, is denoted by $p^+(S)$. A choice set S is **similar** to another choice set S', denoted by $S \sim S'$, if

i. $p^+(S) = p^+(S')$, and

ii. for each $x \in p^+(S)$, $x \perp S$ if and only if $x \perp S'$.

Axiom 7: (Expansion) For each $S, S' \in \Omega$, if $S \sim S'$, then for each $x \in p^+(S)$, $p(x, S) = p(x, S')$.

Next, we state our characterization result. We present the proof in Section 7. In proving this result, as a stepping stone, we show that an RCF that satisfies these axioms satisfies another condition that we call *strong expansion*. As it will be clear from the proof, an alternative characterization follows in which *strong expansion* replaces Axioms 3,4,5, and 7.

Theorem 2 *An RCF p is dPC-rational if and only if p satisfies the modal axioms and Axioms 1–7.*

Axioms1-7 provide an account of the additional structure that dual pro-con model puts on the set contingent probabilities in a dual random utility model. On the other hand, it is clear that every RCF can be recovered via a random utility model with set contingent probabilities, when there is no restriction on the set of orderings. Thus, it follows from our results that a pro-con model is fairly restrictive than a random utility model with set contingent probabilities when the set of orderings is restricted. However, both are inclusive when there is no such restriction.

4 Deterministic pro-con choice

4.1 The model

A (deterministic) choice function C is a mapping that assigns each choice set $S \in \Omega$ a member of S, that is $C : \Omega \to X$ such that $C(S) \in S$. Let

$\succ$ and $\rhd$ stand for two sets of orderings on X as before. A **(deterministic) pro-con model (pcM)** is a pair $\langle \succ, \rhd \rangle$ consisting of the pro-orderings and the con-orderings. As before, define $Pros(x,S) = \{\succ_i \in \succ : x = max(S, \succ_i)\}$ and $Cons(x,S) = \{\rhd_i \in \rhd : x = max(S, \rhd_i)\}$.

Definition 2 *A choice function C is* ***pro-con rational*** *if there is an pcM $\langle \succ, \rhd \rangle$ such that for each choice set $S \in \Omega$ and $x \in S$, $C(S) = x$ if and only if $|Pros(x,S)| > |Cons(x,S)|$.*

Note that our pro-con model is not a direct adaptation of its random counterpart. In that, we require each ordering to have a fixed unit weight, instead of having fractional weights. Moreover, if an agent is pro-con rational, then at each choice set S there should be a single alternative x such that the number of $Pros(x,S)$ is greater than the number of $Cons(x,S)$. Next, to illustrate how the model works, we revisit Luce and Raiffa's dinner example (Luce & Raiffa (1957)) by following a pro-con model.

Example 2 Suppose you choose chicken when the menu consists of steak and chicken only, yet go for the steak when the menu consists of steak (S), chicken (C), and fish (F). Consider the pro-orderings $\succ_1$ and $\succ_2$ that order the three dishes according to their *attractiveness* and *healthiness*, so suppose $S \succ_1 F \succ_1 C$ and $C \succ_2 F \succ_2 S$. As a con-ordering, consider $F \rhd S \rhd C$, which orders the dishes according to their *riskiness*. Since cooking fish requires expertise, it is the most risky one, and since chicken is the safest option. Now, to make a choice from the grand menu, the pros are: "S is the most attractive", "F is the most healthy", but also "F is the most risky". Thus, S is chosen from the grand menu. If only S and C are available, then we have "C is the most healthy", "S is the most attractive", but also "S is the most risky", so C is chosen.

In our Definition 2, although we do not restrict the structure of the pro- and con-orderings, we ask for a rather structured representation that corresponds to one-to-one elimination in Franklin's prudential algebra. We see at least two benefits of the stringency. First, we obtain the uniqueness property presented in Section 4.3. Second, in Section 4.4, we argue that our Theorem 3 fails to hold in the context of multi-valued choice rules. Finally, given our Theorem 3, one can use our representation to identify other inclusive choice models, which otherwise may not be an easy exercise. In Section 4.5, we present an application along these lines, in which we show that each choice function is plurality-rationalizable.

4.2 Main result

By using the construction in the proof of Theorem 1 and two well-known results from integer-programming literature, we show that every choice function is pro-con rational. This result does not directly follow from Theorem 1, since the pro-con model is not a direct adaptation of its random counterpart. Next, we present the result and its proof.

Theorem 3 *Every choice function is pro-con rational.*

Proof. We prove this result by following the construction used to prove Theorem 1. So, we proceed by induction. Note that since C is a deterministic choice function, for each $x_i \in X$, $\lambda^1([\succ^{x_i}]) \in \{0, 1\}$. Next, by proceeding inductively, we assume that for any $k \in \{1, \dots, n-1\}$, there is a signed weight function λ^k that takes values $\{-1, 0, 1\}$ over $\mathcal{P}^k$ and represents C_k. It remains to show that we can construct λ^{k+1} taking values $\{-1, 0, 1\}$ over $\mathcal{P}^{k+1}$, and that represents C_{k+1}. We know from Step 1 of the proof of Theorem 1 that to show this it is sufficient to construct λ^{k+1} such that (RS) and (CS) holds. However, this time, in addition to satisfying (RS) and (CS), we require each $\lambda^{k+1}_{ij} \in \{-1, 0, 1\}$.

First, note that equalities (RS) and (CS) can be written as a system of linear equations: $A\lambda = b$, where $A = [a_{ij}]$ is a $(k! + (n-k)) \times (n-k)k!$ matrix with entries $a_{ij} \in \{0, 1\}$, and $b = [\lambda^k([\succ^k_1]), \dots, \lambda^k([\succ^k_{k!}]), q(x_1, S), \dots, q(x_{n-k}, S)]$ is the column vector of size $k! + (n-k)$. Let Q denote the associated polyhedron, i.e. $Q = \{\lambda \in \mathbb{R}^{(n-k)k!} : A\lambda = b \text{ and } -1 \leq \lambda \leq 1\}$. A matrix is **totally unimodular** if the determinant of each square submatrix is 0, 1 or -1. Following result directly follows from Theorem 2 of Hoffman & Kruskal (2010).

Lemma 1 (Hoffman & Kruskal (2010)) *If matrix A is totally unimodular, then the vertices of Q are integer valued.*

Heller & Tompkins (1956) provide the following sufficient condition for a matrix being totally unimodular.

Lemma 2 (Heller & Tompkins (1956)) *Let A be an $m \times n$ matrix whose rows can be partitioned into two disjoint sets R_1 and R_2. Then, A is totally unimodular if:*

1. *Each entry in A is 0, 1, or -1;*
2. *Each column of A contains at most two non-zero entries;*
3. *If two non-zero entries in a column of A have the same sign, then the row of one is in R_1, and the other is in R_2;*
4. *If two non-zero entries in a column of A have opposite signs, then the rows of both are in R_1, or both in R_2.*

Next, by using Lemma 2, we show that the matrix that is used to define (RS) and (CS) as a system of linear equations is totally unimodular. To see this, let A be the matrix defining the polyhedron Q. Since $A = [a_{ij}]$ is a matrix with entries $a_{ij} \in \{0, 1\}$, (1) and (4) are directly satisfied. To see that (2) and (3) also hold, let $R_1 = [1, \ldots, k!]$ consist of the the first $k!$ rows and $R_2 = [1, \ldots, n-k]$ consist of the the remaining $n-k$ rows of A. Note that for each $i \in R_1$, the i^{th} row A_i is such that $A_i\lambda = \lambda^k([\succ_i^k])$. That is, for each $j \in \{(i-1)k!, \ldots, ik!\}$, $a_{ij} = 1$ and the rest of A_i equals 0. For each $i \in R_2$, the i^{th} row A_i is such that $A_i\lambda = q(x_i, A)$. That is, for each $j \in \{i, i+k!, \ldots, i+(n-k-1)k!\}$, $a_{ij} = 1$ and the rest of A_i equals 0. To see that (2) and (3) hold, note that for each $i, i' \in R_1$ and $i, i' \in R_2$, the non-zero entries of A_i and $A_{i'}$ are disjoint. It follows that for each column there can be at most two rows with value 1, one in R_1 and the other in R_2.

Finally, it follows from the construction in Step 3 of the proof of Theorem 1 that Q is nonempty, since there is λ vector with entries taking values in the $[-1, 1]$ interval. Since, as shown above, A is totally unimodular, it directly follows from Lemma 1 that the vertices of Q are integer valued. Therefore, λ^{k+1} can be constructed such that (RS) and (CS) hold, and each $\lambda_{ij}^{k+1} \in \{-1, 0, 1\}$. ■

Remark 2 The constructed pro-con representation is a rather parsimonious one. To see this, consider a more stringent pro-con model, in which if an alternative x is chosen from a choice set S, it is barely chosen in the sense $|Pros(x,S)| - |Cons(x,S)| = 1$, and if an alternative y is not chosen, it is barely not chosen in the sense $|Pros(y,S)| - |Cons(y,S)| = 0$. It follows from the proof of Theorem 3 that the same anything–goes-result for this more demanding model.

4.3 Uniqueness

For a given pcM $\langle \succ, \triangleright \rangle$, let for each $S \in \Omega$ and $x \in S$, $Pros^k(x,S)$ be the set of pro-orderings at which x is the k^{th}-top-ranked alternative in S. Similarly, let $Cons^k(x,S)$ be the set of con-orderings at which x is the k^{th}-top-ranked alternative in S. We show that for a given choice function, the difference between the number of pro-orderings at which x is the k^{th}-top-ranked alternative in S and the number of con-orderings at which x is the k^{th}-top-ranked alternative in S is the same for each pro-con representation of the given choice function. We obtain this result as a direct corollary to our Proposition 1.

Corollary 1 *If $\langle \succ, \triangleright \rangle$ and $\langle \succ', \triangleright' \rangle$ are pro-con representations of the same choice function C, then for each $S \in \Omega$, $x \in S$, and $k \in \{1, \ldots, n\}$, both representations lead the same $|Pros^k(x,S)| - |Cons^k(x,S)|$ value.*

Proof. Since each pro- and con-ordering has a unit weight at each pro-con representation of a given choice function, $|Pros^k(x,S)| - |Cons^k(x,S)|$ equals $\lambda(x = B_k | S, \succ, \triangleright)$. Then, it follows from Proposition 1 that $|Pros^k(x,S)| - |Cons^k(x,S)|$ is fixed for each pro-con representation. ■

4.4 Extension to multi-valued choice

There are instances in which an agent must choose more than a single alternative from a choice set. For example, consider a school that chooses a cohort from a set of applicants or a professor who chooses a set of questions out of his archive to prepare an exam. As for the random choice, imagine that we have access the support of the random choice function, but not the frequencies, then the observed choice behaviour yields a choice rule.[20]

So far, we have assumed that the observed choice behavior is summarized by a choice function or a random choice rule. Both models rule out the possibility that choice can be multi-valued. Formally, a **choice rule** $\mathbb{C} : \Omega \to \Omega$ such that for each $S \in \Omega$, $\mathbb{C}(S) \subset S$. A choice rule is pro-con rational if there exists a pro-con model $\langle \succ, \triangleright \rangle$ such that for each choice set $S \in \Omega$, $\mathbb{C}(S) = \arg\max_{x \in S}(|Pros(x, S)| - |Cons(x, S)|)$. That is, for each choice set $S \in \Omega$ and $x \in S$, $x \in \mathbb{C}(S)$ if and only if $|Pros(x, S)| - |Cons(x, S)| \geq |Pros(y, S)| - |Cons(y, S)|$ for each $y \in S$.

A natural question is if our result in Theorem 3 extends to choice rules or not. To see that not every choice rule is pro-con rational, consider the choice rule $\mathbb{C}$ defined on $\{x, y, z\}$ such that $\mathbb{C}(\{x, y, z\}) = \{x, y\}$, $\mathbb{C}(\{x, y\}) = \{x\}$, $\mathbb{C}(\{y, z\}) = \{y\}$, and $\mathbb{C}(\{x, z\}) = \{z\}$. It is easy to see that $\mathbb{C}$ is not pro-con rational.[21] The stringency in here derives from the requirement that each ordering can be used only once as a pro- or con-ordering in a pro-con model. In contrast, if we allow multiple appearance of an ordering as a pro- or con-ordering, then every choice rule can be recovered. To see this, let $\mathbb{C}$ be a choice rule, and let p be the associated RCR such that for each $S \in \Omega$ and $x \in \mathbb{C}(S)$, $p(x, S) = 1/|\mathbb{C}(S)|$. It follows from Theorem 1 that there is a random pro-con model $\langle \succ, \triangleright, \lambda \rangle$ which represents p. Moreover, it follows from the construction in the proof of Theorem 1 that if for each $S \in \Omega$ and $x \in \mathbb{C}(S)$, $p(x, S)$

[20] See, for example, Fishburn (1978) who explores a connection in this vein.

[21] The two stage threshold representation analyzed by Manzini et al. (2013) has a similar feature. In that, although each choice function has a two-stage threshold representation, this does not hold for choice rules. That is, for each choice function there is a triplet $\langle f, \theta, g \rangle$ such that for each $S \in \Omega$, the alternative that maximizes $g(x)$ subject to $f(x) \geq \theta(S)$ is chosen, However, such a two stage threshold representation can not be obtained for every choice rule.

is a rational number, then for each $\succ_i \in \succ$ and $\rhd_j \in \rhd$, we can choose $\lambda(\succ_i) = m_i/M$ and $\lambda(\rhd_j) = m_j/M$, where m_i, m_j, M are positive integers. Now, consider a list (or a profile) of pro-orderings with m_i-many copies of $\succ_i$ and m_j-many copies of $\rhd_j$ for each $\succ_i \in \succ$ and $\rhd_j \in \rhd$. It directly follows from this construction that for each $S \in \Omega$ and $x \in S$, $x \in \mathbb{C}(S)$ if and only if x maximizes the difference between number of pro-orderings at which x is top-ranked in S and the number of con-orderings at which x is top-ranked in S.

4.5 Plurality-rationalizable choice rules

We propose a collective decision making model based on plurality voting. It turns out that this model is closely related to our pro-con choice model. To introduce this model, let $[\succ^*] = [\succ_1^*, \ldots, \succ_m^*]$ be a preference profile, which is a list of orderings. In contrast to a set of orderings, denoted by $\succ$ or $\rhd$, an ordering $\succ_i^*$ can appear more than once in a preference profile $[\succ^*]$. For each choice set $S \in \Omega$ and $x \in S$, x is **a plurality winner of** $[\succ^*]$ **in** S if for each $y \in S \setminus \{x\}$, the number of orderings in $[\succ^*]$ that top rank x in S is more than or equal to the number of orderings in $[\succ^*]$ that top rank y in S. That is, for each $y \in S \setminus \{x\}$, $|\{\succ_i^* \in [\succ^*] : \ x = max(S, \succ_i^*)\}| \geq |\{\succ_i^* \in [\succ^*] : \ y = max(S, \succ_i^*)\}|$. Next, we define plurality-rationalizability, then by using our Theorem 3, we show that every choice rule is plurality-rationalizable.

Definition 3 *A choice rule C is **plurality-rationalizable** if there is preference profile $[\succ^*]$ such that for each choice set $S \in \Omega$ and $x \in S$, $x \in \mathbb{C}(S)$ if and only if x is a plurality winner of $[\succ^*]$ in S.*

Proposition 2 *Every choice rule is plurality-rationalizable.*[22]

Proof. Let $\mathbb{C}$ be a choice rule. In Section 4.4, by using Theorem 3, we show that if we allow multiple appearance of an ordering as a pro- or con-ordering, then every choice rule is pro-con rational. First, to formalize this representation, let $\succ$ and $\rhd$ be the set of pro- and con-orderings such that each $\succ_i \in \succ$ ($\rhd_i \in \rhd$) is copied k_i times to

[22]Our initial result was for choice functions. We thank Vicki Knoblauch and an anonymous referee for suggesting the extension to choice rules.

represent $\mathbb{C}$. Then, define for each $S \in \Omega$ and $x \in S$, $SPros(x, S) = \sum_{\{\succ_i \in Pros(x,S)\}} k_i$ and $SCons(x, S) = \sum_{\{\triangleright_i \in Cons(x,S)\}} k_i$, where $Pros(x, S)$ and $Cons(x, S)$ are defined as usual with respect to $\succ$ and $\triangleright$. Now, we know that for each $S \in \Omega$ and $x \in S$, $x \in \mathbb{C}(S)$ if and only if $x \in \arg\max_{x \in S}(|SPros(x, S, \succ^*)| - |SCons(x, S, \triangleright^*)|)$.

Now, to construct the desired preference profile, let $k = max_{\{\triangleright_i \in \triangleright^*\}} k_i$, and begin with the list of all orderings defined on X copied k times. This is preference profile with $kn!$ elements. Then, eliminate k_i copies of the inverse of each ordering $\triangleright_i \in \triangleright$, and add k_i copies of each ordering $\succ_i \in \succ$. Note that since we have k copies of each ordering, the elimination part is well-defined. Let $[\succ^*]$ be the obtained preference profile.

We show that for each $S \in \Omega$ and $x \in S$, $x \in \mathbb{C}(S)$ if and only if x is a plurality winner of $[\succ^*]$ in S. We know that $x \in \mathbb{C}(S)$ if and only if for each $y \in S \setminus \{x\}$, $|SPros(x, S)| - |SCons(x, S)| \geq |SPros(y, S)| - |SCons(y, S)|$. Now, note that by construction of $[\succ^*]$, for each $y \in S$ the number of orderings in $[\succ^*]$ that top rank y in S equals k times the number of all orderings that top rank y in S, added to $|SPros(y, S)| - |SCons(y, S)|$. Since for each $y \in S$, the number of all orderings that top rank y in S is fixed, it follows that $x \in \mathbb{C}(S)$ if and only if x is a plurality winner of $[\succ^*]$ in S. ■

Remark 3 If we restrict our attention to choice functions, then we can consider an even more stringent model. In which, we require that an alternative x is chosen from a choice set S if and only if x is the plurality winner *at the margin*, in the sense that if x receives k votes then each other alternative receives $k - 1$ votes. It follows from Remark 2 and the proof of Proposition 2 that every choice function is plurality-rationalizable via this more demanding model.

In an early paper McGarvey (1953) shows that for each asymmetric and complete binary relation, there exists a preference profile such that the given binary relation is obtained from the preference profile by comparing each pair of alternatives via majority voting.[23] We obtain McGarvey's result, as a corollary to Proposition 2. To

[23]Stearns (1959) finds upper and lower bounds on the number of voters to generate any binary relation. Knoblauch (2016) provides an extension for infinite sets.

see this, note that if we restrict a choice rule to binary choice sets, then we obtain an asymmetric and complete binary relation. Since for binary choices, being a plurality winner means being a majority winner, McGarvey's result directly follows. Conversely, by following McGarvey's construction in his proof one can obtain the result of our Theorem 3 for the binary choice functions. However, this is not directly clear from the statement of his result. In that, first, for a given preference profile a la McGarvey, it is not clear how to classify the preferences as pro-orderings and con-orderings as to be independent of the pair of alternatives to be compared. Second, McGarvey does not explicitly rule out using replica preferences. Moreover, the more stringent representation formulated in Remark 3 is not obtained via McGarvey's construction.

5 Conclusion

Our main results show that the pro-con model–an additive model similar to the RUM–provides a language to describe any choice behavior in terms of structurally-invariant primitives. The structural invariance of the pro-con model reflects itself as a form of uniqueness, which is similar to the uniqueness of a random utility model. Our results and examples indicate that analyzing pro-con choice model with restricted pro- and con-orderings may lead to insightful results. In this vein, we analyzed choice problems in which there are two unobservable orderings of the alternatives that are relevant for choice. Our axiomatic characterization indicates that there is a tight connection between this model and the attraction effect phenomena. Knowing that each choice function is pro-con rational facilitates identification of other inclusive choice models. We present an application along these lines, in which we show that each choice rule is plurality-rationalizable. Although our study covers a rather extensive treatment of the pro-con model, we can hardly claim that it is exhaustive, as it leads to a wide variety of directions yet to be pursued.

6 Proof of Theorem 1

We start by proving some lemmas that are critical for proving the theorem. First, we use a result by Ford Jr & Fulkerson (2015)[24] as Lemma 3. Then, our Lemma 4 follows directly. Next, by using Lemma 4, we prove Lemma 5, which shows that, under suitable conditions, Lemma 3 holds for any real-valued row and column vectors.

Lemma 3 (Ford Jr & Fulkerson (2015)) *Let $R = [r_1, \dots, r_m]$ and $C = [c_1, \dots, c_n]$ be positive real-valued vectors with $\sum_{i=1}^{m} r_i = \sum_{j=1}^{n} c_j$. There is an $m \times n$ matrix $A = [a_{ij}]$ such that A has row sum vector R and column sum vector C, and each entry $a_{ij} \in [0, 1]$ if and only if for each $I \subset \{1, 2, \dots, m\}$ and $J \subset \{1, 2, \dots, n\}$,*

$$|I||J| \geq \sum_{i \in I} r_i - \sum_{j \notin J} c_j. \tag{FF}$$

Lemma 4 *Let $R = [r_1, \dots, r_m]$ and $C = [c_1, \dots, c_n]$ be positive real-valued vectors with $0 \leq r_i \leq 1$ and $0 \leq c_j \leq m$ such that $\sum_{i=1}^{m} r_i = \sum_{j=1}^{n} c_i$. Then there is an $m \times n$ matrix $A = [a_{ij}]$ such that A has row sum vector R and column sum vector C, and each entry $a_{ij} \in [0, 1]$.*

Proof. Given such R and C, since for each $i \in \{1, 2, \dots, m\}$, $0 \leq r_i \leq 1$, we have for each $I \subset \{1, 2, \dots, m\}$, $\sum_{i \in I} r_i \leq |I|$. Then, it directly follows that (FF) holds. ■

Next by using Lemma 4, we prove Lemma 5, which plays a key role in proving Theorem 1.

Lemma 5 *Let $R = [r_1, \dots, r_m]$ and $C = [c_1, \dots, c_n]$ be real-valued vectors with $-1 \leq r_i \leq 1$ and $-m \leq c_j \leq m$ such that $\sum_{i=1}^{m} r_i = \sum_{j=1}^{n} c_j$. If $2m \geq \sum_{i=1}^{m} |r_i| + \sum_{j=1}^{n} |c_j|$, then there is an $m \times n$ matrix $A = [a_{ij}]$ such that:*

i. A has row sum vector R and column sum vector C,

ii. each entry $a_{ij} \in [-1, 1]$, and

[24] This result, as stated in Lemma 3, but with integrality assumptions on R, C, and A follows from Theorem 1.4.2 in Brualdi & Ryser (1991). They report that Ford Jr & Fulkerson (2015) prove, by using network flow techniques, that the theorem remains true if the integrality assumptions are dropped, and the conclusion asserts the existence of a real nonnegative matrix.

iii. for each $j \in \{1, \ldots, n\}$, $\sum_{i=1}^{m} |a_{ij}| \leq |c_j| + \max\{0, \frac{\sum_{i=1}^{m} |r_i| - \sum_{j=1}^{n} |c_j|}{n}\}$.

Proof. Since r_i and c_j values can be positive or negative, although the sum of the rows equals the sum of the column, their absolute values may not be the same. We analyze two cases separately, where $\sum_{i=1}^{m} |r_i| \geq \sum_{j=1}^{n} |c_j|$ and $\sum_{i=1}^{m} |r_i| < \sum_{j=1}^{n} |c_j|$. Before proceeding with these cases, first we introduce some notation and make some elementary observations.

For each real number x, let $x^+ = \max\{x, 0\}$ and $x^- = \min\{x, 0\}$. Note that for each x, $x^+ + x^- = x$. Let $R^+ = [r_1^+, \ldots, r_m^+]$ and $R^- = [r_1^-, \ldots, r_m^-]$. Define the n-vectors C^+ and C^- respectively. Next, let $\Sigma_{R^+} = \sum_{i=1}^{m} r_i^+$, $\Sigma_{R^-} = \sum_{i=1}^{m} r_i^-$, $\Sigma_{C^+} = \sum_{j=1}^{n} c_j^+$ and $\Sigma_{C^-} = \sum_{j=1}^{n} c_j^-$. That is, $\Sigma_{R^+}(\Sigma_{R^-})$ and $\Sigma_{C^+}(\Sigma_{C^-})$ are the sum of the positive (negative) rows in R and columns in C. Since the sum of the rows equals the sum of the columns, we have $\Sigma_{R^+} + \Sigma_{R^-} = \Sigma_{C^+} + \Sigma_{C^-}$.

For each row vector R and column vector C, suppose for each $i \in \{1, \ldots, m_1\}$, $r_i \geq 0$ and for each $i \in \{m_1 + 1, \ldots, m\}$, $r_i < 0$. Similarly, suppose for each $j \in \{1, \ldots, n_1\}$, $c_j \geq 0$ and for each $j \in \{n_1 + 1, \ldots, n\}$, $c_j < 0$. Now, let $R^1(R^2)$ be the m_1-vector ($(m - m_1)$-vector), consisting of the non-negative (negative) components of R. Similarly, for each column vector C, let $C^1(C^2)$ be the n_1-vector ($(n - n_1)$-vector), consisting of the non-negative (negative) components of C. It directly follows from the definitions that $\sum_{i=1}^{m_1} r_i = \sum_{i=1}^{m} r_i^+$ and $\sum_{i=m_1+1}^{m} r_i = \sum_{i=1}^{m} r_i^-$. Similarly, $\sum_{j=1}^{n_1} c_j = \sum_{j=1}^{n} c_j^+$ and $\sum_{j=n_1+1}^{n} c_j = \sum_{j=1}^{n} c_j^-$.

Case 1: Suppose that $\sum_{i=I} |r_i| \geq \sum_{j \in J} |c_j|$. First, for each $j \in \{1, \ldots, n\}$, let

$$\epsilon_j = \frac{\Sigma_{R^+} - \Sigma_{C^+}}{n}.$$

Note that since $\sum_{i=1}^{m} |r_i| \geq \sum_{j=1}^{n} |c_j|$, we have $\Sigma_{R^+} \geq \Sigma_{C^+}$ and $\Sigma_{R^-} \leq \Sigma_{C^-}$. Moreover, since the sum of the rows equals the sum of the columns, we have $\Sigma_{R^+} - \Sigma_{C^+} = \Sigma_{C^-} - \Sigma_{R^-}$. Therefore, by the choice of ϵ_j, we get

$$\sum_{i=1}^{m} r_i^+ = \sum_{j=1}^{n} c_j^+ + \epsilon_j \text{ and } \sum_{i=1}^{m} r_i^- = \sum_{j=1}^{n} c_j^- - \epsilon_j. \quad (5)$$

Next, consider row-column vector pairs $(R^1, C^+ + \epsilon)$ and $(-R^2, -(C^- - \epsilon))$, where ϵ is the non-negative n-vector such that each ϵ_j is as defined above. It follows from (5) that for both pairs the sum of the rows equals the sum of the columns. Now we apply Lemma 4 to the row-column vector pairs $(R^1, C^+ + \epsilon)$ and $(-R^2, -(C^- - \epsilon))$. It directly follows that there exists a positive $m_1 \times n$ matrix A^+ and a negative $(m - m_1) \times n$ matrix A^- that satisfy (i) and (ii). We will obtain the desired matrix A by augmenting A^+ and A^-. We illustrate A^+ and A^- below.

	$(c_1^+ + \epsilon_1)$ $(c_2^+ + \epsilon_2)$ $(c_3^+ + \epsilon_3)$ $\cdots$ $(c_n^+ + \epsilon_n)$	
$r_1 \geq 0$ $r_2 \geq 0$ $\vdots$ $r_{m_1} \geq 0$	A^+	
	A^-	$r_{m_1+1} < 0$ $\vdots$ $r_m < 0$
	$(c_1^- - \epsilon_1)$ $(c_2^- - \epsilon_2)$ $(c_3^- - \epsilon_3)$ $\cdots$ $(c_n^- - \epsilon_n)$	

Since A^+ and A^- satisfy (i) and (ii), A satisfies (i) and (ii). To see that A satisfies (iii), for each $j \in \{1, \ldots, n\}$, consider $\sum_{i=1}^{m} |a_{ij}|$. Note that, by the construction of A^+ and A^-, for each $j \in \{1, \ldots, n\}$,

$$\sum_{i=1}^{m} |a_{ij}| = c_j^+ + \epsilon_j + (-c_j^- + \epsilon_j) = |c_j| + 2\epsilon_j = |c_j| + 2\frac{\Sigma_{R^+} - \Sigma_{C^+}}{n}. \tag{6}$$

Since for each $j \in \{1, \ldots, n\}$, $c_j = c_j^+ + c_j^-$ such that either $c^+ = 0$ or $c_j^- = 0$, we get $|c_j| = c_j^+ - c_j^-$. To see that (iii) holds, observe that $\sum_{i=1}^{m} |r_i| - \sum_{j=1}^{n} |c_j| = \Sigma_{R^+} - \Sigma_{C^+} + \Sigma_{C^-} - \Sigma_{R^-}$. Since the sum of the rows equals the sum of the columns, i.e. $\Sigma_{R^+} + \Sigma_{R^-} = \Sigma_{C^+} + \Sigma_{C^-}$, we also have $\Sigma_{R^+} - \Sigma_{C^+} = \Sigma_{C^-} - \Sigma_{R^-}$. This observation, together with (6), implies that (iii) holds.

Case 2 Suppose that $\sum_{i=1}^{m} |r_i| < \sum_{j=1}^{n} |c_j|$. First, we show that there exists a non-negative m-vector ϵ such that

(E1) for each $i \in \{1, \ldots, m\}$, $r_i^+ + \epsilon_i \leq 1$ and $r_i^- - \epsilon_i \geq -1$, and

(E2) $\sum_{i=1}^{m} r_i^+ + \epsilon_i = \sum_{j=1}^{n} c_j^+$ (equivalently $\sum_{i=1}^{m} r_i^- - \epsilon_i = \sum_{j=1}^{n} c_j^-$) holds.

Step 1: We show that if $\Sigma_{C^+} - \Sigma_{R^+} \leq m - \sum_{i=1}^{m} |r_i|$, then there exists a non-negative m-vector ϵ that satisfies (E1) and (E2). To see this, first note that $m -$

$\sum_{i=1}^{m} |r_i| = \sum_{i=1}^{m}(1 - |r_i|)$. Next, note that, by simply rearranging the terms, we can rewrite (E2) as follows:

$$\sum_{i=1}^{m} \epsilon_i = \Sigma_{C^+} - \Sigma_{R^+} \,. \tag{7}$$

Since $\Sigma_{C^+} - \Sigma_{R^+} \leq \sum_{i=1}^{m}(1 - |r_i|)$, for each $i \in \{1, \ldots, m\}$, we can choose an ϵ_i such that $0 \leq \epsilon_i \leq 1 - |r_i|$ and (7) holds. It directly follows that the associated ϵ vector satisfies (E1) and (E2).

Step 2: We show that since $2m \geq \sum_{i=1}^{m} |r_i| + \sum_{j=1}^{n} |c_j|$, we have $\Sigma_{C^+} - \Sigma_{R^+} \leq m - \sum_{i=1}^{m} |r_i|$. First, it directly follows from the definitions that

$$\sum_{i=1}^{m} |r_i| + \sum_{j=1}^{n} |c_j| = \Sigma_{R^+} - \Sigma_{R^-} + \Sigma_{C^+} - \Sigma_{C^-} \,.$$

Since the sum of the rows equals the sum of the columns, i.e. $\Sigma_{R^+} + \Sigma_{R^-} = \Sigma_{C^+} + \Sigma_{C^-}$, we also have $\Sigma_{R^+} - \Sigma_{C^-} = \Sigma_{C^+} - \Sigma_{R^-}$. It follows that

$$\Sigma_{C^+} - \Sigma_{R^-} \leq m.$$

Finally, if we subtract $\sum_{i=1}^{m} |r_i|$ from both sides of this equality, we obtain $\Sigma_{C^+} - \Sigma_{R^+} \leq m - \sum_{i=1}^{m} |r_i|$, as desired.

It follows from Step 1 and Step 2 that there exists a non-negative m-vector ϵ that satisfies (E1) and (E2). Now, consider the row-column vector pairs $(R^+ + \epsilon, C^1)$ and $(-(R^- - \epsilon), -C^2)$. Since ϵ satisfies (E1) for each $i \in \{1, \ldots, m\}$, $r_i^+ + \epsilon_i \in [0, 1]$ and $r_i^- - \epsilon_i \in [-1, 0]$. Since ϵ satisfies (E2), for both of the row-column vector pairs the sum of the rows equals the sum of the columns. Therefore, we can apply Lemma 4 to row-column vector pairs $(R^+ + \epsilon, C^1)$ and $(-(R^- - \epsilon), -C^2)$. It directly follows that there exists a positive $m \times n_1$ matrix A^+ and a negative $m \times (n - n_1)$ matrix A^- that satisfy (i) and (ii). We obtain the desired matrix A by augmenting A^+ and A^-. We illustrate A^+ and A^- below.

	$c_1 \quad c_2 \quad \cdots \quad c_{n_1} \geq 0$		
$(r_1^+ + \epsilon_1)$			$(r_1^- - \epsilon_1)$
$(r_2^+ + \epsilon_2)$			$(r_2^- - \epsilon_2)$
$\vdots$	A^+	A^-	$\vdots$
$\vdots$			$\vdots$
$(r_m^+ + \epsilon_m)$			$(r_m^- - \epsilon_m)$
		$c_{n_1+1} < 0 \quad \cdots \quad c_n$	

Since A^+ and A^- satisfy (i) and (ii), A satisfies (i) and (ii). In this case, since we did not add anything to the columns and each entry in $A^+(A^-)$ is non-negative (negative), for each $j \in \{1, \ldots, n\}$, $\sum_{i=1}^m |a_{ij}| = |c_j|$. Therefore, A also satisfies (iii). ■

To prove Theorem 1, let p be an RCF and $\mathcal{P}$ denote the collection of all orderings on X. First, we show that there is a **signed weight function** $\lambda : \mathcal{P} \to [-1, 1]$ that **represents** p, i.e. for each $S \in \Omega$ and $x \in S$, $p(x, S)$ is the sum of the weights over $\{\succ_i \in \mathcal{P} \ : \ x = max(S, \succ_i)\}$. Note that λ can assign negative weights to orderings. Once we obtain this signed weight function λ, let $\succ$ be the collection of orderings that receive positive weights, and let $\rhd'$ be the collection of orderings that receive negative weights. Let $\rhd$ be the collection of the inverse of the orderings in $\rhd'$. Finally, let λ^* be the weight function obtained from λ by assigning the absolute value of the weights assigned by λ. It directly follows that p is pro-con rational with respect to the RpcM $\langle \succ, \rhd, \lambda^* \rangle$. We first introduce some notation and present crucial observations to construct the desired signed weight function λ.

Let p be a given RCF and Let $q : X \times \Omega \to \mathbb{R}$ be a mapping such that for each $S \in \Omega$ and $a \notin S$, $q(a, S) = q(a, S \cup \{a\})$ holds. Next, we present a result that is directly obtained by applying the *Möbius inversion*.[25]

Lemma 6 *For each choice set $S \in \Omega$, and alternative $a \in S$,*

$$p(a, S) = \sum_{S \subset T \subset X} q(a, T) \tag{8}$$

[25] See Stanley (1997), Section 3.7. See also Fiorini (2004), who makes the same observation.

if and only if

$$q(a,S) = \sum_{S\subset T\subset X} (-1)^{|T|-|S|} p(a,T) \tag{9}$$

Proof. For each alternative $a \in X$, note that $p(a,\cdot)$ and $q(a,\cdot)$ are real-valued functions defined on the domain consisting of all $S \in \Omega$ with $a \in S$. Then, by applying the Möbius inversion, we get the conclusion. ■

Lemma 7 *For each choice set $S \in \Omega$ with $|S| = n-k$,*

$$\sum_{a\in X} |q(a,S)| \leq 2^k. \tag{10}$$

Proof. First, note that (10) can be written as follows:

$$\sum_{a\in S} |q(a,S)| + \sum_{b\notin S} |-q(b,S)| \leq 2^k. \tag{11}$$

For a set of real numbers,$\{x_1, x_2, \ldots x_n\}$, to show $\sum_{i=1}^n |x_i| \leq 2d$, it suffices to show that for each $I \subset \{1, 2, \cdots, n\}$, we have $-d \leq \sum_{i\in I} x_i \leq d$. Now, as the set of real numbers, consider $\{q(a,S)\}_{a\in X}$. It follows that to show that (11) holds, it suffices to show that for each $S_1 \subset S$ and $S_2 \subset X \setminus S$,

$$-2^{k-1} \leq \sum_{a\in S_1} q(a,S) - \sum_{b\in S_2} q(b,S) \leq 2^{k-1}$$

holds. To see this, first, for each $S_1 \subset S$ and $S_2 \subset X \setminus S$, it follows from Lemma 6 that for each $a \in S_1$ and for each $b \in S_2$, we have

$$q(a,S) = \sum_{S\subset T\subset X} (-1)^{|T|-|S|} p(a,T) \text{ and } q(b,S) = \sum_{S\subset T\subset X} (-1)^{|T|-|S|-1} p(b,T). \tag{12}$$

Note that we obtain the second equality from Lemma 6, since for each $b \notin S$, by definition of $q(b,S)$, we have $q(b,S) = q(b, S\cup\{b\})$. Next, note that for each $T \in \Omega$ with $S \subset T$, $a \in S$, and $b \notin S$, $p(a,T)$ has the opposite sign of $p(b,T)$. Now, suppose for each $b \in S_2$, we multiply $q(b,S)$ with -1. Then, it follows from (12) that

$$\sum_{a\in S_1} q(a,S) - \sum_{b\in S_2} q(b,S) = \sum_{S\subset T\subset X} (-1)^{|T|-|S|} \sum_{a\in S_1\cup S_2} p(a,T). \tag{13}$$

Note that, for each $T \in \Omega$ such that $S \subset T$, $\sum_{a \in S_1 \cup S_2} p(a,T) \in [0,1]$. Therefore, the term $(-1)^{|T|-|S|} \sum_{a \in S_1 \cup S_2} p(a,T)$ adds at most 1 to the right-hand side of (13) if $|T|-|S|$ is even, and at least -1 if $|T|-|S|$ is odd. Since $|S| = n-k$, for each m with $n-k \leq m \leq n$, there are $\binom{k}{m-n+k}$ possible choice sets $T \in \Omega$ such that $S \subset T$ and $|T| = m$. Moreover, for each $i \in \{1, \ldots, k\}$, there are $\binom{k}{i}$ possible choice sets T such that $S \subset T$ and $|T| = n-k+i$. Now, the right-hand side of (13) reaches its maximum (minimum) when the negative (positive) terms are 0 and the positive (negative) terms are $1(-1)$. Thus, we get

$$-\sum_{i=0}^{\lfloor \frac{k-1}{2} \rfloor} \binom{k}{2i+1} \leq \sum_{S \subset T \subset X} (-1)^{|T|-|S|} \sum_{a \in S_1 \cup S_2} p(a,T) \leq \sum_{i=0}^{\lfloor \frac{k}{2} \rfloor} \binom{k}{2i}.$$

It follows from the *binomial theorem* that both leftmost and rightmost sums are equal to 2^{k-1}. This, combined with (13), implies

$$-2^{k-1} \leq \sum_{a \in S_1} q(a,S) - \sum_{b \in S_2} q(b,S) \leq 2^{k-1}.$$

Then, as argued before, it follows that $\sum_{a \in X} |q(a,S)| \leq 2^k$. ■

Now, we are ready to complete the proof of Theorem 1. Recall that we assume $|X| = n$. For each $k \in \{1, \ldots, n\}$, let $\Omega_k = \{S \in \Omega \ : \ |S| > n-k\}$. Note that $\Omega_n = \Omega$ and $\Omega_1 \subset \Omega_2 \subset \cdots \subset \Omega_n$. For each pair of orderings $\succ_1, \succ_2 \in \mathcal{P}$, $\succ_1$ is **k-identical** to $\succ_2$, denoted by $\succ_1 \sim_k \succ_2$, if the first k-ranked alternatives are the same. Note that $\sim_k$ is an equivalence relation on $\mathcal{P}$. Let $\mathcal{P}^k$ be the collection of orderings, such that each set (equivalence class) contains orderings that are k-identical to each other ($\mathcal{P}^k$ is the quotient space induced from $\sim_k$). For each $k \in \{1, \ldots, n\}$, let $[\succ^k]$ denote an **equivalence class** at $\mathcal{P}^k$, where $\succ^k$ linearly orders a fixed set of k alternatives in X.

Note that for each $k \in \{1, \ldots, n\}$, $S \in \Omega_k$ and $\succ_1, \succ_2 \in \mathcal{P}$, if $\succ_1 \sim_k \succ_2$, then since S contains more than $n-k$ alternatives, $\max(\succ_1, S) = \max(\succ_2, S)$. Therefore, for each $S \in \Omega_k$, it is sufficient to specify the weights on the equivalence classes contained in $\mathcal{P}^k$ instead of all the weights over $\mathcal{P}$. Let p_k be the restriction of p to Ω_k. Similarly, if λ is a signed weight function over $\mathcal{P}$, then let λ^k be the restriction of λ to $\mathcal{P}^k$, i.e. for each $[\succ^k] \in \mathcal{P}^k$, $\lambda^k[\succ^k] = \sum_{\succ_i \in [\succ^k]} \lambda(\succ_i)$. It directly follows that λ represents p if and only if for each $k \in \{1, \ldots, n\}$, λ^k represents p_k. In what follows, we inductively show that for each $k \in \{1, \ldots, n\}$, there is a signed weight function λ^k over $\mathcal{P}^k$ that represents p_k. For $k = n$ we obtain the desired λ.

For $k = 1$, $\Omega_1 = \{X\}$ and $\mathcal{P}^1$ consists of n-many equivalence classes such that each class contains all the orderings that top rank the same alternative, irrespective of whether these are chosen with positive probability. That is, if $X = \{x_1, \ldots, x_n\}$, then $\mathcal{P}^1 = \{[\succ^{x_1}], \cdots, [\succ^{x_n}]\}$, where for each $i \in \{1, \ldots, n\}$ and $\succ_i \in [\succ^{x_i}]$, $\max(X, \succ_i) = x_i$. Now, for each $x_i \in X$, define $\lambda^1([\succ^{x_i}]) = p(x_i, X)$. It directly follows that λ^1 is a signed weight function over $\mathcal{P}^1$ that represents p_1.

For $k = 2$, $\Omega_2 = \{X\} \cup \{X \setminus \{x\}\}_{x \in X}$ and $\mathcal{P}^2$ consists of $\binom{n}{2}$-many equivalence classes such that each class contains all the orderings that top rank the same two alternatives. Now, for each $[\succ_i^2] \in \mathcal{P}^2$ such that x_{i1} is the first-ranked alternative and x_{i2} is the second-ranked alternative, define $\lambda^2([\succ_i^2]) = p(x_{i2}, X \setminus \{x_{i1}\}) - p(x_{i2}, X)$. It directly follows that λ^2 is a signed weight function over $\mathcal{P}^2$ that represents p_2. Next, by our inductive hypothesis, we assume that for each $k \in \{1, \ldots, n-1\}$, there is a signed weight function λ^k over $\mathcal{P}^k$ that represents p_k. Next, we show that we can construct λ^{k+1} over $\mathcal{P}^{k+1}$ that represents p_{k+1}.

Note that $\mathcal{P}^{k+1}$ is a refinement of $\mathcal{P}^k$, in which each equivalence class $[\succ^k] \in \mathcal{P}^k$ is divided into sub-equivalence classes $\{[\succ_1^{k+1}], \cdots [\succ_{n-k}^{k+1}]\} \subset \mathcal{P}^{k+1}$. Given λ^k, we require λ^{k+1} satisfy for each $[\succ^k] \in \mathcal{P}^k$ the following

$$\lambda^k([\succ^k]) = \sum_{j=1}^{n-k} \lambda^{k+1}([\succ_j^{k+1}]). \tag{14}$$

If λ^{k+1} satisfies (14), then since induction hypothesis implies that λ^k represents p_k, we get for each $S \in \Omega_k$ and $x \in S$, $p(x, S) = \lambda^{k+1}\big(\{[\succ_j] \in \mathcal{P}^{k+1} \ : x = max(S, \succ_j)\}\big)$.

Next, we show that λ^{k+1} can be constructed such that (14) holds, and for each $S \in \Omega_{k+1} \backslash \Omega_k$, λ^{k+1} represents $p_{k+1}(S)$. To see this, pick any $S \in \Omega_{k+1} \backslash \Omega_k$. It follows that $|S| = n - k$. Let $S = \{x_1, , .., x_{n-k}\}$ and $X \setminus S = \{y_1, y_2, \cdots y_k\}$. Recall that each $[\succ^k] \in \mathcal{P}^k$ linearly orders a fixed set of k-many alternatives. Let $\{\succ^k\}$ denote the set of k alternatives ordered by $\succ^k$. Now, there exist $k!$-many $[\succ^k] \in \mathcal{P}^k$ such that $\{\succ^k\} = X \setminus S$. Let $\big\{[\succ_1^k], \cdots, [\succ_{k!}^k]\big\}$ be the collection of all such classes. Each ordering that belongs to one of these classes is a different ordering of the same set of

k alternatives.

Now, let $I = \{1, \ldots, k!\}$ and $J = \{1, \ldots, n-k\}$. For each $i \in I$ and $j \in J$, suppose that $\succ_{ij}^{k+1}$ linearly orders $X \setminus S$ as in $\succ_i^k$ and ranks x_j in the $k+1^{th}$ position. Consider the associated equivalence class $[\succ_{ij}^{k+1}]$. Next, we specify $\lambda^{k+1}([\succ_{ij}^{k+1}])$, the signed weight of $[\succ_{ij}^{k+1}]$, such that the resulting λ^{k+1} represents p_{k+1}. To see this, we proceed in two steps.

Step 1: First, we show that for each $S \in \Omega_{k+1} \backslash \Omega_k$, if the associated $\{\lambda_{ij}^{k+1}\}_{ij \in I \times J}$ satisfies the following two equalities for each $i \in I$ and $j \in J$,

$$\sum_{j \in J} \lambda_{ij}^{k+1} = \lambda^k([\succ_i^k]) \tag{RS}$$

$$\sum_{i \in I} \lambda_{ij}^{k+1} = q(x_j, S) \tag{CS}$$

then λ^{k+1} represents $p_{k+1}(S)$. For each $S \in \Omega$ and $x_j \in S$, $q(x_j, S)$ is as defined in (9) by using the given RCF p.

For each $S \in \Omega$ and $a \in S$, let $B(a, S)$ be the collection of all orderings at which a is the top-ranked alternative in S, and for each $k \in \mathbb{N}$ such that $n - k \leq |S|$, $\mathbf{B}^{k+1}(a, S)$ be the set of associated equivalence classes in $\mathcal{P}^{k+1}$, i.e. $B(a, S) = \{\succ \in \mathcal{P} : a = max(S, \succ)\}$ and $\mathbf{B}^{k+1}(a, S) = \{[\succ^{k+1}] \in \mathcal{P}^{k+1} : [\succ^{k+1}] \subset B(a, S)\}$. To prove the result we have to show that for each $x_j \in S$,

$$p(x_j, S) = \sum_{\{[\succ^{k+1}] \in \mathbf{B}^{k+1}(x_j, S)\}} \lambda^{k+1}([\succ^{k+1}]). \tag{15}$$

To see this, for each $\succ \in \mathcal{P}$ and $a \in X$, let $W(\succ, a)$ denote the set of alternatives that are worse than a at $\succ$ and a itself, i.e. $W(\succ, a) = \{x \in X : a \succ x\} \cup \{a\}$. For each $S \in \Omega$ with $a \in X$. Let $Q(a, S)$ be the collection of all orderings such that $W(\succ, a)$ is exactly $S \cup \{a\}$ and for each $k \in \mathbb{N}$ such that $n - k \leq |S|$, $\mathbf{Q}^{k+1}(a, S)$ be the set of associated equivalence classes in $\mathcal{P}^{k+1}$, i.e. $Q(a, S) = \{\succ \in \mathcal{P} : W(\succ, a) = S \cup \{a\}\}$ and $\mathbf{Q}^{k+1}(a, S) = \{[\succ^{k+1}] \in \mathcal{P}^{k+1} : [\succ^{k+1}] \subset Q(a, S)\}$. Note that, for each $x_j \in S$, we have $Q(x_j, S) = \bigcup_{i \in I} [\succ_{ij}^{k+1}]$. Moreover, it directly follows from the definitions of $Q(x_j, \cdot)$ and $B(x_j, \cdot)$ that

$$B(x_j, S) = \bigcup_{S \subset T} Q(x_j, T). \tag{16}$$

It follows from this observation that the right-hand side of (15) can be written as

$$\sum_{S\subset T}\sum_{\{[\succ^{k+1}]\in\mathbf{Q}^{k+1}(x_j,T)\}}\lambda^{k+1}([\succ^{t+1}]). \tag{17}$$

i. Since (CS) holds, we have

$$q(x_j,S)=\sum_{\{[\succ^{k+1}]\in\mathbf{Q}^{k+1}(x_j,S)\}}\lambda^{k+1}([\succ^{k+1}]). \tag{18}$$

ii. Next, we argue that for each $T\in\Omega$ such that $S\subsetneq T$,

$$q(x_j,T)=\sum_{\{[\succ^{k+1}]\in\mathbf{Q}^{k+1}(x_j,T)\}}\lambda^{k+1}([\succ^{k+1}]). \tag{19}$$

To see this, recall that by definition of $q(x_j,T)$ (9), we have

$$q(x_j,T)=\sum_{T\subset T'}(-1)^{|T'|-|T|}p(x_j,T'). \tag{20}$$

Since by the induction hypothesis, λ^k represents p_k, we have

$$p(x_j,T')=\sum_{\{[\succ^{k}]\in\mathbf{B}^{k}(x_j,T')\}}\lambda^{k}([\succ^{k}]). \tag{21}$$

Next, suppose that we substitute (21) into (20). Now, consider the set collection $\{B(x_j,T')\}_{\{T\subset T'\}}$. Note that if we apply the *principle of inclusion-exclusion* to this set collection, then we obtain $Q(x_j,T)$. It follows that

$$\sum_{T\subset T'}(-1)^{|T'|-|T|}\sum_{\{[\succ^{k}]\in\mathbf{B}^{k}(x_j,T')\}}\lambda^{k}([\succ^{k}])=\sum_{\{[\succ^{k}]\in\mathbf{Q}^{k}(x_j,T)\}}\lambda^{k}([\succ^{k}]). \tag{22}$$

Since (RS) holds, we have

$$\sum_{\{[\succ^{k}]\in\mathbf{Q}^{k}(x_j,T)\}}\lambda^{k}([\succ^{k}])=\sum_{\{[\succ^{k+1}]\in\mathbf{Q}^{k+1}(x_j,T)\}}\lambda^{k+1}([\succ^{k+1}]). \tag{23}$$

Thus, if we combine (20)-(23), then we obtain that (19) holds.

Now, (17) combined with (18) and (19) imply that the right-hand side of (15) equals to $\sum_{S\subset T}q(x_j,T)$. Finally, it follows from Lemma 6 that

$$p(x_j,S)=\sum_{S\subset T}q(x_j,T). \tag{24}$$

Thus, we obtain that (15) holds.

In what follows we show that for each $S \in \Omega_{k+1} \backslash \Omega_k$, there exists $k! \times (n-k)$ matrix $\lambda = [\lambda_{ij}^{k+1}]$ such that both (RS) and (CS) holds, and each $\lambda_{ij}^{k+1} \in [-1, 1]$. To prove this we use Lemma 5. For this, for each $i \in I$ let $r_i = \lambda^k([\succ_i^k])$ and for each $j \in J$ let $c_j = q(x_j, S)$. Then, let $R = [r_1, \ldots, r_{k!}]$ and $C = [c_1, \ldots, c_{n-k}]$. In Step 2, we show that the sum of C equals the sum of R. In Step 3, we show that for each $k > 1$, $2k! \geq \sum_{i=1}^{k!} |r_i| + \sum_{j=1}^{n-k} |c_j|$.

Step 2: We show that the sum of C equals the sum of R, i.e.

$$\sum_{j \in J} q(x_j, S) = \sum_{i \in I} \lambda^k [\succ_i^k]. \tag{25}$$

First, if we substitute (9) for each $q(x_j, S)$, then we get

$$\sum_{j \in J} q(x_j, S) = 1 + \sum_{j \in J} \sum_{S \subsetneq T} (-1)^{|T|-|S|} p(x_j, T). \tag{26}$$

Now, let $F(x_j)$ be the collection of orderings $\succ$ such that there exists $T \in \Omega$ such that $S \subsetneq T$ and x_j is the $\succ$-top-ranked alternative in T, i.e. $F(x_j) = \{\succ \in \mathcal{P} \ : \ max(T, \succ) = x_j \text{ for some } S \subsetneq T\}$. For each $k \in \mathbb{N}$ such that $n - k \leq |S|$, let $\mathbf{F}(x_j)$ be the set of associated equivalence classes in $\mathcal{P}^k$. Next, we show that for each $x_j \in S$,

$$\sum_{S \subsetneq T} (-1)^{|T|-|S|+1} p(x_j, T) = \sum_{\{[\succ^k] \in \mathbf{F}(x_j)\}} \lambda^k([\succ^k]). \tag{27}$$

To see this, first, since by the induction hypothesis, λ^k represents p_k, we can replace each $p(x_j, T)$ with $\sum_{\{[\succ^k] \in \mathbf{B}^k(x_j, T)\}} \lambda^k([\succ^k])$. Next, consider the set collection $\{B(x_j, T)\}_{\{S \subsetneq T\}}$. Since $\cup_{\{S \subsetneq T\}} B(x_j, T) = F(x_j)$, it follows from the *principle of inclusion-exclusion* that (27) holds. Next, when we substitute (27) in (26), we obtain

$$\sum_{j \in J} q(x_j, S) = 1 - \sum_{\{[\succ^k] \in \mathbf{F}(x_j)\}} \lambda^k([\succ^k]). \tag{28}$$

Then, since, by the induction hypothesis, λ^k represents p_k, we can replace 1 with $\sum_{\{[\succ^k] \in \mathcal{P}^k\}} \lambda^k([\succ^k])$. Finally, note that an equivalence class $[\succ^k] \notin \cup_{j \in J} \mathbf{F}(x_j)$ if and only if $\{\succ^k\} \cap S = \emptyset$. This means $\mathcal{P}^k \setminus \cup_{j \in J} \mathbf{F}(x_j) = \{[\succ_i^k]\}_{\{i \in I\}}$. It follows that (25) holds.

Step 3: To show that the base of induction holds, we showed that for $k = 1$ and $k = 2$, the desired signed weight functions exist. To get the desired signed weight functions for each $k+1 > 2$, we will apply Lemma 5. To apply Lemma 5, we have to show that for each $k \geq 2$, $\sum_{i=1}^{k!} |r_i| + \sum_{j=1}^{n-k} |c_j| \leq 2k!$. In what follows we show that this is true. That is, we show that for each $S \in \Omega_{k+1} \setminus \Omega_k$

$$\sum_{i \in I} |\lambda^k([\succ_i^k])| + \sum_{j \in J} |q(x_j, S)| \leq 2k!. \tag{29}$$

To see this, first we will bound the term $\sum_{i \in I} |\lambda^k([\succ_i^k])|$. As noted before, each $i \in I = \{1, \ldots, k!\}$ corresponds to a specific linear ordering of $X \setminus S$. For each $y \notin S$, there are $k-1!$ such different orderings that rank y at the k^{th} position. So, there are $k-1!$ different equivalence classes in $\mathcal{P}^k$ that rank y at the k^{th} position. Let $I(y)$ be the index set of these equivalence classes. Since $\{I(y)\}_{y \notin S}$ partitions I, we have

$$\sum_{i \in I} |\lambda^k([\succ_i^k])| = \sum_{y \notin S} \sum_{i \in I(y)} |\lambda^k([\succ_i^k])|. \tag{30}$$

Now, fix $y \notin S$ and let $T = S \cup \{y\}$. Since for each $i \in I(y)$, $[\succ_i^k] \in \mathbf{Q}^k(y, T)$ and vice versa, we have

$$\sum_{i \in I(y)} |\lambda^k([\succ_i^k])| = \sum_{[\succ_i^k] \in \mathbf{Q}^k(y,T)} |\lambda^k([\succ_i^k])|. \tag{31}$$

Recall that by the definition of $q(y, T)$, we have

$$q(y, T) = \sum_{[\succ_i^k] \in \mathbf{Q}^k(y,T)} \lambda^k([\succ_i^k]). \tag{32}$$

Next, consider the construction of the values $\{\lambda^k([\succ_i^k]\}_{\{i \in I(y)\}}$ from the previous step. For $k = 2$, as indicated in showing the base of induction, there is only one row; that is, there is a single $\{[\succ_i^k]\} = \mathbf{Q}^k(y, T)$. Therefore, we directly have $|\lambda^k([\succ_i^k])| = |q(y, T)|$. For $k > 2$, we construct λ^k by applying Lemma 5. It follows from iii of Lemma 5 that

$$\sum_{[\succ_i^k] \in \mathbf{Q}^k(y,T)} |\lambda^k([\succ_i^k])| \leq |q(y, T)| + \frac{(k-1)!}{n-k+1}. \tag{33}$$

Now, if we sum (33) over $y \notin S$, we get

$$\sum_{y \notin S} \sum_{[\succ_i^k] \in \mathbf{Q}^k(y,S \cup y)} |\lambda^k([\succ^k])| \leq \left(\sum_{y \notin S} |q(y, S \cup y)| \right) + \frac{k!}{n-k+1}. \tag{34}$$

Recall that by definition, we have $\mathbf{Q}^k(y, S \cup y) = \mathbf{Q}^k(y, S)$ and $q(y, S \cup y) = q(y, S)$. Similarly, since each $j \in J = \{1, \ldots, n\}$ denotes an alternative $x_j \in S$, we have $\sum_{x \in S} |q(x, S)| = \sum_{j \in J} |q(x_j, S)|$. Now, if we add $\sum_{j \in J} |q(x_j, S)|$ to both sides of (34), then we get

$$\sum_{i \in I} |\lambda^k([\succ_i^k])| + \sum_{j \in J} |q(x_j, S)| \leq \sum_{x \in X} |q(x, S)| + \frac{k!}{n - k + 1}. \tag{35}$$

Since by Lemma 7, $\sum_{x \in X} |q(x, S)| \leq 2^k$, we get

$$\sum_{i \in I} |\lambda^k([\succ_i^k])| + \sum_{j \in S} |q(x_j, S)| \leq 2^k + \frac{k!}{n - k + 1}. \tag{36}$$

Finally, note that since for each k such that $2 < k < n$ $2^k \leq \frac{(2n-2k+1)k!}{n-k+1}$ holds, we have $2^k + \frac{k!}{n-k+1} \leq 2k!$. This, together with (36), implies that (29) holds. Thus, we complete the inductive construction of the desired signed

7 Proof of Theorem 2

We leave it to the reader to show that if an RCF p is dual–pro-con rational, then p satisfies our axioms. Conversely, let p be an RCF that satisfies our axioms.

Step 1: We show that since p satisfies the *modal axioms* and *binary WARP*, there exists $(\succ_1, \succ_2)$ such that:

(1) for each $S \in \Omega$, $p^+(S) = \{max(S, \succ_1), max(S, \succ_2)\}$, and

(2) for each $x, y \in X$, $x \; R \; y$ if and only if $x \succ_1 y$ and $y \succ_2 x$.

It follows from Theorem 2 of Manzini & Mariotti (2018) that if p satisfies the *modal axioms*, then there exists $(\succ_1, \succ_2)$ such that (1) holds. It follows that p^+, as a deterministic choice rule, is a *top-and-the-top* procedure and satisfies the related axioms of Eliaz et al. (2011). Therefore, the algorithm used by Eliaz et al. (2011) also yields a pair $(\succ_1, \succ_2)$ such that (1) holds. Next, we slightly modify this algorithm and argue that the resulting $(\succ_1, \succ_2)$ satisfies both (1) and (2).

Let n be the number of elements in X. We construct inductively two orderings $a_1 \succ_1 \ldots \succ_1 a_n$ and $b_1 \succ_2 \ldots \succ_2 b_n$. For each $i, j \in \{1, \ldots, n\}$, let $A_i \equiv \{a_i, \; \ldots, \; a_j\}$

and $B_j \equiv \{b_i,\ \ldots,\ b_j\}$. Assume that we have deduced the first $m \geq 0$ alternatives of both orderings, then the algorithm finds the $m+1^{st}$ alternative of each ordering by considering the following three cases:

Case 1: If $|p^+(X \backslash A_m)| = 1$, then set a_{m+1} to be $p^+(X \backslash A_m)$.

Case 2: If $|p^+(X \backslash A_m)| = 2$ and $B_m = A_m$, then set a_{m+1} to be one of the two members of $p^+(X \backslash A_m)$ that is stochastically preferred to the other, i.e. we have $a_{m+1}\ R\ b_{m+1}$. If none is stochastically preferred to the other, then set a_{m+1} to be any one of two members of $p^+(X \backslash A_m)$.

Case 3: $|p^+(X \backslash A_m)| = 2$ and $B_m \neq A_m$, then there exists $b_i \in X \backslash A_m$ where $i \leq m$. Let $b_i^* \in X \backslash A_m$ with the minimal index. Note that by construction, $b_i^* \in p^+(X \backslash B_{i^*-1})$ and $X \backslash A_m \subset X \setminus B_{i^*-1}$. By *modal monotonicity*, we have $b_i^* \in p^+(X \setminus A_m)$. Define a_{m+1} to be the other alternative in $p^+(X \setminus A_m)$.

To determine b_{m+1} , apply the above algorithm to $p^+(X \setminus B_m)$. Note that if $|p^+(X \setminus A_m)| = 2$ and $B_m = A_m$, then $p^+(X \setminus B_m) = p^+(X \setminus A_m)$, and applying the above algorithm to either $p^+(X \setminus A_m)$ or $p^+(X \setminus B_m)$ yields the same pair$(a_{m+1},\ b_{m+1})$.

In defining this algorithm, we depart from the Eliaz et al.'s algorithm only in Case 2. That is, we choose set a_{m+1} to be one of the two members of $p^+(X \backslash A_m)$ that is stochastically preferred to the other, whereas in the original one this choice is arbitrary. Therefore, it directly follows from Eliaz et al. (2011) that $(\succ_1, \succ_2)$ satisfies (1). Next, we show that $(\succ_1, \succ_2)$ satisfies (2).

We first show that for each $a, b \in X$ if $a \succ_1 b$ and $b \succ_2 a$, then $a\ R\ b$. Suppose not, and let $V = \{(a, b) : a \succ_1 b \text{ and } b \succ_2 a, \text{ but } \neg(a\ R\ b)\}$. Now, for each $a, b \in V$, there exists $i, j \in \{1, \ldots, m\}$ such that $a = a_i$ and $b = b_j$. First let j be the lowest index such that $(a_i, b_j) \in V$ for some $a_i \in X$. Then, for given b_j, let i be the lowest index such that $(a_i, b_j) \in V$.

Next, by using *binary WARP*, we argue that $a_i\ R\ b_j$. To see this we analyze three cases separately. First, suppose $i = j$, then it directly follows from the construction of $(\succ_1, \succ_2)$ that $a_i\ R\ b_j$. Next, suppose $i < j$, and consider $X \setminus A_i$. We know that $a_i \in p^+(X \setminus A_i)$. Since $b_j \succ_2 a_i$, a_i is determined by Case 2 or Case 3 of the algorithm. It follows that there exists $b_m \neq a_i$ such that $b_m \in p^+(X \setminus A_i)$ and $m \leq i$. Note that $a_i \succ_1 b_m$ and $b_m \succ_2 a_i$. Moreover, since $m \leq i$, by our choice of b_j, $a_i\ R\ b_m$. Now consider the choice set $\{a_i, b_j, b_m\}$. Since $(\succ_1, \succ_2)$ satisfies (1), we get $p^+(\{a_i, b_m, b_j\}) = \{a_i, b_m\}$. Note that we have $a_i \perp b_m$ and $a_i \perp b_j$. Since $a_i \in p^+(\{a_i, b_m, b_j\})$ and $a_i\ R\ b_m$, it follows from *binary WARP* that $a_i\ R\ b_m$. Finally, suppose $i > j$, by similar reasoning, consider $X \setminus B_j$. We know that $b_j \in p^+(X \setminus B_j)$. Since $a_i \succ_1 b_j$, b_j is determined by Case 2 or Case 3 of the algorithm. It follows that there exists $a_m \neq b_j$ such that $a_m \in p^+(X \setminus B_j)$ and $m \leq i$. Note that $a_m \succ_1 b_j$ and $b_j \succ_2 a_i$. Moreover, since $m \leq i$, by our choice of a_i, $a_m\ R\ b_j$. Now consider the choice set $\{a_i, a_m, b_j\}$. Since $(\succ_1, \succ_2)$ satisfies (1), $p^+(\{a_i, a_m, b_j\}) = \{a_i, b_j\}$. Note that we have $b_j \perp a_m$ and $b_j \perp a_i$. Since $b_j \in p^+(\{a_i, a_m, b_j\})$ and $a_m\ R\ b_j$, it follows from *binary WARP* that $a_i\ R\ b_j$.

To complete Step 1, we show that for each $a, b \in X$, if $a\ R\ b$, then $a \succ_1 b$ and $b \succ_2 a$. First, note that since $(\succ_1, \succ_2)$ satisfies (1), $a\ R\ b$ implies $a \perp b$. Therefore, we can not have $b \succ_1 a$ and $b \succ_2 a$, or $a \succ_1 b$ and $a \succ_2 b$. Now, if $b \succ_1 a$ and $a \succ_2 b$, then it follows from our previous observation that $b\ R\ a$. Therefore, we must have $a \succ_1 b$ and $b \succ_2 a$. In the rest of the proof we show that p is pro-con rational w.r.t. $(\succ_1, \succ_2)$.

Step 2: First, we introduce a strengthening of our *expansion* axiom, and then show that p satisfies this axiom. To introduce *strong expansion*, a choice set S is **weakly-isomorphic** to another choice set S', denoted by $S \approx^\pi S'$, if there is a one-to-one mapping π between between $p^+(S)$ and $p^+(S')$ such that for each $x, y \in p^+(S)$

W1. $x\ R\ y$ if and only if $\pi(x)\ R\ \pi(y)$, and

W2. $x \perp S$ if and only if $\pi(x) \perp S'$.

Strong expansion: For each $S, S' \in \Omega$, if $S \approx^\pi S'$, then for each $x \in p^+(S)$, $p(x, S) = p(\pi(x), S')$.

To see that p satisfies *strong expansion*, first note that, by construction, $\approx^\pi$ is transitive, i.e. for each $S, S', S'' \in \Omega$, if $S \approx^\pi S' \approx^\pi S''$, then $S \approx^\pi S''$. Second, since by Step 1 for each $S \in \Omega$, $|p^+(S)| \leq 2|$, it follows from weak isomorphism that each $S \in \Omega$ belongs to one of the following six categories. In the rest, we analyze each of these categories separately, and show that *strong expansion* holds. Since $\approx^\pi$

is transitive, it will follow that p satisfies *strong expansion*. To keep notation simple, for each $S \approx^{\pi} S'$ and $x \in p^{+}(S)$, we generically denote $\pi(x)$ by x'.

C1: $S \approx^{\pi} \{x\}$ for some $x \in X$. Since there is a one-to-one mapping π between S and x, $p^{+}(S) = \{x'\}$ for some $x' \in X$, which directly implies that $p(x', S) = 1 = p(x, \{x\})$.

C2: $S \approx^{\pi} \{x, y\}$ for some distinct $x, y \in X$ such that $x \ R \ y$. It follows that $p^{+}(S) = \{x', y'\}$ for some distinct $x', y' \in S$ such that $x' \ R \ y'$. Since p satisfies *binary invariance*, $p(x', \{x', y'\}) = p(x, \{x, y\})$. Now, since $S \sim \{x', y'\}$ and p satisfies *expansion*, $p(x', \{x', y'\}) = p(x', S)$. It directly follows that $p(x, \{x, y\}) = p(x', S)$.

C3: $S \approx^{\pi} \{x, y, z\}$ for some distinct $x, y, z \in X$ such that $x \ R \ y$, and z is a decoy for x when y is available. It follows that there exist distinct $x', y', z' \in S$ such that $p^{+}(S) = \{x', y'\}$ with $x' \ R \ y'$, and z' is a decoy for x' when y' is available. Since p satisfies *decoy invariance*, $p(x', \{x', y', z'\}) = p(x, \{x, y, z\})$. Now, since $S \sim \{x', y', z'\}$ and p satisfies *expansion*, $p(x', \{x', y', z'\}) = p(x', S)$. It directly follows that $p(x, \{x, y, z\}) = p(x', S)$.

C4: $S \approx^{\pi} \{x, y, w\}$ for some distinct $x, y, w \in X$ such that $x \ R \ y$, and w is a decoy for y when x is available. It follows that there exist distinct $x', y', w' \in S$ such that $p^{+}(S) = \{x', y'\}$ with $x' \ R \ y'$, and w' is a decoy for y' when x' is available. Since p satisfies *decoy invariance*, $p(y', \{x', y', z'\}) = p(y, \{x, y, z\})$. Now, since $S \sim \{x', y', z'\}$ and p satisfies *expansion*, $p(y', \{x', y', z'\}) = p(y', S)$. Since $p^{+}(S) = \{x', y'\}$, it directly follows that $p(x, \{x, y, z\}) = p(x', S)$.

C5: $S \approx^{\pi} \{x, y, z\}$ for some distinct $x, y, z \in X$ such that both x and y dominate z. It follows that there exist distinct $x', y', z' \in S$ such that $p^{+}(S) = \{x', y'\}$ with $x' \ R \ y'$, and both x' and y' dominate z'. Since p satisfies *dominated alternative invariance*, $p(x', \{x', y', z'\}) = p(x, \{x, y, z\})$. Now, since $S \sim \{x', y', z'\}$ and p satisfies *expansion*, $p(x', \{x', y', z'\}) = p(x', S)$. It directly follows that $p(x, \{x, y, z\}) = p(x', S)$.

C6: $S \approx^{\pi} \{x, y, z, w\}$ for some distinct $x, y, z, w \in X$ such that $x \ R \ y$, z is a decoy for x when y is available, and w is a decoy for y when x is available. It follows that there exist distinct $x', y', z', w' \in S$ such that $p^{+}(S) = \{x', y'\}$ with $x' \ R \ y'$, z' is a decoy for x' when y' is available, and w' is a decoy for y' when x' is available. Since the choice set $\{x', y', z'\}$ belongs to C3, $\{x', y', w'\}$ belongs to C4, and $\{x', y'\}$ belongs to C2, we know that $p(x', \{x', y', z'\}) = p(x, \{x, y, z\})$, $p(x', \{x', y', w'\}) = p(x, \{x, y, w\})$, and $p(x', \{x', y'\}) = p(x, \{x, y\})$. Then, it follows from *likelihood ratio invariance* that $p(x', \{x', y', z', w'\}) = p(x, \{x, y, z, w\})$. Now, since $S \sim \{x', y', z', w'\}$ and p satisfies *expansion*, $p(x', \{x', y', z', w'\}) = p(x', S)$. It follows that $p(x, \{x, y, z, w\}) = p(x', S)$.

Step 3: We introduce an auxiliary condition *best-worst expansion* w.r.t. a pair of orderings, and show that p satisfies *best-worst expansion* w.r.t. $(\succ_1, \succ_2)$ obtained in Step 1. For the given $(\succ_1, \succ_2)$, *best-worst expansion* requires that if two choice sets are similar to each other in the sense that the $(\succ_1, \succ_2)$-best alternatives in S can be renamed as to obtain the configuration of the $(\succ_1, \succ_2)$-best alternatives in S' in the best and worst positions, then the choice probabilities should be preserved under this renaming. Formally, a choice set S is **best-worst isomorphic** to another choice set S', denoted by $S \risingdotseq^{\pi} S'$, if there is a one-to-one mapping π between the $(\succ_1, \succ_2)$-best alternatives in S and the $(\succ_1, \succ_2)$-best alternatives S' such that for each $i, j \in \{1, 2\}$ and $x \in max(S, \succ_i)$,

BW1. $x = max(S, \succ_i)$ if and only if $\pi(x) = max(S', \succ_i)$, and

BW2. $x = min(S, \succ_j)$ if and only if $\pi(x) = min(S', \succ_j)$.

Best-worst expansion: For each $S, S' \in \Omega$, if $S \risingdotseq^{\pi} S'$, then for each $i \in \{1, 2\}$ and $x \in max(S, \succ_i)$, $p(x, S) = p(\pi(x), S')$.

Since p satisfies *strong expansion,* to conclude that p satisfies *best-worst expansion*, it is sufficient to show that for each $S, S' \in \Omega$, if $S \risingdotseq^{\pi} S'$, then $S \approx^{\pi} S'$. Suppose $S \risingdotseq^{\pi} S'$, to see that W1 holds, note that for each $S \in \Omega$ and $x, y \in p^{+}(S)$, since $(\succ_1, \succ_2)$ satisfies (2) by Step 1, $x \ R \ y$ if and only if $x \succ_1 y$ and $y \succ_2 x$. Therefore, W1 follows from BW1 and BW2. To see that W2 holds, since $(\succ_1, \succ_2)$ satisfies (1) by Step 1, for each $S \in \Omega$ and $x \in p^{+}(S)$, $x \perp S$ if and only if $x = max(S, \succ_i)$ for some $i \in \{1, 2\}$ and $x = min(S, \succ_j)$ for $j \neq i$. Therefore, W2 follows from BW1 and BW2.

Step 4: We first define the weights on $(\succ_1, \succ_2)$ as to render a pro-con representation for the given p. To define the weights, let us make a key observation. Consider the five types of configurations below that can be obtained by restricting $\succ_1$ and $\succ_2$. To clarify the terminology, for each $S \in \Omega$, S is called **a type** i **choice set** if we obtain

Type 0		Type 1		Type 2		Type 3		Type 4	
$\succ_1$	$\succ_2$	$\succ_1$	$\succ_2$	$\succ_1$	$\succ_2$	$\succ_1$	$\succ_2$	$\succ_1$	$\succ_2$
x	x	x	y	x	y	x	y	x	y
y	y	y	x	y	x	z	x	y	w
				z	z	y	z	w	x

the configuration type i when $\succ_1$ and $\succ_2$ are restricted to S.

First, note that since by Step 1 $(\succ_1, \succ_2)$ satisfies (1), for each type i choice set S_i, if $x = max(S_i, \succ_1)$ and $y = max(S_i, \succ_2)$, then $p(x, S_i) + p(y, S_i) = 1$. Next, note that for each $S \in \Omega$, there exists a type i choice set S_i, for $i \in \{0, \ldots, 4\}$, such that S is best-worst isomorphic to S_i. Then, it follows from *best-worst expansion* that if we define the weights as to obtain $p(x, S_i)$ and $p(y, S_i)$, then by using the same weights we obtain $p(\pi(x), S)$ and $p(\pi(y), S)$. Therefore, to render a pro-con representation for p with respect to $(\succ_1, \succ_2)$, it is sufficient to define the four weights, namely λ_1, λ_2, δ_1, and δ_2, as to generate the choice probabilities for these five types of choice sets.

Note that depending on X and $(\succ_1, \succ_2)$, we may not observe each type of choice set. In what follows, we analyze the problem case by case. First, let us make some primitive observations to rule out the trivial cases. If $X = \{x, y\}$, then the construction is trivial, so we assume that X has at least three alternatives. In the rest of the proof, we assume that there exist distinct $x, y \in X$ with $x \succ_1 y$, and $y \succ_2 x$. If not, then $\succ_1 = \succ_2$, and Step 1 implies that for each $S \in \Omega$, $p(max(S, \succ_1), S) = 1$. So, we can choose the weights in an arbitrary way. If for each $x, y \in X$, $x \perp y$, then this means $\succ_1$ is the inverse of $\succ_2$, and each $S \in \Omega$ is a type 1 choice set. Then, it is sufficient to set the weights such that $p(x, \{x, y\} = \frac{\lambda_1 - \delta_2}{\lambda_2 - \delta_1}$. So, we suppose that there exists $x, y, z \in X$ such that $x \succ_1 y$, $y \succ_2 x$, and $x \; D \; z$ or $y \; D \; z$. For each $S \in \Omega$ that is best-worst isomorphic to a type 0 choice set, the alternative that is $\succ_1$- and $\succ_2$-best is chosen with probability one, irrespective of the weight function. Therefore, we disregard these choice sets in the rest of the proof.

Case 1: Suppose that for each $i \in \{1, 2, 3, 4\}$, there is a type i choice set. This is equivalent to suppose that there exist $x, y, z, w \in X$ such that z is a decoy for x when

y is available, and w is a decoy for y when x is available. Suppose w.l.o.g. that $x \succ_1 y$ and $y \succ_2 x$. Next, we define the weights as to generate p for these choice sets. First, define $\lambda_1 = p(x, \{x, y, z, w\})$ and $\lambda_2 = p(y, \{x, y, z, w\})$. Note that for each $x', y' \in X$, since $(\succ_1, \succ_2)$ satisfy (1), $x' \; D \; y'$ if and only if $x' \succ_1 y'$ and $x' \succ_2 y'$. Since $x \; D \; z$ and $y \; D \; w$, it follows from Step 1 that $\lambda_1 + \lambda_2 = 1$. Next, consider the set $\{x, y, w\}$, and define $\delta_2 = \frac{p(y,\{x,y,w\}) - \lambda_2}{p(y,\{x,y,w\})}$. Since z is a decoy for x when y is available and p satisfies *attraction,* $p(x, \{x, y, z, w\} \geq p(x, \{x, y, z\}$. This, together with our choice of λ_2, implies that $p(y, \{x, y, w\}) - \lambda_2 \geq 0$. Therefore, $\delta_2 \geq 0$, and we get $p(y, \{x, y, w\}) = \frac{\lambda_2}{1-\delta_2}$. To define δ_1, consider the set $\{x, y, z\}$ and define $\delta_1 = \frac{p(x,\{x,y,z\}) - \lambda_1}{p(x,\{x,y,z\})}$. Similarly, *since* p *satisfies attraction,* $\delta_1 \geq 0$ and we get $p(x, \{x, y, z)\} = \frac{\lambda_1}{1-\delta_1}$. Finally, consider the set $\{x, y\}$. It follows from *likelihood ratio invariance* that if we substitute the defined weights for the choice likelihoods except $L(x, \{x, y\})$, then we obtain that $\frac{p(x,\{x,y\})}{p(y,\{x,y\})} = \frac{\lambda_1 - \delta_2}{\lambda_2 - \delta_1}$. Thus, the pro-con model w.r.t. $(\succ_1, \succ_2)$ with the defined weights represent p.

In the rest of this step, we consider cases in which some types of choices sets are absent. This creates additional degree of freedom in defining the weights. Therefore, it is intuitively clear the we can also define the weights in the desired way. Yet, for sake of completeness, we go through these cases.

Case 2: Suppose there is no $x, y, z, w \in X$ such that $x \succ_1 y$ and $y \succ_2 x$, and $x \; D \; z$ and $y \; D \; w$. Equivalently, suppose there is no type 2 choice set, but a type i choice set may exist for $i \in \{1, 3, 4\}$. Next, we define the weights as to generate p for these choice sets. We assumed that there exist $x, y \in X$ such that $x \succ_1 y$ and $y \succ_2 x$, i.e. a type 1 choice set always exists. If there exists $z \in X$ such that z is a decoy for x when y is available, then define $\lambda_1 = p(x, \{x, y, z\})$ and $\lambda_2 = p(y, \{x, y, z\})$. Additionally, if there exist $x', y', w \in X$ such that $x' \succ_1 y'$, $y' \succ_2 x'$, and w is a decoy for y' when x' is available, then define δ_2 as to satisfy $p(y', \{x', y', w\}) = \frac{\lambda_2}{1-\delta_2}$. For given λ_1 and λ_2 such δ_2 is unique. Finally, define δ_1 as to satisfy $p(x, \{x, y\}) = p(x', \{x', y'\}) = \frac{\lambda_1 - \delta_2}{\lambda_2 - \delta_1}$. If there is no such $x', y', w \in X$, then we can freely define both δ_1 and δ_2 as to satisfy $p(x, \{x, y\}) = \frac{\lambda_1 - \delta_2}{\lambda_2 - \delta_1}$.

Case 3: Suppose that Case 1 and Case 2 fail to hold. Since Case 2 fails to hold, there exist $x, y, z, w \in X$ such that $x \succ_1 y$ and $y \succ_2 x$, and $x\ D\ z$ and $y\ D\ w$. Since Case 1 fails to hold, three scenarios can happen: (i) Both x and y dominate z and w, i.e. there exist only type 1 and type 2 choice sets, (ii) w is a decoy for y when x is available, and $y\ D\ z$, i.e. there exist only type 1 and type 4 choice sets, (iii) z is a decoy for x when y is available, and $x\ D\ w$, i.e. there exist only type 1 and type 3 choice sets.

Suppose that scenario (i) holds. First, define $\lambda_1 = p(x, \{x, y, z, w\})$ and $\lambda_2 = p(y, \{x, y, z, w\})$. Then, since there is no alternative that is a decoy for another in the availability of a third one, we can freely define δ_1 and δ_2 as to satisfy $p(x, \{x, y\}) = \frac{\lambda_1 - \delta_2}{\lambda_2 - \delta_1}$. Note that the same reasoning applies irrespective of z and w are distinct or not.

Suppose that scenario (ii) holds, then we follow a construction similar to that of Case 2. First, define $\lambda_1 = p(x, \{x, y, z, w\})$ and $\lambda_2 = p(y, \{x, y, z, w\})$. Define δ_2 as to satisfy $p(y, \{x, y, w\}) = \frac{\lambda_2}{1-\delta_2}$. Next, since z is not a decoy for x when y is available, we can freely define δ_1 as to satisfy $p(x, \{x, y\}) = \frac{\lambda_1 - \delta_2}{\lambda_2 - \delta_1}$. Finally, for scenario (iii), a symmetric construction works. Thus, we can define the weights as to render a pro-con representation for p w.r.t. $(\succ_1, \succ_2)$.

Step 5: To complete the proof, we argue that the weights can be chosen as to guarantee that P1 and P2 hold. Thus, we will conclude that p is dPC-rational. To see that P1 holds, by Step 1, $(\succ_1, \succ_2)$ satisfies (1), i.e. for each $S \in \Omega$, $p^+(S) = \{max(S, \succ_1), max(S, \succ_2)\}$. Now, by contradiction suppose $\lambda_1 \leq \delta_2$ or $\lambda_2 \leq \delta_1$. Suppose w.l.o.g. that $\lambda_1 \leq \delta_2$. Then, for each $x, y \in X$, $p^+(\{x, y\}) = max(\{x, y\}, \succ_2)$. Since (1) holds, this is possible only if $\succ_1 = \succ_2$. But, then we can choose the weights in any arbitrary way.

To see that P2 holds, since the constructed $(\succ_1, \succ_2)$ and the defined weights represent p, if $\lambda_1 - \delta_2 = \lambda_2 - \delta_1$, then for each $x, y \in X$, $p(x, \{x, y\}) = 1/2$. That means for each $x, y \in X$, if $x \perp y$, then $x\ R\ y$ and $y\ R\ x$. Now, (I) suppose there exist $x, y, z, w \in X$ such that z is a decoy for x when y is available, and w is a decoy for y when x is available. Then, since p satisfies *decoy invariance*, $p(x, \{x, y, z\}) = p(y, \{x, y, w\}) = 1/2$. It follows that $\delta_1 = \delta_2$, which implies $\lambda_1 = \lambda_2$. (II) Suppose there exist $x, y, z \in X$ such that both x and y dominate z. Then it follows from *dominated alternative invariance* that $p(x, \{x, y, z\}) = p(y, \{x, y, z\}) = 1/2$. Then, we get $\lambda_1 = \lambda_2$. Finally, if both (I) and (II) fail to hold, then for the remaining cases we can freely define δ_1 or δ_2, and respectively λ_1 or λ_2. weight function λ. This completes the proof.

References

Allison, G. T. (1971), *Essence of decision: Explaining the Cuban missile crisis*, number 327.5 (729.1), Little, Brown and Company,. 4

Barberá, S. & Pattanaik, P. K. (1986), 'Falmagne and the rationalizability of stochastic choices in terms of random orderings', *Econometrica: Journal of the Econometric Society* pp. 707–715. 6, 10

Barberá, S. & Sonnenschein, H. (1978), 'Preference aggregation with randomized social orderings', *Journal of Economic Theory* **18**(2), 244–254. 6

Block, H. & Marschak, J. (1960), *Contributions to Probability and Statistics*, Stanford University Press, Stanford, CA. 6

Bossert, W. & Sprumont, Y. (2013), 'Every choice function is backwards-induction rationalizable', *Econometrica* **81**(6), 2521–2534. 7

Brualdi, R. A. & Ryser, H. J. (1991), *Combinatorial Matrix Theory*, Vol. 39, Cambridge University Press. 11, 29

De Clippel, G. & Eliaz, K. (2012), 'Reason-based choice: A bargaining rationale for the attraction and compromise effects', *Theoretical Economics* **7**(1), 125–162. 15

Eliaz, K., Richter, M. & Rubinstein, A. (2011), 'Choosing the two finalists', *Economic Theory* **46**(2), 211–219. 17, 41, 42

Falmagne, J.-C. (1978), 'A representation theorem for finite random scale systems', *Journal of Mathematical Psychology* **18**(1), 52–72. 6, 10, 13

Fiorini, S. (2004), 'A short proof of a theorem of Falmagne', *Journal of mathematical psychology* **48**(1), 80–82. 11, 33

Fishburn, P. C. (1978), 'Choice probabilities and choice functions', *Journal of Mathematical Psychology* **18**(3), 205–219. 24

Ford Jr, L. R. & Fulkerson, D. R. (2015), *Flows in networks*, Princeton University Press. 6, 10, 11, 29

Franklin, B. (1887), *The Complete Works of Benjamin Franklin: 1772-1775*, Vol. 5, GP Putnam's Sons. 3

Gelb, L. H. & Betts, R. K. (2016), *The irony of Vietnam: The system worked*, Brookings Institution Press. 4

Glass, B. (1988), *The Correspondence of Charles Darwin. Volume 2: 1837-1843*, University of Chicago Press. 3

Heller, I. & Tompkins, C. (1956), 'An extension of a theorem of dantzig's', *Linear inequalities and related systems* **38**, 247–252. 22

Hoffman, A. J. & Kruskal, J. B. (2010), Integral boundary points of convex polyhedra, *in* '50 Years of Integer Programming 1958-2008', Springer, pp. 49–76. 22

Huber, J. & Puto, C. (1983), 'Market boundaries and product choice: Illustrating attraction and substitution effects', *The Journal of Consumer Research* **10(1)**, 31–44. 14

Intriligator, M. D. (1973), 'A probabilistic model of social choice', *The Review of Economic Studies* **40**(4), 553–560. 6

Intriligator, M. D. (1982), 'Probabilistic models of choice', *Mathematical Social Sciences* **2**(2), 157–166. 6

Kalai, G., Rubinstein, A. & Spiegler, R. (2002), 'Rationalizing choice functions by multiple rationales', *Econometrica* **70**(6), 2481–2488. 7

Knoblauch, V. (2016), 'Elections generate all binary relations on infinite sets', *Mathematical Social Sciences* **84**, 105–108. 27

Luce, R. D. & Raiffa, H. (1957), 'Games and decisions: Introduction and critical survey (new york, 1957)', *Chs. vi and xiv* . 20

Manzini, P. & Mariotti, M. (2014), 'Stochastic choice and consideration sets', *Econometrica* **82**(3), 1153–1176. 7

Manzini, P. & Mariotti, M. (2018), 'Dual random utility maximisation', *Journal of Economic Theory* **177**, 162–182. 16, 17, 41

Manzini, P., Mariotti, M. & Tyson, C. J. (2013), 'Two-stage threshold representations', *Theoretical Economics* **8**(3), 875–882. 25

McFadden, D. (1978), 'Modeling the choice of residential location', *Transportation Research Record* (673). 6

McGarvey, D. C. (1953), 'A theorem on the construction of voting paradoxes', *Econometrica: Journal of the Econometric Society* pp. 608–610. 7, 27

Pattanaik, P. K. & Peleg, B. (1986), 'Distribution of power under stochastic social choice rules', *Econometrica: Journal of the Econometric Society* pp. 909–921. 6

Payne, J. W. & Puto, C. (1982), 'Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis', *Journal of Consumer Research* **9(1)**, 90–98. 14

Rieskamp, J., Busemeyer, J. R. & Mellers, B. A. (2006), 'Extending the bounds of rationality: evidence and theories of preferential choice', *Journal of Economic Literature* pp. 631–661. 14

Saito, K. (2017), 'Axiomatizations of the mixed logit model', *mimeo* . 8, 10, 12

Shafir, E. (1993), 'Choosing versus rejecting: Why some options are both better and worse than others', *Memory & cognition* **21**(4), 546–556. 5

Shafir, E., Simonson, I. & Tversky, A. (1993), 'Reason-based choice', *Cognition* **49**(1), 11–36. 3, 4

Stanley, R. P. (1997), *Enumerative Combinatorics. Vol. 1, vol. 49 of Cambridge Studies in Advanced Mathematics*, Cambridge University Press. 33

Stearns, R. (1959), 'The voting problem', *The American Mathematical Monthly* **66**(9), 761–763. 27

Telhami, S. (1990), 'Israeli foreign policy: A static strategy in a changing world', *Middle East Journal* **44**(3), 399–416. 4

Tversky, A. (1972), 'Elimination by aspects: A theory of choice.', *Psychological review* **79**(4), 281. 5