Operations Research

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Every Choice Function Is Pro-Con Rationalizable

Serhat Dogan, Kemal Yildiz

informs
http://pubsonline.informs.org/journal/opre

OPERATIONS RESEARCH
Articles in Advance, pp. 1–14
ISSN 0030-364X (print), ISSN 1526-5463 (online)

Methods

Every Choice Function Is Pro-Con Rationalizable

Serhat Dogan,[a] Kemal Yildiz[a,*]

[a] Department of Economics, Bilkent University, Ankara 06800, Turkey
*Corresponding author

Contact: dserhat@bilkent.edu.tr, https://orcid.org/0000-0002-0848-8130 (SD); kemal.yildiz@bilkent.edu.tr, https://orcid.org/0000-0003-4352-3197 (KY)

Received: August 30, 2021
Revised: January 26, 2022
Accepted: May 2, 2022
Published Online in Articles in Advance: June 22, 2022

Area of Review: Decision Analysis

https://doi.org/10.1287/opre.2022.2312

Abstract. We consider an agent who is endowed with two sets of orderings: pro- 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, the probability that an alternative is chosen equals the difference between the total weights of its pros and cons. We show that every nuance of the rich human choice behavior can be captured via this structured model. Our technique requires a generalization of the Ford-Fulkerson theorem, which may be of independent interest. As an application of our results, we show that every choice rule is plurality-rationalizable.

Funding: K. Yildiz is grateful for the hospitality of New York University, Department of Economics during his visit in the 2017–2018 academic year, and the support from the Scientific and Research Council of Turkey (TUBITAK) [Grant 1059B191601712].

Keywords: choice function • random choice • attraction effect • additivity • integer programming

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" and "charms of music"—and *cons*—such as "fewer conversations with clever people" and "no books"—regarding the potential impact of marriage on his life. With this list of pros and cons, Darwin seems to follow a choice procedure ascribed to Benjamin Franklin. Here we present the choice procedure of Franklin (1887) 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.

Choice models most commonly used in economics are based on maximization of preferences. An alternative mode of choice, which is common for the scholarly work in other social disciplines such as history, law, and political science, is the less formal *reason-based analysis* (Shafir et al. 1993). Reason-based analysis is also commonly used for the analysis of "case studies" in business and law schools. In the vein of Franklin's prudential algebra, first, various arguments that support or oppose an alternative are identified, then the balance of these arguments determines the choice. We formulate and analyze the *pro-con choice model* that connects these two approaches by presenting a reason-based choice model, in which the reasons are formed via a preference-based language.

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-orderings*. We require that an ordering cannot both serve as a pro- and con-ordering. Given an pcM $\langle \succ, \rhd \rangle$, 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. Conversely, 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.

Our central new concept is the following: A *choice function* is *pro-con rational(izable)* if there is an pcM $\langle \succ, \rhd \rangle$ 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.

A *random pro-con model* (RpcM) is a triplet $\langle \succ, \rhd, \lambda \rangle$, where $\succ$ and $\rhd$ stand for the sets of pro-orderings and con-orderings, as before. The weight function λ assigns to each pro-ordering $\succ_i \in \succ$ a value in the $[0,1]$ interval and con-ordering $\rhd_i \in \rhd$ a value in the $[-1,0]$ interval, which we interpret as a measure of the salience of each ordering. In this case, the probability that an alternative is chosen equals the total weight of its pros and cons.

The most familiar stochastic choice model in economics is the *random utility model* (RUM), which assumes that an agent is endowed with a probability measure μ over a set of orderings $\succ$ such that agent 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. Both the RpcM and the RUM are *additive* models, in the sense that the choice probability of an alternative is calculated by 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 $\langle \succ, \rhd, \lambda \rangle$ to make a choice from each choice set. These two features of RUM reflect themselves in its characterization.[1] Despite the similarity between the RpcM and the RUM, in Theorem 2, we show that every random choice function is pro-con rational. Our technique is build on our Theorem 3, which presents an original extension of the seminal result in optimization theory of Ford and Fulkerson (2015). Then, using the construction in Theorem 2's proof together with two key results from the integer-programming literature, in Theorem 1, we show that each (deterministic) choice function is pro-con rational.[2]

The remaining observations in the paper are as follows. In Section 2.3, we observe that our Theorem 1 fails to hold in the context of multivalued choice rules unless we allow multiple appearance of an ordering as a pro- or con-ordering. In Section 2.4, we illustrate that our results facilitate identification of other inclusive choice models, by showing that each choice function is *plurality-rationalizable* as a corollary to our Theorem 1. For the uniqueness of representation, the RpcM has characteristics similar to the RUM, which we present and discuss in Section 3.3.

1.1. Related Literature

Since Arrow and Raynaud (1986), Franklin's moral algebra is well known in the field of multicriteria decision making. Köksalan et al. (2011) reports Franklin's moral algebra as one of the earliest examples of multicriteria decision making. Hammond et al. (1998) proposes the related "even swaps" method as a practical way of making tradeoffs among any set of objectives across a range of alternatives. Our approach diverges from this literature in taking the route of revealed preference analysis. That is, we recover the primitive objectives and the way that they are reconciled from observed choices.

In the deterministic choice literature, previous choice models proposed by Kalai et al. (2002) and Bossert and 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 *backward-induction rationalizable* (Bossert and Sprumont 2013) if there is an extensive-form game such that for each choice set the backward-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. For random choice functions, Manzini and Mariotti (2014) provide an anything-goes result for the *menu-dependent random consideration set rules*, in which an agent keeps a single preference relation and attaches to each alternative a choice-set–specific attention parameter. Then, the agent chooses an alternative with the probability that no more-preferable alternative grabs attention. In contrast to these models, we believe that the pro-con model is more structured and exhibits limited context dependency. An agent following a pro-con model restricts the pro- and con-orderings to the given choice set to make a choice.

It may be 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. As a partial list, one can consider Intriligator (1973), Barberá and Sonnenschein (1978), Pattanaik and 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 and 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 $\langle \{\succ_i\}_{i \in N}, \mu \rangle$. 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.

In a contemporary paper, Saito (2017) offers characterizations of the mixed logit model. Our Theorem 2 can be obtained from a result in this paper, which is proved by using a different approach independently.[3] We discuss the connection at the end of Section 3.2.

2. Deterministic Pro-Con Choice

2.1. 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 (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 $C(S) \in S$. An *ordering,* denoted generically by $\succ_i$ or $\rhd_i$, is a complete, transitive, and antisymmetric binary relation on X.

A *(deterministic) pro-con model (pcM)* is a pair $\langle \succ, \rhd \rangle$, where $\succ = \{\succ_1, \cdots, \succ_m\}$ and $\rhd = \{\rhd_1, \cdots, \rhd_q\}$ are sets of pro- and con-orderings on X. Given an pcM $\langle \succ, \rhd \rangle$, 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. Conversely, 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 require that an ordering cannot both serve as a pro- and con-ordering.[4] Note that, $\succ$ and $\rhd$ are defined as sets of orderings rather than lists or profiles of orderings, therefore each ordering can be used only once as a pro- or con-ordering. 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. 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 1. 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

$$|Pros(x,S)| > |Cons(x,S)|.$$

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 $Cons(x, S)$. We ask for a rather structured representation that corresponds to one-to-one elimination in Franklin's prudential algebra, in that each ordering has a fixed unit weight, instead of a fractional weight. Next, to illustrate how the model works, first, we revisit Luce and Raiffa's dinner example (Luce and Raiffa 1957) by following a pro-con model.

Example 1. 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 $F \succ_2 C \succ_2 S$. As a con-ordering, consider $F \rhd S \rhd C$, which orders the dishes according to their *riskiness*. Cooking fish requires expertise, therefore it is the most risky one and 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. Similarly, we have S is chosen compared with F and F is chosen compared with C. Thus, we also observe a binary choice cycle between the three alternatives.

2.2. Main Result

We show that every choice function is pro-con rational. In the language of mathematical programming, in Theorem 2, we show that the relaxed (convex) problem has a solution. In Appendix A.3, we prove Theorem 1, which translates into finding an integer solution, using the construction in Theorem 2's proof together with two key results from the integer-programming literature, the ones developed by Hoffman and Kruskal (2010) and Heller and Tompkins (1956).

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

Remark 1. 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 in S, it is barely not chosen in the sense $|Pros(y,S)| - |Cons(y,S)| = 0$. It follows from the proof of Theorem 1 that the same anything-goes result holds for this model.

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 elimination continues until an alternative is singled out. Our model is a specific sequential pro-con model in which all the alternatives but the chosen one are eliminated in the first step. Therefore, an all-goes result applies to this sequential and less structured pro-con model.

2.3. Extension to Multivalued Choice

There are instances in which an agent chooses several alternatives from a choice set. For example, consider a school admitting a cohort from a set of applicants or a professor selecting a set of questions out of his archive to prepare an exam. As for the random choice, imagine that we observe the support of the random choice function, but not the frequencies, then the observed choice behavior yields a choice rule.[5]

Thus 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 multivalued. Formally, a *choice rule* $\mathbb{C}:\Omega\rightarrow\Omega$ 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 $\langle \succ, \rhd\rangle$ 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 1 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.[6] 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 random choice function such that for each $S\in\Omega$ and $x\in\mathbb{C}(S)$, $p(x,S) = 1/|\mathbb{C}(S)|$. It follows from Theorem 2 that there is a random pro-con model $\langle \succ, \rhd, \lambda\rangle$ that represents p. Moreover, it follows from the construction in the proof of Theorem 2 that if for each $S\in\Omega$ and $x\in\mathbb{C}(S)$, $p(x, S)$ 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.

2.4. Plurality-Rationalizable Choice Rules

We analyze 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 1, we show that every choice rule is plurality-rationalizable. This observation also follows from a rather general result of Saari (1989) proved using different techniques.[7] Our proof is based on our Theorem 1 and the related discussion in Section 2.3.

Definition 2. 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 1. *Every choice rule is plurality-rationalizable.*

Proof. Let $\mathbb{C}$ be a choice rule. In Section 2.3, by using Theorem 1, 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 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 $SCons(x,S) = \sum_{\{\rhd_i\in Cons(x,S)\}} k_i$, where $Pros(x, S)$ and $Cons(x, S)$ are defined as usual with respect to $\succ$ and $\rhd$. Now, we know that for each $S\in\Omega$ and $x\in S$, $x\in\mathbb{C}(S)$ if and only if $x\in \arg\max_{x\in S}(|SPros(x,S,\succ^*)| - |SCons(x,S,\rhd^*)|)$.

Now, to construct the desired preference profile, let $k = max_{\{\rhd_i\in\rhd^*\}} k_i$ 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 of each ordering $\rhd_i\in\rhd$, and add k_i copies of each ordering $\succ_i\in\succ$. Because 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, 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 $|SPros(y,S)| - |SCons(y,S)|$. Because 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. □

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 1 and the proof of Proposition 1 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. We obtain McGarvey's result, as a corollary to Proposition 1. To see this, if we restrict a choice rule to binary choice sets, then we obtain an asymmetric and complete binary relation. For binary choices, being a plurality winner means being a majority winner, therefore McGarvey's result directly follows.

3. Random Pro-Con Choice

3.1. Model

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.

A *random pro-con model* (RpcM) is a triplet $\langle \succ, \rhd, \lambda \rangle$, where $\succ$ and $\rhd$ stand for the sets of pro- and con-orderings on X as before. 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, that is, $\sum_{\{\succ_i \in \succ\}} \lambda(\succ_i) + \sum_{\{\rhd_i \in \rhd\}} \lambda(\rhd_i) = 1$. 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. We interpret the weight assigned to each pro-ordering or con-ordering as a measure of the strength of that ordering, and define pro-con rationality of an RCF as follows.

Definition 3. An RCF p is *pro-con rational* if there is an RpcM $\langle \succ, \rhd, \lambda \rangle$ such that for each choice set $S \in \Omega$ and $x \in S$,

$$p(x,\ S) = \lambda(Pros(x,S)) + \lambda(Cons(x,S)), \qquad (1)$$

where $\lambda(Pros(x,S))$ and $\lambda(Cons(x,S))$ are the sum of the weights over *Pros(x, S)* and *Cons(x, S)* as defined in Section 2.1.

As the reader would easily notice not every RpcM yields an RCF. For this to be true, for each choice set S and $x \in S$, the expression in (1) should be nonnegative and sum up to one. These additional requirements are imposed on the model by Definition 3. The requirement that the total weighted sum of pro-orderings and con-orderings be unity is 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.

Next, we provide an equivalent, but less structured formulation of random pro-con rationality that always yields an RCF. 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 $\lambda(x,S)$, that is, $\lambda(x,S) = \lambda(Pros(x,S)) + \lambda(Cons(x,S))$. For each choice set $S \in \Omega$, let S^+ be the set of alternatives in S that receives a positive total weight, that is, $S^+ = \{x \in S : \lambda(x,S) > 0\}$. An RCF p is pro-con rational if there is an RpcM $\langle \succ, \rhd, \lambda \rangle$ such that for each choice set $S \in \Omega$ and $x \in S$,

$$p(x,S) = max\left\{0,\ \frac{\lambda(x,S)}{\sum_{\{y \in S^+\}} \lambda(y,S)}\right\}. \qquad (2)$$

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. Our proof of Theorem 2 clarifies this equivalence. To illustrate this alternative formulation of RpcM, we focus on a particular choice problem in which there are only two orderings $(\succ_1, \succ_2)$ that are relevant for choice, such as price and quality, and present an *attraction effect* scenario.[8] In this scenario, when we introduce an asymmetrically dominated alternative, called a *decoy*, the choice probability of the dominating alternative goes up. This choice behavior, known as the *attraction effect*, is incompatible with any RUM.

Example 2 (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 $\langle \succ, \rhd, \lambda \rangle$, in which there is single pair of orderings used both as the pro- and con-orderings, with weights shown in parenthesis. We can interpret this ordering pair as two distinct criteria that order the alternatives.

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

Now, because for both criteria x is better than z, we get $p(x, \{x, z\}) = 1$. Because x and y fail to dominate each other, and y fail to dominate z, we get $p(y, \{x, y\}) = p(y, \{y, z\}) = 1/2$. That is, z is a "decoy" for x when y is available. When only x and y are available, because 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. Thus, our model captures the intuition that the choice probability of an alternative may increase when a *decoy* is added, because this alternative may no longer be the worst one according to a relevant attribute.

3.2. 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 Appendix Section A.2. As a notable technical contribution, we generalize the Ford-Fulkerson theorem (Ford and Fulkerson 2015) from combinatorial matrix theory to prove the result. Next, we state the theorem and present an overview of the proof. Then, we discuss the connection to Saito (2017).

Theorem 2. *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 $\mathcal{P}^1$ 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 and Pattanaik (1986) for the characterization of RUM.

To understand (RS), while moving from the kth step to the $(k + 1)$th step, each $[\succ^k]$ is decomposed into a collection $\{[\succ^{k+1}_j]\}_{j \in J}$ such that for each $[\succ^{k+1}_j]$ there exists an alternative x_j that is not linearly ordered by $[\succ^k]$, but placed at $[\succ^{k+1}_j]$ right on top of the alternatives that are not linearly ordered by $[\succ^k]$. Therefore, the sum of the weights assigned to $\{[\succ^{k+1}_j]\}_{j \in J}$ 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 A.2, which we obtain by using the *Mobius inversion*.[9]

Our proof is based on two intertwined 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 $k = 2$. However, 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 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 $\boldsymbol{R} = [r_1, \ldots, r_m]$ and $\boldsymbol{C} = [c_1, \ldots, c_n]$ be two real-valued vectors such that the sum of $\boldsymbol{R}$ equals to the sum of $\boldsymbol{C}$. Now, for which $\boldsymbol{R}$ and $\boldsymbol{C}$ can we find an $m \times n$ matrix $A = [a_{ij}]$ such that A has row sum vector $\boldsymbol{R}$ and column sum vector $\boldsymbol{C}$, with each entry $a_{ij} \in [-1, 1]$? Ford and Fulkerson (2015) provide a full answer to this question when $\boldsymbol{R}$ and $\boldsymbol{C}$ are positive real valued.[10] However, there are two issues peculiar to our problem. First issue is that the row and column sums can take negative real values. 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 Theorem 3, we provide a generalization of Ford-Fulkerson theorem that paves the way for our proof by solving the two issues.

Theorem 3 (Generalized Ford-Fulkerson Theorem). *Let $\boldsymbol{R} = [r_1, \ldots, r_m]$ and $\boldsymbol{C} = [c_1, \ldots, c_n]$ 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 $\boldsymbol{R}$ and column sum vector $\boldsymbol{C}$,*
ii. *Each entry $a_{ij} \in [-1, 1]$, and*
iii. *For each $j \in \{1, \ldots, n\}$, $\sum_{i=1}^m |a_{ij}| \leq |c_j| + \max\left\{0, \frac{\sum_{i=1}^m |r_i| - \sum_{j=1}^n |c_j|}{n}\right\}$.*

To get an intuition for Theorem 3, it is easy to see that the sum of the absolute values of the rows and columns should be bounded to extend the result to real-valued vectors. Therefore, in Theorem 3, 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 in item (iii) of Theorem 3. This solution turns out be the forwarding looking solution, which solves the second issue.

The 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 A.3, we show that the row and column vectors associated with (RS) and (CS) satisfy the premises of our Theorem 3. This completes the construction of the desired signed weight function.

As noted in Section 1.1, Saito (2017) independently shows that each RCF can be expressed as an affine combination of two random utility functions by using different techniques. Our Theorem 2 can be obtained from this observation.[11] Additionally, his result holds even when the random choice function is not defined on every subset of the alternative set, here we assume that the random choice function is observed for each choice set.

In our case, 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 prove that the weights can be chosen from $[-1,1]$ interval, we need to apply Theorem 3 by following a deliberate induction argument supported by other structural results, such as Lemma A.3. We believe that our technique build on Theorem 3 can be fruitful in solving similar problems in operations research that are linked to choice models.

3.3. 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 kth 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 $\langle \succ, \rhd, \lambda\rangle$ 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.[12] As for the general case, Proposition 2 provides a uniqueness result for the RpcM, which can be thought as the counterpart of Falmagne's result for the RUM. We would like to highlight that this uniqueness result is not applicable for the alternative and less structured formulation of pro-con rationality via Identity (2) presented in Section 3.1.

For a given RpcM $\langle \succ, \rhd, \lambda\rangle$, let for each $S\in\Omega$ and $x\in S$, $\lambda(x=B_k|S,\succ,\rhd)$ be the sum of the weights assigned to the pro- and con-orderings at which x is the kth top-ranked alternative in S. In our next result, we show that for each RCF the sum of the weights assigned to the orderings at which x is the kth top-ranked alternative in S is the same for each pro-con representation of the given RCF. That is, $\lambda(x=B_k|S,\succ,\rhd)$ is fixed for each RpcM $\langle \succ, \rhd, \lambda\rangle$ that represents the given RCF.

Proposition 2. *If $\langle \succ, \rhd, \lambda\rangle$ and $\langle \succ', \rhd', \lambda'\rangle$ are random pro-con representations of the same RCF p, then for each $S\in\Omega$ and $x\in S$,*

$$\lambda(x=B_k|S,\succ,\rhd)=\lambda'(x=B_k|S,\succ',\rhd'). \tag{3}$$

Proof. Let $\langle \succ, \rhd, \lambda\rangle$ and $\langle \succ', \rhd', \lambda'\rangle$ 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 2. 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 $\lambda(x=B_k|S,\succ,\rhd)$ can be expressed in terms of $q(x,\cdot)$, then (3) follows. To see this, let $\langle \succ, \rhd, \lambda\rangle$ be any RpcM that represents p. Next, for each $S\in\Omega, x\in S$, and $k\in\{1,\dots,|S|\}$, 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 $\lambda(x\,|\,S_1,S_2,\succ,\rhd)$ 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. For each such ordering, x is the kth top-ranked alternative in S. Now, it follows that we have

$$\lambda(x=B_k|S,\succ,\rhd)=\sum_{\{(S_1,S_2)\in\mathbb{P}(S,x,k)\}}\lambda(x\,|\,S_1,S_2,\succ,\rhd). \tag{4}$$

For each $T\in\Omega$ such that $S_2\subset T$ and $T\subset X\setminus S_1$, thus, 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

$$\sum_{\mathbb{P}(S,x,k)}\lambda(x\,|\,S_1,S_2,\succ,\rhd)=\sum_{\mathbb{P}(S,x,k)}\sum_{S_2\subset T\subset X\setminus S_1}q(x,T). \tag{5}$$

Finally, if we substitute (4) in (5), then we express $\lambda(x=B_k\,|\,S,\succ,\rhd)$ only in terms of $q(x,\cdot)$, as desired. □

4. Final Remarks

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. Knowing that each choice function is pro-con rational facilitates identification of other permissive 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. Next, we briefly discuss some of these directions and their relevance.

Discrete choice models play a significant role in analyzing operations management problems that are linked to consumer behavior, such as inventory control, demand estimation, and assortment and price optimization.[13] As for the potential use of the pro-con model in similar analysis, we find the contribution of Farias et al. (2013) related. Arguing that parametric models are prone to issues of overfitting and underfitting, they propose a data-driven approach to efficiently estimate a general random utility model over products that generates revenue predictions. Considering the structural similarity between the random pro-con model and the random utility model, we wonder if a similar exercise can be conducted with the random pro-con model successfully. For these application, a major concern for the fruitful use of choice models is their ease of identification. In this vein, to sharpen our uniqueness result (Proposition 2), in our future research, we aim to analyze *sparse* pro-con representations, in which the number of orderings used in the representation is minimized.[14] Another intriguing question is about the pro-con representation of the *monotone* random choice functions, in which the choice probability of an existing alternative does not increase as new alternatives are added to a choice set. We leave this as an open problem.

Acknowledgments
The authors thank Ariel Rubinstein, Vicki Knoblauch, Efe Ok, Yusufcan Masatlioglu, Erkut Ozbay, Ed Green, Kfir Eliaz, Mustafa Celebi Pinar, Miguel Ballester, Ran Spiegler, Burak Pac, Paola Manzini, Marco Mariotti, Herve Moulin, 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, the editor, and the referees for helpful comments and suggestions.

Appendix

A.1. Proof of Theorem 3
We start by proving some lemmas that are critical for proving the theorem. First, we report the original theorem of Ford and Fulkerson (2015).[15] Then, we prove Theorem 3, which offers a generalization of the result to any real-valued row and column vectors.

Theorem A.1 (Ford and Fulkerson 2015). *Let $R = [r_1, \ldots, r_m]$ and $C = [c_1, \ldots, c_n]$ 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 $I \subset \{1,2,\ldots,m\}$ and $J \subset \{1,2,\ldots,n\}$,*

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

Lemma A.1. *Let $R = [r_1, \ldots, r_m]$ and $C = [c_1, \ldots, c_n]$ be positive real-valued vectors with $0 \le r_i \le 1$ and $0 \le c_j \le m$ 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, because for each $i \in \{1,2,\ldots,m\}$, $0 \le r_i \le 1$, we have for each $I \subset \{1,2,\ldots,m\}$, $\sum_{i\in I} r_i \le |I|$. Then, it directly follows that (FF) holds and the conclusion follows from Theorem A.1. □

By using Lemma A.1, we prove Theorem 3 that is formulated and discussed in Section 3.2.

Proof of Theorem 3 (Generalized Ford-Fulkerson Theorem). Because 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 $\sum_{i=1}^m |r_i| \ge \sum_{j=1}^n |c_j|$ 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 $x^- = \min\{x, 0\}$. 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. Because 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 $i \in \{1,\ldots,m_1\}$, $r_i \ge 0$ and for each $i \in \{m_1+1,\ldots,m\}$, $r_i < 0$. Similarly, suppose for each $j \in \{1,\ldots,n_1\}$, $c_j \ge 0$ 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 nonnegative (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 nonnegative (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 $\sum_{i=I} |r_i| \ge \sum_{j\in J} |c_j|$ and let

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

Because $\sum_{i=1}^m |r_i| \ge \sum_{j=1}^n |c_j|$, we have $\Sigma_{R^+} \ge \Sigma_{C^+}$ and $\Sigma_{R^-} \le \Sigma_{C^-}$. Moreover, because 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 δ, we get

$$\sum_{i=1}^m r_i^+ = \sum_{j=1}^n c_j^+ + \delta \text{ and } \sum_{i=1}^m r_i^- = \sum_{j=1}^n c_j^- - \delta. \tag{A.1}$$

Next, consider row-column vector pairs $(R^1, C^+ + \epsilon)$ and $(-R^2, -(C^- - \epsilon))$, where ϵ is the nonnegative n-vector such that for each $j \in \{1,\ldots,n\}$, $\epsilon_j = \delta$. It follows from (A.1) that for both pairs the sum of the rows equals the sum of the columns. Now we apply Lemma A.1 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^- here:

$$
\begin{array}{r|c|l}
 & (c_1^+ + \epsilon_1)\ \ (c_2^+ + \epsilon_2)\ \ (c_3^+ + \epsilon_3)\ \ \cdots\ \ (c_n^+ + \epsilon_n) & \\
\hline
r_1 \geq 0 & & \\
r_2 \geq 0 & A^+ & \\
\vdots & & \\
r_{m_1} \geq 0 & & \\
\hline\hline
 & & r_{m_1+1} < 0 \\
 & A^- & \vdots \\
 & & r_m < 0 \\
\hline
 & (c_1^- - \epsilon_1)\ \ (c_2^- - \epsilon_2)\ \ (c_3^- - \epsilon_3)\ \ \cdots\ \ (c_n^- - \epsilon_n) &
\end{array}
$$

Because 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 $\sum_{i=1}^{m} |a_{ij}|$. By the construction of A^+ and A^-, for each $j \in \{1, \ldots, n\}$,

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

Because 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 $\sum_{i=1}^{m} |r_i| - \sum_{j=1}^{n} |c_j| = \Sigma_{R^+} - \Sigma_{C^+} + \Sigma_{C^-} - \Sigma_{R^-}$. Because the sum of the rows equals the sum of the columns, that is, $\Sigma_{R^+} + \Sigma_{R^-} = \Sigma_{C^+} + \Sigma_{C^-}$, we also have $\Sigma_{R^+} - \Sigma_{C^+} = \Sigma_{C^-} - \Sigma_{R^-}$. This observation, together with (A.2), 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 nonnegative m-vector ϵ such that, the following two conditions hold.

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

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

Step 1: We show that if $\Sigma_{C^+} - \Sigma_{R^+} \leq m - \sum_{i=1}^{m} |r_i|$, then there exists a nonnegative 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, by simply rearranging the terms, we can rewrite (E2) as follows:

$$\sum_{i=1}^{m} \epsilon_i = \Sigma_{C^+} - \Sigma_{R^+}. \quad \text{(A.3)}$$

Because $\Sigma_{C^+} - \Sigma_{R^+} \leq \sum_{i=1}^{m} (1 - |r_i|)$, for each $i \in \{1, \ldots, m\}$, we can choose an ϵ_i such that $0 \leq \epsilon_i \leq 1 - |r_i|$ and (A.3) holds. It directly follows that the associated ϵ vector satisfies (E1) and (E2).

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

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

Because the sum of the rows equals the sum of the columns, that is, $\Sigma_{R^+} + \Sigma_{R^-} = \Sigma_{C^+} + \Sigma_{C^-}$, 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 nonnegative m-vector ϵ that satisfies (E1) and (E2). Now, consider the row-column vector pairs $(R^+ + \epsilon, C^1)$ and $(-(R^- - \epsilon), -C^2)$. Because ϵ satisfies (E1) for each $i \in \{1, \ldots, m\}$, $r_i^+ + \epsilon_i \in [0,1]$ and $r_i^- - \epsilon_i \in [-1, 0]$. Because 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 A.1 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^- here:

$$
\begin{array}{c|c|c|c}
 & c_1\ \ c_2\ \ \cdots\ \ c_{n_1} \geq 0 & & \\
(r_1^+ + \epsilon_1) & & & (r_1^- - \epsilon_1) \\
\hline
(r_2^+ + \epsilon_2) & & & (r_2^- - \epsilon_2) \\
\vdots & A^+ & A^- & \vdots \\
\vdots & & & \vdots \\
(r_m^+ + \epsilon_m) & & & (r_m^- - \epsilon_m) \\
 & & c_{n_1+1} < 0\ \ \cdots\ \ c_n & \\
\hline
\end{array}
$$

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

A.2. Proof of Theorem 2

To prove Theorem 2, 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, that is, 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*.[16]

Lemma A.2. *For each choice set $S \in \Omega$, and alternative $a \in S$,*

$$p(a, S) = \sum_{S \subset T \subset X} q(a, T) \quad \text{(A.4)}$$

if and only if

$$q(a, S) = \sum_{S \subset T \subset X} (-1)^{|T| - |S|} p(a, T). \quad \text{(A.5)}$$

Proof. For each alternative $a \in X$, $p(a,\cdot)$ 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 A.3. *For each choice set* $S \in \Omega$ *with* $|S| = n-k$,

$$\sum_{a\in X} |q(a,S)| \le 2^k. \tag{A.6}$$

Proof. First, (A.6) can be written as follows:

$$\sum_{a\in S} |q(a,S)| + \sum_{b\notin S} |-q(b,S)| \le 2^k. \tag{A.7}$$

For a set of real numbers,$\{x_1, x_2, \dots x_n\}$, to show $\sum_{i=1}^n |x_i| \le 2d$, it suffices to show that for each $I \subset \{1,2,\cdots,n\}$, we have $-d \le \sum_{i\in I} x_i \le d$. Now, as the set of real numbers, consider $\{q(a,S)\}_{a\in X}$. It follows that to show that (A.12) holds, it suffices to show that for each $S_1 \subset S$ and $S_2 \subset X\setminus S$,

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

holds. To see this, first, for each $S_1 \subset S$ and $S_2 \subset X\setminus S$, it follows from Lemma A.2 that for each $a \in S_1$ and for each $b \in S_2$, we have

$$\begin{aligned} q(a,S) &= \sum_{S\subset T\subset X} (-1)^{|T|-|S|} p(a,T) \text{ and} \\ q(b,S) &= \sum_{S\subset T\subset X} (-1)^{|T|-|S|-1} p(b,T). \end{aligned} \tag{A.8}$$

We obtain the second equality from Lemma A.2, because for each $b \notin S$, by definition of $q(b,\ S)$, we have $q(b,S) = q(b, S\cup\{b\})$. Next, 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 (A.8) that

$$\sum_{a\in S_1} q(a,S) - \sum_{b\in S_2} q(b,S) = \sum_{S\subset T\subset X} (-1)^{|T|-|S|} \sum_{a\in S_1\cup S_2} p(a,T). \tag{A.9}$$

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 (A.9) if $|T|-|S|$ is even, and at least -1 if $|T|-|S|$ is odd. Because $|S| = n-k$, for each m with $n-k \le m \le n$, there are $\binom{k}{m-n+k}$ possible choice sets $T \in \Omega$ such that $S \subset T$ and $|T| = m$. Moreover, for each $i \in \{1,\dots,k\}$, 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 (A.9) reaches its maximum (minimum) when the negative (positive) terms are 0 and the positive (negative) terms are $1(-1)$. Thus, we get

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

It follows from the *binomial theorem* that both leftmost and rightmost sums are equal to 2^{k-1}. This, combined with (A.9), implies

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

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

Now, we are ready to complete the proof of Theorem 2. Recall that we assume $|X| = n$. For each $k \in \{1,\dots,n\}$, let $\Omega_k = \{S \in \Omega : |S| > n-k\}$. Note that $\Omega_n = \Omega$ and $\Omega_1 \subset \Omega_2 \subset \dots \subset \Omega_n$. 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 $k \in \{1,\dots,n\}$, let $[\succ^k]$ denote an *equivalence class* at $\mathcal{P}^k$, where $\succ^k$ linearly orders a fixed set of k alternatives in X.

For each $k \in \{1,\dots,n\}$, $S \in \Omega_k$ and $\succ_1, \succ_2 \in \mathcal{P}$, if $\succ_1 \sim_k \succ_2$, then because 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 $\mathcal{P}^k$, that is, 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 $k \in \{1,\dots,n\}$, λ^k represents p_k. In what follows, we inductively show that for each $k \in \{1,\dots,n\}$, 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 $X = \{x_1,\dots,x_n\}$, then $\mathcal{P}^1 = \{[\succ^{x_1}],\cdots,[\succ^{x_n}]\}$, where for each $i \in \{1,\dots,n\}$ 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 $k \in \{1,\dots,n-1\}$, 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 subequivalence 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

$$\lambda^k([\succ^k]) = \sum_{j=1}^{n-k} \lambda^{k+1}([\succ_j^{k+1}]). \tag{A.10}$$

If λ^{k+1} satisfies (A.10), then, because the induction hypothesis implies that λ^k represents p_k, we get for each $S \in \Omega_k$ and $x \in S$, $p(x,\ S) = \lambda^{k+1}(\{[\succ_j] \in \mathcal{P}^{k+1} : x = max(S, \succ_j)\})$.

Next, we show that λ^{k+1} can be constructed such that (A.10) holds, and for each $S \in \Omega_{k+1}\setminus\Omega_k$, 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 $\{[\succ_1^k], \cdots, [\succ_{k!}^k]\}$ 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 (A.5) 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 $x_j \in S$,

$$p(x_j, S) = \sum_{\{[\succ^{k+1}] \in \mathbf{B}^{k+1}(x_j, S)\}} \lambda^{k+1}([\succ^{k+1}]). \tag{A.11}$$

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, that is, $W(\succ, a) = \{x \in X : a \succ x\} \cup \{a\}$. 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 $\mathcal{P}^{k+1}$, that is, $Q(a, S) = \{\succ \in \mathcal{P} : W(\succ, a) = S \cup \{a\}\}$ and $\mathbf{Q}^{k+1}(a, S) = \{[\succ^{k+1}] \in \mathcal{P}^{k+1} : [\succ^{k+1}] \subset Q(a, S)\}$. For each $x_j \in S$, we have $Q(x_j, S) = \cup_{i \in I} [\succ_{ij}^{k+1}]$. Moreover, it directly follows from the definitions of $Q(x_j, \cdot)$ and $B(x_j, \cdot)$ that

$$B(x_j, S) = \underset{S \subset T}{\cup} Q(x_j, T). \tag{A.12}$$

It follows from this observation that the right-hand side of (A.11) can be written as

$$\sum_{S \subset T} \sum_{\{[\succ^{k+1}] \in \mathbf{Q}^{k+1}(x_j, T)\}} \lambda^{k+1}([\succ^{t+1}]). \tag{A.13}$$

i. Because (CS) holds, we have

$$q(x_j, S) = \sum_{\{[\succ^{k+1}] \in \mathbf{Q}^{k+1}(x_j, S)\}} \lambda^{k+1}([\succ^{k+1}]). \tag{A.14}$$

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

$$q(x_j, T) = \sum_{\{[\succ^{k+1}] \in \mathbf{Q}^{k+1}(x_j, T)\}} \lambda^{k+1}([\succ^{k+1}]). \tag{A.15}$$

To see this, recall that by definition of $q(x_j, T)$ (A.5), we have

$$q(x_j, T) = \sum_{T \subset T'} (-1)^{|T'| - |T|} p(x_j, T'). \tag{A.16}$$

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

$$p(x_j, T') = \sum_{\{[\succ^k] \in \mathbf{B}^k(x_j, T')\}} \lambda^k([\succ^k]). \tag{A.17}$$

Next, suppose that we substitute (A.17) into (A.16). Now, consider the set collection $\{B(x_j, T')\}_{\{T \subset T'\}}$. If we apply the *principle of inclusion-exclusion* to this set collection, then we obtain $Q(x_j, T)$. It follows that

$$\sum_{T \subset T'} (-1)^{|T'| - |T|} \sum_{\{[\succ^k] \in \mathbf{B}^k(x_j, T')\}} \lambda^k([\succ^k]) = \sum_{\{[\succ^k] \in \mathbf{Q}^k(x_j, T)\}} \lambda^k([\succ^k]). \tag{A.18}$$

Because (RS) holds, we have

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

Thus, if we combine (A.16)–(A.19), then we obtain that (A.15) holds.

Now, (A.13) combined with (A.14) and (A.15) imply that the right-hand side of (A.11) equals to $\sum_{S \subset T} q(x_j, T)$. Finally, it follows from Lemma A.2 that

$$p(x_j, S) = \sum_{S \subset T} q(x_j, T). \tag{A.20}$$

Thus, we obtain that (A.11) 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 Theorem 3. 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, that is,

$$\sum_{j \in J} q(x_j, S) = \sum_{i \in I} \lambda^k [\succ_i^k]. \tag{A.21}$$

First, if we substitute (A.5) for each $q(x_j, S)$, then we get

$$\sum_{j \in J} q(x_j, S) = 1 + \sum_{j \in J} \sum_{S \subsetneq T} (-1)^{|T| - |S|} p(x_j, T). \tag{A.22}$$

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, that is, $F(x_j) = \{\succ \in \mathcal{P} : max(T, \succ) = x_j$ 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 $x_j \in S$,

$$\sum_{S \subsetneq T} (-1)^{|T| - |S| + 1} p(x_j, T) = \sum_{\{[\succ^k] \in \mathbf{F}(x_j)\}} \lambda^k([\succ^k]). \tag{A.23}$$

To see this, first, because 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 $\{B(x_j,T)\}_{\{S\subsetneq T\}}$. Because $\cup_{\{S\subsetneq T\}}B(x_j,T)=F(x_j)$, it follows from the *principle of inclusion-exclusion* that (A.23) holds. Next, when we substitute (A.23) in (A.22), we obtain

$$\sum_{j\in J}q(x_j,S)=1-\sum_{\{[\succ^k]\in\mathbf{F}(x_j)\}}\lambda^k([\succ^k]). \tag{A.24}$$

Then, because, by the induction hypothesis, λ^k represents p_k, we can replace one with $\sum_{\{[\succ^k]\in\mathcal{P}^k\}}\lambda^k([\succ^k])$. Finally, 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 $\mathcal{P}^k\setminus\cup_{j\in J}\mathbf{F}(x_j)=\{[\succ^k_i]\}_{\{i\in I\}}$. It follows that (A.21) 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 Theorem 3. To apply Theorem 3, 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 $S\in\Omega_{k+1}\setminus\Omega_k$,

$$\sum_{i\in I}|\lambda^k([\succ^k_i])|+\sum_{j\in J}|q(x_j,S)|\leq 2k!. \tag{A.25}$$

To see this, first we will bound the term $\sum_{i\in I}|\lambda^k([\succ^k_i])|$. 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 kth position. Therefore, there are $k-1!$ different equivalence classes in $\mathcal{P}^k$ that rank y at the kth position. Let $I(y)$ be the index set of these equivalence classes. Because $\{I(y)\}_{y\notin S}$ partitions I, we have

$$\sum_{i\in I}|\lambda^k([\succ^k_i])|=\sum_{y\notin S}\sum_{i\in I(y)}|\lambda^k([\succ^k_i])|. \tag{A.26}$$

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

$$\sum_{i\in I(y)}|\lambda^k([\succ^k_i])|=\sum_{[\succ^k_i]\in\mathbf{Q}^k(y,T)}|\lambda^k([\succ^k_i])|. \tag{A.27}$$

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

$$q(y,\ T)=\sum_{[\succ^k_i]\in\mathbf{Q}^k(y,T)}\lambda^k([\succ^k_i]). \tag{A.28}$$

Next, consider the construction of the values $\{\lambda^k([\succ^k_i]\}_{\{i\in I(y)\}}$ 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 $\{[\succ^k_i]\}=\mathbf{Q}^k(y,T)$. Therefore, we directly have $|\lambda^k([\succ^k_i])|=|q(y,T)|$. For $k>2$, we construct λ^k by applying Theorem 3. It follows from (iii) of Theorem 3 that

$$\sum_{[\succ^k_i]\in\mathbf{Q}^k(y,T)}|\lambda^k([\succ^k_i])|\leq|q(y,T)|+\frac{(k-1)!}{n-k+1}. \tag{A.29}$$

Now, if we sum (A.29) over $y\notin S$, we get

$$\sum_{y\notin S}\sum_{[\succ^k_i]\in\mathbf{Q}^k(y,S\cup y)}|\lambda^k([\succ^k])|\leq\left(\sum_{y\notin S}|q(y,S\cup y)|\right)+\frac{k!}{n-k+1}. \tag{A.30}$$

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, because 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 (A.30), we get

$$\sum_{i\in I}|\lambda^k([\succ^k_i])|+\sum_{j\in J}|q(x_j,S)|\leq\sum_{x\in X}|q(x,S)|+\frac{k!}{n-k+1}. \tag{A.31}$$

Because by Lemma A.3, $\sum_{x\in X}|q(x,S)|\leq 2^k$, we get

$$\sum_{i\in I}|\lambda^k([\succ^k_i])|+\sum_{j\in S}|q(x_j,S)|\leq 2^k+\frac{k!}{n-k+1}. \tag{A.32}$$

Finally, because 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 (A.32), implies that (A.25) holds. Thus, we complete the inductive construction of the desired signed weight function.

A.3. Proof of Theorem 1

We prove this result by following the construction used to prove Theorem 2. Therefore, we proceed by induction. Because 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 $k\in\{1,\ldots,n-1\}$, 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 2 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, 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 $b=[\lambda^k([\succ^k_1]),\ldots,\lambda^k([\succ^k_{k!}]),q(x_1,S),\ldots,q(x_{n-k},S)]$ is the column vector of size $k!+(n-k)$. Let Q denote the associated polyhedron, that is, $Q=\{\lambda\in\mathbb{R}^{(n-k)k!}:A\lambda=b\text{ and }-1\leq\lambda\leq 1\}$. A matrix is *totally unimodular* if the determinant of each square submatrix is 0, 1, or –1. The following result directly follows from theorem 2 of Hoffman and Kruskal (2010).

Lemma A.4 (Hoffman and Kruskal 2010). *If matrix A is totally unimodular, then the vertices of Q are integer valued.*

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

Lemma A.5 (Heller and 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 nonzero entries;*
3. *If two nonzero 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 nonzero 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 A.5, 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. Because $A = [a_{ij}]$ 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 $R_1 = [1,\dots,k!]$ consist of the first $k!$ rows and $R_2 = [1,\dots,n-k]$ consist of the the remaining $n-k$ rows of A. For each $i \in R_1$, the ith row A_i is such that $A_i\lambda = \lambda^k([\succ_i^k])$. That is, for each $j \in \{(i-1)k!,\dots,ik!\}$, $a_{ij} = 1$ and the rest of A_i equals zero. For each $i \in R_2$, the ith row A_i is such that $A_i\lambda = q(x_i, S)$. That is, for each $j \in \{i, i+k!,\dots,i+(n-k-1)k!\}$, $a_{ij} = 1$ and the rest of A_i equals zero. To see that 2 and 3 hold, for each $i, i' \in R_1$ and $i, i' \in R_2$, the nonzero entries of A_i and $A_{i'}$ are disjoint. It follows that for each column there can be at most two rows with a value of one: one in R_1 and the other in R_2.

Finally, it follows from the construction in Step 3 of the proof of Theorem 2 that Q is nonempty because there is λ vector with entries taking values in the $[-1,1]$ interval. Because, as shown previously, A is totally unimodular, it directly follows from Lemma A.4 that the vertices of Q are integer valued. Therefore, λ^{k+1} can be constructed such that (RS) and (CS) hold, and each $\lambda_{ij}^{k+1} \in \{-1,0,1\}$.

Endnotes

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

[2] This result does not directly follow from Theorem 2, because 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.

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

[4] We allow both an ordering and its inverse to form pro- or conorderings. For a scenario demonstrating its plausibility, suppose that alternatives are bottles of wine and a relevant criterion is their price. Now, a wine being the cheapest might be a con for its choice, because it might be plonk, or its choice signals stinginess. However, a wine being the most expensive one can also be a con, because it might cost a lot, or choosing the most expensive one might be snobbish.

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

[6] 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 cannot be obtained for every choice rule.

[7] Our initial result was for choice functions.We thank Vicki Knoblauch and an anonymous referee for suggesting the extension to choice rules and bringing the veiled connection to Saari (1989) to our awareness.

[8] Experimental evidence for the attraction effect is first presented by Payne and Puto (1982) and Huber and 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).

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

[10] Brualdi and Ryser (1991) provides a detailed account of similar results.

[11] In the latest version of Saito (2017) (https://saito.sites.caltech.edu/documents/3687/revised_mixed_logit11.pdf), this observation is reported as corollary 6, item (ii). We refer the reader to footnote 22 in the version for the precise argument showing how our Theorem 2 can be obtained from this statement.

[12] This directly follows from the construction used to establish the base of induction in Theorem 2's proof.

[13] For some recent examples, see Alptekinoğlu and Semple (2016) for the connection of choice models to assortment and price optimization and Berbeglia et al. (2022) for a systematic, empirical study of different choice-based demand models and estimation algorithms.

[14] In the high-dimensional statistics literature, the success and tractability of sparse models are demonstrated by several studies following Donoho and Elad (2003) and Candes and Tao (2005), including Farias et al. (2013) from the operations management literature.

[15] This result, as stated in Lemma A.4, but with integrality assumptions on R, C, and A, follows from corollary 1.4.2 in Brualdi and Ryser (1991). They report that Ford and 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.

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

References

Alptekinoğlu A, Semple JH (2016) The exponomial choice model: A new alternative for assortment and price optimization. *Oper. Res.* 64(1):79–93.

Arrow KJ, Raynaud H (1986) *Social Choice and Multicriterion Decision-Making* (MIT Press, Cambridge, MA).

Barberá S, Pattanaik PK (1986) Falmagne and the rationalizability of stochastic choices in terms of random orderings. *Econometrica* 54(3):707–715.

Barberá S, Sonnenschein H (1978) Preference aggregation with randomized social orderings. *J. Econom. Theory* 18(2):244–254.

Berbeglia G, Garassino A, Vulcano G (2022) A comparative empirical study of discrete choice models in retail operations. *Management Sci.* 68(6):4005–4023.

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

Bossert W, Sprumont Y (2013) Every choice function is backward-induction rationalizable. *Econometrica* 81(6):2521–2534.

Brualdi RA, Ryser HJ (1991) *Combinatorial Matrix Theory*, vol. 39 (Cambridge University Press, Cambridge, UK).

Candes E, Tao T (2005) Decoding by linear programming. *IEEE Trans. Inform. Theory* 51(12):4203–4215.

Donoho DL, Elad M (2003) Optimally sparse representation in general (nonorthogonal) dictionaries via l1 minimization. *Proc. National Acad. Sci. USA* 100(5):2197–2202.

Falmagne J-C (1978) A representation theorem for finite random scale systems. *J. Math. Psych.* 18(1):52–72.

Farias VF, Jagabathula S, Shah D (2013) A nonparametric approach to modeling choice with limited data. *Management Sci.* 59(2):305–322.

Fiorini S (2004) A short proof of a theorem of Falmagne. *J. Math. Psych.* 48(1):80–82.

Fishburn PC (1978) Choice probabilities and choice functions. *J. Math. Psych.* 18(3):205–219.

Ford LR Jr, Fulkerson DR (2015) *Flows in Networks* (Princeton University Press, Princeton, NJ).

Franklin B (1887) *The Complete Works of Benjamin Franklin: 1772–1775*, vol. 5 (GP Putnam's Sons).
Hammond JS, Keeney RL, Raiffa H (1998) Even swaps: A rational method for making trade-offs. *Harvard Bus. Rev.* 76:137–150.
Heller I, Tompkins C (1956) An extension of a theorem of Dantzig's. *Linear Inequalities Related Systems* 38:247–252.
Hoffman AJ, Kruskal JB (2010) Integral boundary points of convex polyhedra. *50 Years of Integer Programming 1958–2008* (Springer, Berlin), 49–76.
Huber J, Puto C (1983) Market boundaries and product choice: Illustrating attraction and substitution effects. *J. Consumer Res.* 10(1):31–44.
Intriligator MD (1973) A probabilistic model of social choice. *Rev. Econom. Stud.* 40(4):553–560.
Intriligator MD (1982) Probabilistic models of choice. *Math. Social Sci.* 2(2):157–166.
Kalai G, Rubinstein A, Spiegler R (2002) Rationalizing choice functions by multiple rationales. *Econometrica* 70(6):2481–2488.
Köksalan MM, Wallenius J, Zionts S (2011) *Multiple Criteria Decision Making: From Early History to the 21st Century* (World Scientific).
Luce RD, Raiffa H (1957) *Games and Decisions: Introduction and Critical Survey* (Wiley, New York).
Manzini P, Mariotti M (2014) Stochastic choice and consideration sets. *Econometrica* 82(3):1153–1176.
Manzini P, Mariotti M, Tyson CJ (2013) Two-stage threshold representations. *Theoretical Econom.* 8(3):875–882.
McFadden D (1978) Modeling the choice of residential location. *Transporation Res. Rec.* 673:72–77.
McGarvey DC (1953) A theorem on the construction of voting paradoxes. *Econometrica* 21(4):608–610.
Pattanaik PK, Peleg B (1986) Distribution of power under stochastic social choice rules. *Econometrica.* 54(4):909–921.
Payne JW, Puto C (1982) Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. *J. Consumer Res.* 9(1):90–98.
Rieskamp J, Busemeyer JR, Mellers BA (2006) Extending the bounds of rationality: Evidence and theories of preferential choice. *J. Econom. Literature* 44(3):631–661.
Saari DG (1989) A dictionary for voting paradoxes. *J. Econom. Theory* 48(2):443–475.
Saito K (2017) Axiomatizations of the mixed logit model. Working paper, Caltech. https://authors.library.caltech.edu/99361/.
Shafir E (1993) Choosing vs. rejecting: Why some options are both better and worse than others. *Memory Cognition* 21(4): 546–556.
Shafir E, Simonson I, Tversky A (1993) Reason-based choice. *Cognition* 49(1):11–36.
Stanley RP (1997) *Enumerative Combinatorics*, vol. 1 (Cambridge University Press, Cambridge, UK).

Kemal Yildiz is an assistant professor in the Department of Economics at Bilkent University. He is an economic theorist interested in how individual and institutional choices are made and implemented. His work has been published in *Theoretical Economics*, *Journal of Economic Theory*, *Games and Economic Behavior*, and *Mathematics of Operations Research*.

Serhat Dogan is a PhD graduate from the Department of Economics at Bilkent University. His research interests include choice theory and contest theory. His work has been published in the *Journal of Economic Theory, Games and Economic Behavior*, and *Mathematics of Operations Research*.