A new choice model is presented. An agent is endowed with two sets preferences: pro-preferences and con-preferences. For each choice set, if an alternative is the best (worst) for a pro-preference (con-preference), then this is a pro (con) for choosing that alternative. The alternative that has more pros than cons is chosen from each choice set. Each preference may have a weight reflecting its salience. In this case, each alternative is chosen with a probability proportional to the difference between the weights of its pros and cons. We show that this model accommodates every choice behavior.

JEL Classification Numbers: D01

Keywords: Choice function, random choice, collective-decision making.
# Contents

1 Introduction 3

1.1 Related literature 7

2 Prudential random choice functions 8

2.1 The model 8

2.2 Examples 10

2.3 Main Result 12

2.4 Uniqueness 14

3 Prudential deterministic choice functions 16

3.1 The model 16

3.2 Main result 17

3.3 Plurality-rationalizable choice functions 19

4 Conclusion 20

5 Proof of Theorem 1 21
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, to make up his mind whether to marry or not was a hard decision for Darwin. Just a few months ago, he had scribbled on a piece of paper a carefully considered list of pros such as “constant companion”, “charms of music”, “female chit-chat” and cons such as “may be quarrelling”, “less 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 the choice procedure proposed by Franklin, 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.

In this paper, we formulate and analyze a choice model that we call prudential choice inspired by the Franklin’s prudential algebra. We formulate this model both in

\(^1\)See Glass (1988) for the full list.
\(^2\)In 1772 a man named Joseph Priestley wrote a letter to Benjamin Franklin asking for his advice on a decision he was trying to make. Franklin writes back indicating that he could not tell him what to do, but he could tell him how to make his decision, and suggests his prudential algebra.
the deterministic and stochastic choice setups. In doing so, we extend the Franklin’s prudential algebra as to allow an agent’s choices yield a probability distribution over choice sets with possibly more than two alternatives. Although the deterministic choice framework is a special case of the stochastic one, our formulation of deterministic prudential choice is more restrictive than a direct adaptation of its stochastic counterpart. In our analysis, we observe that prudential choice models provide an inclusive and structured representation for both deterministic and stochastic choice functions.

First we formulate the prudential choice model in the deterministic choice setup. Let \( X \) be a nonempty finite alternative set and any nonempty subset \( S \) be a choice set. A choice function \( C \) singles out an alternative from each choice set. A (deterministic) prudential model (PM) is a pair \( (\succ, \preceq) \) such that \( \succ = \{\succ_1, \cdots, \succ_m\} \) is a collection of pro-preferences\(^3\) and \( \preceq = \{\preceq_1, \cdots, \preceq_q\} \) is a collection of con-preferences. Given an PM \( (\succ, \preceq) \), for each choice set \( S \) and alternative \( x \in S \), if \( x \) is the \( \succ_i \)-best 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 \( \preceq_i \)-worst alternative in \( S \) for some \( \preceq_i \in \preceq \), then we interpret this as a ”con” for choosing \( x \) from \( S \). More formally, let \( \text{Pros}(x, S) \) denote the set of pro-preferences \( (\succ_i \in \succ) \) at which \( x \) is the best alternative in \( S \) and \( \text{Cons}(x, S) \) denote the set of con-preferences \( (\preceq_i \in \preceq) \) at which \( x \) is the worst alternative in \( S \).

Our central new concept is the following: A choice function is prudential if there is an PM \( (\succ, \preceq) \) such that for each choice set \( S \), an alternative \( x \) is chosen from \( S \) if and only if the number of \( \text{Pros}(x, S) \) is greater than the number of \( \text{Cons}(x, S) \).\(^4\)

To see how the model works, let us revisit the Luce and Raiffa’s Dinner (Luce & Raiffa (1957)) in which they choose chicken when the menu consists of steak and chicken only, yet go for the steak when the menu consists of steak (\( S \)), chicken

---

\(^3\)A preference is a complete, transitive, and antisymmetric binary relation on \( X \).

\(^4\)This formulation corresponds to the Franklin’s prudential algebra in which each pro and con item has equal weight. We propose PM as a plausible individual choice model, but also it turns out that a PM can also be viewed as a collective-decision making model based on plurality voting. We present the model in Section 3.3. As a corollary to our Theorem 2, we show that every choice function is plurality-rationalizable.
(C), and fish (F). Consider the pro-preferences \( \succ_1 \) 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-preference consider \( C \succ S \succ F \), which orders the dishes according to their riskiness. Since cooking fish requires expertise, it is the most risky one, and since chicken is the safe option, it is the least risky one. In short, we have risk averse agents who like attractive and healthy food. 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 as in the story of Luce and Raiffa.

Next, we formulate the prudential model in the stochastic choice setup. In this setup, an agent’s repeated choices or a group’s choices are summarized by a random choice function (RCF) \( p \), which assigns to each choice set \( S \), a probability measure over \( S \). For each choice set \( S \) and alternative \( x \), we denote by \( p(x, S) \) the probability that alternative \( x \) is chosen from choice set \( S \). A random prudential model (RPM) is a triplet \( (\succ, \succ, \lambda) \), where \( \succ \) and \( \succ \) stand for pro-preferences and con-preferences as before. The weight function \( \lambda \) assigns to each \( \succ_i \in \succ \) and \( \succ_i \in \succ \), a value in the \((0, 1]\) interval, which we interpret as a measure of the salience of each preference. An RCF \( p \) is prudential if there is an RPM \( (\succ, \succ, \lambda) \) such that for each choice set \( S \) and alternative \( x \in S \),

\[
p(x, S) = \lambda(\text{Pros}(x, S)) - \lambda(\text{Cons}(x, S)),
\]

where \( \lambda(\text{Pros}(x, S)) \) and \( \lambda(\text{Cons}(x, S)) \) are the sum of the weights over \( \text{Pros}(x, S) \) and \( \text{Cons}(x, S) \).\(^5\)

The most familiar stochastic choice model in economics is the random utility model (RUM),\(^6\) which assumes that an agent is endowed with a probability measure

\(^5\)Note that every RPM \( (\succ, \succ, \lambda) \) does not yield an RCF. For an equivalent description of the RPM that yields an RCF, for each choice set \( S \in \Omega \) and \( x \in S \), let \( \lambda(x, S) = \lambda(\text{Pros}(x, S)) - \lambda(\text{Cons}(x, S)) \) and \( S^+ \) be the alternatives in \( S \) with \( \lambda(x, S) > 0 \), then require that \( p(x, S) = \frac{\lambda(x, S)}{\sum_{y \in S^+} \lambda(y, S)} \) if \( \lambda(x, S) > 0 \), and \( p(x, S) = 0 \) otherwise.

\(^6\)See Thurstone (1927), Marschak et al. (1959), Harsanyi (1973), McFadden (1978).
μ over a set of preferences ≻ such that he randomly selects a preference to be maximized from ≻ according to μ. The connection between RUM and RPM is clear, since each RUM (≻, μ) is a RPM in which there is no set of con-preferences. As an alternative model, Tversky (1972) proposes elimination by aspects (EBA), in which an agent views each alternative as a set of attributes and makes his choice by following a probabilistic process that eliminates alternatives based on their attributes.\footnote{Tversky (1972) argues that EBA reflects the choice process followed by agents more precisely than the classical choice models.} In the vein of EBA, if an alternative $x$ is not the worst alternative in choice set $S$ for some con-preference $≿_i$, then this can be interpreted as “$x$ has attribute $i$ in choice set $S$”. Then, each alternative without attribute $i$ in choice set $S$ is eliminated with a probability proportional to the weight of attribute $i$. Thus, RPM offers a choice procedure that both carries the act of probabilistic selection of a preference to be maximized as in the RUM, and elimination of the alternatives as in the EBA.

As for the similarity between the RPM and the RUM, both models are additive in the sense that choice probability of an alternative is calculated by summing up the weights assigned to the preferences. The primitives of both RPM and RUM are structurally invariant in the sense that the decision maker uses the same $(≿, μ)$ and $(≿, ≻_i, λ)$ to make a choice from each choice set. This feature of RUM brings stringency in its identification, which reflects itself in its characterization. Namely, the RCFs that render a random utility representation are the ones that satisfy the Block-Marschak polynomials.\footnote{See Block & Marschak (1960), Falmagne (1978), McFadden (1978), and Barberá & Pattanaik (1986).} On the other hand, despite the similarity between RPM and RUM, in our Theorem 1, we show that every random choice function is prudential. Although RPM offers a structurally invariant additive choice model similar to RUM, this result indicates that RPM is permissive enough to accommodate any choice behavior. By using the construction in the proof of Theorem 1 and two well-known results from integer-programming literature, we show that each (deterministic) choice function
is prudential.\footnote{Note that this result does not directly follow from Theorem 1, since a prudential model is not a direct adaptation of the random prudential model. In that we require each preference to have a fixed unit weight, instead of having fractional weights.}

For a choice model, a theoretical and practical concern is if the primitives of the model can be precisely identified from the observed choices. On this front, RPM has characteristics that are similar to RUM. An RCF may have different random utility representations even with disjoint sets of preferences. However, Falmagne (1978) argues that random utility representation is essentially unique, in the sense that the sum of the probabilities assigned to the preferences at which an alternative $x$ is the $k^{th}$-best in a choice set $S$ is the same for all random utility representations of the given RCF. In the vein of Falmagne's result, we show that for each RCF the difference between the sum of the weights assigned to the pro-preferences at which $x$ is the $k^{th}$-best alternative in $S$ and the sum of the weights assigned to the con-preferences at which $x$ is the $k^{th}$-worst alternative in $S$ is the same for each prudential representation of the given RCF.

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 a preference 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 that subset of alternatives coincides with the choice from that subset. 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 all goes result for their 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
there is no more preferable alternative that grabs his attention. In contrast to these
models, we believe that prudential model is more structured and exhibits limited
context dependency in that an agent following a prudential model only restricts the
pro-preferences and con-preferences to the given choice set to make a choice.

2 Prudential random choice functions

2.1 The model

Given a nonempty finite alternative set $X$, any nonempty subset $S$ is called a choice
set. Let $\Omega$ 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 $S$. A preference, denoted generically by $\succ_i$ or $\triangleright_i$, is a complete,
transitive, and antisymmetric binary relation on $X$.

A random prudential model (RPM) is a triplet $(\succ, \triangleright, \lambda)$, where $\succ = \{\succ_1,
\cdots, \succ_m\}$ and $\triangleright = \{\triangleright_1, \cdots, \triangleright_q\}$ are sets of preferences on $X$, and $\lambda$ is a weight function
such that for each $\succ_i \in \succ$ and $\triangleright_i \in \triangleright$, we have $\lambda(\succ_i) \in (0, 1]$ and $\lambda(\triangleright_i) \in (0, 1]$.

Given an RPM $(\succ, \triangleright, \lambda)$, for each choice set $S$ and alternative $x \in S$, if $x$ is the $\succ_i$-best 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 $\triangleright_i$-worst alternative in $S$ for some $\triangleright_i \in \triangleright$, then we interpret this as a "con" for choosing $x$ from $S$. We interpret the weight assigned to each pro-preference or con-preference as a measure of the salience of that preference. To define when an RCF is called prudential, let $Pros(x, S) = \{\succ_i \in \succ : x = \max(S, \succ_i)\}$ and $Cons(x, S) = \{\triangleright_i \in \triangleright : x = \min(S, \triangleright_i)\}$. 


Definition 1 An RCF \( p \) is **prudential** if there is an RPM \((\succ, \succ^\mathcal{P}, \lambda)\) such that for each choice set \( S \in \Omega \) and \( x \in S \),

\[
p(x, S) = \lambda(\text{Pros}(x, S)) - \lambda(\text{Cons}(x, S)),
\]

where \( \lambda(\text{Pros}(x, S)) \) and \( \lambda(\text{Cons}(x, S)) \) are the sum of the weights over \( \text{Pros}(x, S) \) and \( \text{Cons}(x, S) \).

As the reader would easily notice every RPM \((\succ, \succ^\mathcal{P}, \lambda)\) does not yield an RCF. For this to be true, for each choice set \( S \in \Omega \) and \( x \in S \), (1) should be nonnegative and sum up to one. These additional requirements are imposed on the model by our Definition 1. For an equivalent description of the RPM that always yields an RCF, consider a triplet \((\succ, \succ^\mathcal{P}, \lambda')\) such that the difference between the weighted sum of pro-preferences and con-preferences is one, i.e. \( \sum_{\{\succ_i \in \succ\}} \lambda'(\succ_i) - \sum_{\{\succ_i \in \succ^\mathcal{P}\}} \lambda'({\succ^\mathcal{P}}_i) = 1 \). This restricted weight function \( \lambda' \) acts like a probability measure over the set of preferences that can assign negative values.\(^{10}\) Now, given an RPM \((\succ, \succ^\mathcal{P}, \lambda')\), let for each \( S \in \Omega \) and \( x \in S \), \( \lambda'(x, S) = \lambda'(\text{Pros}(x, S)) - \lambda'(\text{Cons}(x, S)) \), and \( S^+ = \{ x \in S : \lambda'(x, S) > 0 \} \). Next, we provide an equivalent formulation of a prudential RCF.

Definition 2 An RCF \( p \) is **prudential** if there is an RPM \((\succ, \succ^\mathcal{P}, \lambda')\) such that for each choice set \( S \in \Omega \) and \( x \in S \),

\[
p(x, S) = \begin{cases} 
\frac{\lambda'(x, S)}{\sum_{y \in S^+} \lambda'(y, S)} & \text{if } \lambda'(x, S) > 0 \\
0 & \text{if } \lambda'(x, S) \leq 0
\end{cases}
\]

(2)

To make a choice from each choice set \( S \), a prudential agent considers the alternatives with positive \( \lambda'(x, S) \) score, and chooses each alternative from this consideration set with a probability proportional to its weight.

As mentioned in the introduction, 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 \( \mu \) over a set of preferences \( \succ \) such that he randomly selects a preference to be maximized from \( \succ \) according to \( \mu \). It directly

\(^{10}\)In measure theory, it is called charge or signed measure.
follows that each RUM is a RPM without any con-preferences. As an alternative to RUM, Tversky (1972) proposes elimination by aspects (EBA), in which an agent views each alternative as a set of attributes, then at each stage he selects an attribute with probability proportional to its weight and eliminates all the alternatives without the selected attribute. The elimination process continues until all alternatives but one are eliminated. To highlight the connection between EBA and RPM, consider a con-preference $\triangleright_i$, if an alternative $x$ is not the $\triangleright_i$-worst alternative in a choice set $S$, then say that $x$ is acceptable according to $\triangleright_i$ in $S$. Now, we can interpret the statement “$x$ has attribute $i$ in choice set $S$” as $x$ is acceptable according to $\triangleright_i$ in $S$. Thus, for given RPM, each alternative without attribute $i$ in choice set $S$ is eliminated with a probability proportional to the weight of attribute $i$. In line with this interpretation, we illustrate in our Example 2 and Example 3 that each preference in an RPM can be interpreted as an attribute or a relevant criterion for the choice. The attitude of the agent is different to these criteria in that if it corresponds to a pro-preference then the agent seeks maximization; if it corresponds to a con-preference, then the agent is satisfied by the elimination of the worst alternative.

### 2.2 Examples

First, we present an example in which all the preferences have a weight of one. Therefore, the resulting choice is deterministic and illustrates the deterministic counterpart of the RPM.

**Example 1 (Binary choice cycles)** Suppose $X = \{x, y, z\}$ and consider the following RPM $(\succ, \triangleright, \lambda)$. Note that $x$ is chosen from the grand set and when compared to $y$, $y$ is chosen when compared to $z$, but $z$ is chosen when compared to $x$. That is, the given PM generates the choice behavior of an agent who exhibits a binary choice cycle between $a, b, c$, and chooses $a$ from the grand set.

**Example 2 (Similarity Effect)** Suppose $X = \{x_1, x_2, y\}$, where $x_1$ and $x_2$ are similar alternatives, such as recordings of the same Beethoven symphony by different conductors, while $y$ is a distinct alternative, such as a Debussy suite. Suppose between
any pair of the three recordings our classical music aficionado chooses with equal probabilities, and he chooses from the set \( \{x_1, x_2, y\} \) with probabilities 0.25, 0.25, and 0.5 respectively.\(^\text{11}\) Consider the RPM \((\succ, \succsim, \lambda)\) presented below.

\[
\begin{array}{cccc}
(1/4) & (1/4) & (1/2) & (1/2) \\
\succ_1 & \succ_2 & \succ_3 & \succ_4 \\
y & y & x_1 & x_2 \\
x_1 & x_2 & x_2 & x_1 \\
x_2 & x_1 & y & y \\
\end{array}
\]

We choose \((\succ_1, \succsim_1)\) and \((\succ_2, \succsim_2)\) as the same preferences, and assign the same weight. In the story, the composer has a primary importance, whereas the conductor only has a secondary importance. In line with this observation, all the preferences in the given RPM ranks the recordings first according to the composer, then according to the conductor. One can easily verify that the induced RCF generates our classical music aficionado’s choices.

In Example 2 there are two copies of the same alternative which are slightly different. If substitution is not so extreme, then an agent may exhibit a choice pattern incompatible with RUM. In this vein, next example illustrates that when we introduce an asymmetrically dominated alternative, the choice probability of the dominating alternative may even go up. This choice behavior, known as the "attraction effect", is

\[^{11}\text{Debreu (1960) proposes this example to highlight a shortcoming of the Luce rule (Luce (1959)). This phenomena is later referred to as the similarity effect or duplicates effect. See Gul et al. (2014) for a random choice model that accommodates the similarity effect.}\]
incompatible with any RUM.\footnote{Experimental evidence for attraction effect is first presented by \cite{PaynePuto1982} and \cite{HuberPuto1983}. Following their work, evidence for attraction effect has been observed in a wide variety of different settings. For a list of these results one can consult \cite{Rieskamp2006}. On the theory side, \cite{Echenique2013} propose a Luce type of model and \cite{Natenzon2012} proposes a learning model that accommodate attraction effect in the random choice setup.}

**Example 3 (Attraction Effect)** Suppose $X = \{x, y, z\}$, where $x$ and $y$ are two competing alternatives such that none of which clearly dominates the other, and $z$ is another alternative that is dominated by $x$ but not by $y$. To illustrate the attraction effect, we follow the formulation in our Definition 2 with RPM $(\succ, \sqsupset, \lambda)$ in which there is single pair of preferences used both as the pro-preferences and con-preferences. We can interpret this preference pair as two distinct criteria that order the alternatives.

<table>
<thead>
<tr>
<th></th>
<th>$\succ_1$</th>
<th>$\succ_2$</th>
<th>$\sqsupset_1$</th>
<th>$\sqsupset_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x$</td>
<td>$\succ_1$</td>
<td>$\succ_2$</td>
<td>$\sqsupset_1$</td>
<td>$\sqsupset_2$</td>
</tr>
<tr>
<td>$y$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$z$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$x$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$y$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$x$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$y$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Now, since for both criteria $x$ is better than $z$, we get $p(x, z) = [1, 0]$. That is $z$ is a decoy for $x$. Since $x$ and $y$ does not dominate each other, and $y$ does not dominate $z$, we get $p(x, y) = [1/2, 1/2]$ and $p(y, z) = [2/3, 1/3]$. Now note that when only $x$ and $y$ are available, since $x$ is the $\sqsupset_2$-worst alternative, $x$ is eliminated with weight $1/2$. However, when the decoy $z$ is added to the choice set, then this is no longer the case. So, availability of the decoy $z$ increases the choice probability of $x$. Thus, the proposed RPM presents an attraction effect scenario.

### 2.3 Main Result

In our main result, we show that every random choice function is prudential. Although RPM offers a structurally invariant additive choice model similar to RUM,
this result indicates that RPM is permissive enough as to accommodate any choice behavior. We present the proof in Section 5.

**Theorem 1** Every random choice function is prudential.

Here, we present an overview of the proof. For a given RCF \( p \), we first show that there is a signed weight function \( \lambda \), which assigns each preference \( \succ_i \), a value \( \lambda(\succ_i) \in [-1, 1] \) such that \( \lambda \) represents \( p \). That is for each choice set \( S \) and \( x \in S \), \( p(x, S) \) is the sum of the weights over preferences at which \( x \) is the top-ranked alternative. Once we obtain this signed weight function \( \lambda \), let \( \succ \) be the collection of preferences that receive positive weights, and \( \succ \) be the collection of the inverses of the preferences that receive negative weights. Let \( \lambda^\ast \) be the weight function obtained from \( \lambda \) by assigning the absolute value of the weights assigned by \( \lambda \). It directly follows that \( p \) is prudential with respect to the RPM \((\succ, \succ \ast, \lambda^\ast)\). Therefore, to prove the theorem it is sufficient to show that there exists a signed weight function that represents \( p \). 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 preferences that top rank the same alternative, irrespective of whether they are chosen with positive probability or not. That is, for \( X = \{x_1, \ldots, x_n\} \), we have \( \mathcal{P}^1 = \{[\succ x_1], \ldots, [\succ x_n]\} \), where for each \( i \in \{1, \ldots, n\} \) and preference \( \succ_i \in [\succ x_i] \), \( \max(X, \succ_i) = x_i \). Now for each \( x_i \in X \), define \( \lambda^1([\succ x_i]) = p(x_i, X) \). It directly follows that \( \lambda^1 \) is a signed weight function over \( \mathcal{P}^1 \) that represents \( p_1 \). By proceeding inductively, it remains to show that we can construct \( \lambda^{k+1} \) over \( \mathcal{P}^{k+1} \) that represents \( p_{k+1} \). In Step 1 of the proof we show that finding such a \( \lambda^{k+1} \) pins down to finding a solution to the system of equalities described by row sums (RS) and column sums (CS).

\[13\] 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.
the alternatives that are not linearly ordered by \([\succ^k]\). Therefore, sum of the weights assigned to \(\{\succ_j^{k+1}\}\) should be equal to the weight assigned to \([\succ^k]\). This gives us the set of equalities formulated in (RS). To get an intuition for (CS), let \(S\) be the set of alternatives that are not linearly ordered by \([\succ^k]\). Now, we should design \(\lambda^{k+1}\) such that for each \(x_j \in S\), \(p(x_j, S)\) should be equal to the sum of the weights assigned to preferences at which \(x_j\) is the top-ranked alternative in \(S\). The set of equalities formulated in (CS) guarantees this.\footnote{A key observation that is related is our Lemma 6, which we obtain by using the Mobius inversion.}

Next, we observe 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 sum of \(R\) equals to 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 \(C\), and each entry \(a_{ij} \in [-1, 1]\). \textit{Ford Jr \\& Fulkerson (2015)} provides a full answer to this question when \(R\) and \(C\) are positive real valued.\footnote{Brualdi \\& Ryser (1991) provides a detailed account of similar results.} However, a peculiarity of our problem is that the corresponding row and column values can be negative. Indeed, we get nonnegative valued rows and columns only if Block-Marschak polynomials hold, that is the given \(p\) renders a RU representation. In our Lemma 5, we provide an extension of \textit{Ford Jr \\& Fulkerson (2015)}’s result that paves the way for our proof.\footnote{Roughly, for extending the result for real valued vectors, sum of the absolute values of the rows and columns should respect specific bounds.} Then, 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.

### 2.4 Uniqueness

The primitives of the RUM model are \textit{structurally invariant} in the sense that the agent uses the same \(\succ\) and \(\mu\) to make a choice from each choice set. This feature of RUM
brings precision for the identification of the choice behavior. To elaborate on this, although an RCF may have different random utility representation even with disjoint set of preferences, Falmagne (1978) argues that random utility representation is essentially unique. In that, the sum of the probabilities assigned to the preferences at which an alternative \( x \) is the \( k^{th} \)-best in a choice set \( S \) is the same for all random utility representations of the given RCF. Similarly, the primitives of a RPM are structurally invariant in the sense that the agent uses the same triplet \( (\succ,\succeq,\lambda) \) to make a choice from each choice set. In our Proposition 1, we provide a result for RPM that is similar to Falmagne’s result.

For a given RPM \((\succ,\succeq,\lambda)\), let for each \( S \in \Omega \) and \( x \in S \), \( \lambda(x = B_k|S,\succ) \) be the sum of the weights assigned to the pro-preferences at which \( x \) is the \( k^{th} \)-best alternative in \( S \). Similarly, let \( \lambda(x = W_k|S,\succeq) \) be the sum of the weights assigned to the con-preferences at which \( x \) is the \( k^{th} \)-worst alternative in \( S \). In our next result, we show that for each RCF the difference between the the sum of the weights assigned to the pro-preferences at which \( x \) is the \( k^{th} \)-best alternative in \( S \) and the sum of the weights assigned to the con-preferences at which \( x \) is the \( k^{th} \)-worst alternative in \( S \) is the same for each prudential representation of the given RCF. That is, \( \lambda(x = B_k|S,\succ) - \lambda(x = W_k|S,\succeq) \) is fixed for each RPM \((\succ,\succeq,\lambda)\) that represents the given RCF.

**Proposition 1** If \((\succ,\succeq,\lambda)\) and \((\succ',\succeq',\lambda')\) are random prudential representations of the same RCF \( p \), then for each \( S \in \Omega \) and \( x \in S \),

\[
\lambda(x = B_k|S,\succ) - \lambda(x = W_k|S,\succeq) = \lambda'(x = B_k|S,\succ') - \lambda'(x = W_k|S,\succeq')
\]  

**Proof.** Let \((\succ,\succeq,\lambda)\) and \((\succ',\succeq',\lambda')\) be two RPMs that represent the same RCF \( p \). Now, for each choice set \( S \in \Omega \), both \( \lambda \) and \( \lambda' \) 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 \( \lambda \) and \( \lambda' \) generates the same \( q(x,S) \) value. While proving Theorem 1, we have also shown that for each RPM that represents a RCF \( p \), \( q(x,S) \) gives the difference between the sum of the weights of the pro-preferences at which \( x \) is the best alternative in \( S \) and sum of the weights of the con-preferences at which \( x \) is the worst alternative in \( S \). Therefore, if we can show that \( \lambda(x = B_k|S,\succ) \) can be expressed in terms of \( q(x,\cdot) \), then \( 3 \)
To see this, let $(\succ, \triangleright, \lambda)$ be any RPM 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 $\mathcal{P}(S, x, k)$ be the collection of all these partitions. Now, for each fixed $(S_1, S_2) \in \mathcal{P}(S, x, k)$, let $\lambda(x|S_1, S_2, \succ)$ be the sum of the weights of the pro-preferences at which $x$ is the best alternative in $S_2$ and the worst alternative in $S_1$. Note that for each such pro-preference $x$ is the $k^{th}$-best alternative in $S$. Similarly, let $\lambda(x|S_1, S_2, \triangleright)$ be the sum of the weights of the con-preferences at which $x$ is the best alternative in $S_1$ and the worst alternative in $S_2$. Note that for each such con-preference $x$ is the $k^{th}$-worst alternative in $S$. Now, it follows that we have:

$$\lambda(x = B_k|S, \succ) = \sum_{\{(S_1, S_2) \in \mathcal{P}(S, x, k)\}} \lambda(x|S_1, S_2, \succ)$$

(4)

$$\lambda(x = B_k|S, \triangleright) = \sum_{\{(S_1, S_2) \in \mathcal{P}(S, x, k)\}} \lambda(x|S_1, S_2, \triangleright)$$

(5)

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 difference between the sum of the weights of the pro-preferences at which $x$ is the best alternative in $S$ and sum of the weights of the con-preferences at which $x$ is the worst alternative in $S$, it follows that

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

(6)

Finally, if we substitute (4) and (5) in (6), then we express $\lambda(x = B_k|S, \succ) - \lambda(x = B_k|S, \triangleright)$ only in terms of $q(x, \cdot)$ as desired. ■

3 Prudential deterministic choice functions

3.1 The model

A (deterministic) choice function $C$ is a mapping that assigns each choice set $S \in \Omega$ a member of $S$, i.e. $C : \Omega \rightarrow X$ such that $C(S) \in S$. Let $\succ$ and $\triangleright$
stand for two collections of preferences on $X$ as before. A **(deterministic) prudential model (PM)** is a pair $(\succ, \succsim)$ consisting of the pro-preferences and the con-preferences. As before, define $Pros(x, S) = \{ \succ_i \in \succ : x = \max(S, \succ_i) \}$ and $Cons(x, S) = \{ \succsim_i \in \succsim : x = \min(S, \succsim_i) \}$.

**Definition 3** A choice function $C$ is **prudential** if there is an PM $(\succ, \succsim)$ such that for each choice set $S \in \Omega$ and $x \in S$, $C(S) = x$ if and only if $|Pros(x, S)| > |Cons(x, S)|$.

It follows that if an agent is prudential, 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 $Cons(x, S)$.

### 3.2 Main result

In our next result, by using the construction in the proof of Theorem 1 and two well-known results from integer-programming literature, we show that each choice function is prudential. Note that this result does not directly follow from Theorem 1, since our prudential model not a direct adaptation of its random counterpart. In that we require each preference to have a fixed unit weight, instead of having fractional weights.

**Theorem 2** Every choice function is prudential.

**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 $x_i \in X$, $\lambda^1(\langle x_i \rangle) \in \{0, 1\}$. Next, by proceeding inductively, we assume that for any $k \in \{1, \ldots, n - 1\}$, there is a signed weight function $\lambda^k$ taking values $\{-1, 0, 1\}$ over $P^k$ and represents $C_k$. It remains to show that we can construct $\lambda^{k+1}$ taking values $\{-1, 0, 1\}$ over $P^{k+1}$ 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 $\lambda^{k+1}$ 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 an $(k! + (n - k)) \times (n - k)k!$ matrix with entries $a_{ij} \in \{0, 1\}$, and $b = [\lambda^k([> k])_1, \dots, \lambda^k([> k])_{k!}, q(x_1, S), \dots, q(x_{n-k}, S)]$ is the column vector of size $k! + (n - k)$. Let $Q$ denote the associated polyhedron, i.e. $Q = \{\lambda \in \mathbb{R}^{(n-k)k!} : A\lambda = b$ and $-1 \leq \lambda \leq 1\}$. A matrix is totally unimodular if the determinant of every square submatrix is 0, 1 or −1. Following result directly follows from Theorem 2 of Hoffman & Kruskal (2010).

**Lemma 1 (Hoffman & Kruskal (2010))** If the 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. Every entry in $A$ is 0, 1, or $-1$;
2. Every 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 $R_1 = [1, \dots, k!]$ consist of the the first $k!$ rows and $R_2 = [1, \dots, n-k]$ 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 $A_i\lambda = \lambda^k([> k])$. That is, for each $j \in \{(i-1)k!, \dots, ik!\}$, $a_{ij} = 1$ 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 of which is in \( R_1 \) and the other is 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 \( \lambda \) 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, \( \lambda^{k+1} \) can be constructed such that (RS) and (CS) holds, and each \( \lambda^{k+1}_{ij} \in \{-1, 0, 1\} \).

3.3 Plurality-rationalizable choice functions

We propose a collective-decisive making model based on plurality voting. It turns out that this model is closely related to our prudential choice model. To introduce this collective-decisive making model, let \( \succeq^* = [\succeq^*_1, \ldots, \succeq^*_m] \) be a preference profile, which is a list of preferences. In contrast to a collection of preferences, denoted by \( \succ \), a preference \( \succ_i \) can appear more than once in a preference profile \( \succ^* \). For each choice set \( S \in \Omega \) and \( x \in S \), \( x \) is the plurality winner of \( \succ^* \) in \( S \) if for each \( y \in S \setminus \{x\} \), the number of preferences in \( \succ^* \) that top ranks \( x \) in \( S \) is more than the number of preferences in \( \succ^* \) that top ranks \( y \) in \( S \), i.e. for each \( y \in S \setminus \{x\} \), \(|\{\succeq^*_i : x = \max(S, \succeq^*_i)\}| > |\{\succeq^*_i : y = \max(S, \succeq^*_i)\}|. \)

Next, we define plurality-rationalizability, then by using our Theorem 2, we show that every choice function is plurality-rationalizable.

**Definition 4** A choice function \( C \) is **plurality-rationalizable** if there is preference profile \( \succ^* \) such that for each choice set \( S \in \Omega \) and \( x \in S \), \( C(S) = x \) if and only if \( x \) is the plurality winner of \( \succ^* \) in \( S \).

**Proposition 2** Every choice function is plurality-rationalizable.
Proof. Let $C$ be a choice function. It follows from Theorem 2 that $C$ is prudential. Let the PM $(\succ,\succeq)$ be such that for each choice set $S \in \Omega$ and $x \in S$, $C(S) = x$ if and only if $|\text{Pros}(x, S)| > |\text{Cons}(x, S)|$. Now, to construct the desired preference profile, first consider the list of all preferences defined on $X$. Then eliminate any preference that belongs to $\succeq$ and add any preference that belongs to $\succ$. Let $\succ^*$ be the obtained preference profile. Next, consider a choice set $S \in \Omega$ and suppose $C(S) = x$. In what follows we show that $x$ is the plurality winner of $\succ^*$ in $S$. We know that $|\text{Pros}(x, S)| > |\text{Cons}(x, S)|$ and for each $y \in S \setminus \{x\}$, $|\text{Pros}(y, S)| \leq |\text{Cons}(y, S)|$. It follows that for each $y \in S \setminus \{x\}$, $|\text{Pros}(x, S)| - |\text{Cons}(x, S)| > |\text{Pros}(y, S)| - |\text{Cons}(y, S)|$. Now, note that by construction of $\succ^*$, for each $y \in S$ the number of preferences in $\succ^*$ that top ranks $y$ in $S$ equals the number of all preferences that top ranks $y$ in $S$ added to $|\text{Pros}(y, S)| - |\text{Cons}(y, S)|$. Since for each $y \in S$, the number of all preferences that top ranks $y$ in $S$ is fixed, it follows that $x$ is the plurality winner of $\succ^*$ in $S$. ■

Remark 1 One can consider an even more stringent model in which we require an alternative $x$ is chosen from a choice set $S$ if and only if for each $y \in S \setminus \{x\}$, $|\{\succ^*_i : x = \max(S, \succ^*_i)\}| - |\{\succ^*_i : y = \max(S, \succ^*_i)\}| = 1$. We obtain the same "all goes" result with this more demanding model, by following the proof of Proposition 2.

4 Conclusion
5 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)\textsuperscript{17} 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 $R = [r_1, \ldots, r_m]$ 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 $J \subset \{1, 2, \ldots, n\}$,

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

(FF)

Lemma 4 Let $R = [r_1, \ldots, r_m]$ and $C = [c_1, \ldots, c_n]$ be positive real valued vectors with $0 \leq r_i \leq 1$ and $0 \leq c_j \leq m$ such that $\sum_{i=1}^{m} r_i = \sum_{j=1}^{n} c_j$. 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 $i \in \{1, 2, \ldots, m\}$, $0 \leq r_i \leq 1$, we have for each $I \subset \{1, 2, \ldots, m\}$, $\sum_{i \in I} r_i \leq |I|$. Then, it directly follows that (FF) holds. \hfill \blacksquare

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

Lemma 5 Let $R = [r_1, \ldots, r_m]$ and $C = [c_1, \ldots, c_n]$ 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

1. $A$ has row sum vector $R$ and column sum vector $C$,

\textsuperscript{17}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) on page 12. Brualdi & Ryser (1991) reports that Ford Jr & Fulkerson (2015) proves, 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.
ii. each entry $a_{ij} \in [-1, 1]$, and

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

**Proof.** Since $r_i$ and $c_j$ values can be positive or negative, although sum of the rows equals 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 few 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 positive (negative) rows in $R$ and columns in $C$. Since sum of the rows equals 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 $\sum_{i=1}^{m_1} r_i = \sum_{i=1}^{n_1} c_j^+$ and $\sum_{i=m_1+1}^{m} r_i = \sum_{i=1}^{n_1} c_j^-$. Similarly, $\sum_{j=1}^{n_1} c_j = \sum_{j=1}^{n} c_j^+$ and $\sum_{j=n_1+1}^{n} c_j = \sum_{j=1}^{n} c_j^-$. 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 sum of the rows equals sum of the columns, we have $\Sigma_{R^+} - \Sigma_{C^+} = \Sigma_{C^-} - \Sigma_{R^-}$.

Therefore, by the choice of $\epsilon_j$, we get

$$\sum_{i=1}^{m} r_i^+ = \sum_{j=1}^{n} c_j^+ + \epsilon_j \quad \text{and} \quad \sum_{i=1}^{m} r_i^- = \sum_{j=1}^{n} c_j^- - \epsilon_j$$  \hspace{1cm} (7)
Next, consider row-column vector pairs \((R^1, C^+ + \epsilon)\) and \((-R^2, -(C^- - \epsilon))\), where \(\epsilon\) is the non-negative \(n\)-vector such that each \(\epsilon_j\) is as defined above. It follows from (7) that for both pairs sum of the rows equals the sum of the columns. Now we apply Lemma 4 to 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.

\[
\begin{array}{cccc}
(c_1^+ + \epsilon_1) & (c_2^+ + \epsilon_2) & (c_3^+ + \epsilon_3) & \cdots & (c_n^+ + \epsilon_n) \\
\hline
r_1 \geq 0 & A^+ & r_{m_1+1} < 0 \\
r_2 \geq 0 & \vdots & \vdots & \vdots \\
r_{m_1} \geq 0 \end{array}
\]

\[
\begin{array}{cccc}
(c_1^- - \epsilon_1) & (c_2^- - \epsilon_2) & (c_3^- - \epsilon_3) & \cdots & (c_n^- - \epsilon_n) \\
\hline
\end{array}
\]

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 \(j \in \{1, \ldots, n\}\),

\[
\sum_{i=1}^{m} |a_{ij}| = c_j^+ + \epsilon_j + (-c_j^- + \epsilon_j) = |c_j| + 2\epsilon_j = |c_j| + \frac{2\sum R^+ - \sum C^+}{n} \tag{8}
\]

Since for each for each \(j \in \{1, \ldots, n\}\), \(c_j = c_j^+ + c_j^-\) with either \(c_j^+ = 0\) 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 sum of the rows equals 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 (8) 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 \(\epsilon\) 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 \( \epsilon \) 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:

\[
\sum_{i=1}^{m} \epsilon_i = \Sigma_{C^+} - \Sigma_{R^+}
\]

(9)

Since \( \Sigma_{C^+} - \Sigma_{R^+} \leq \sum_{i=1}^{m} (1 - |r_i|) \), for each \( i \in \{1, \ldots, m\} \), we can choose an \( \epsilon_i \) such that \( 0 \leq \epsilon_i \leq 1 - |r_i| \) and (9) holds. It directly follows that the associated \( \epsilon \) 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 sum of the rows equals 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 \( \epsilon \) that satisfies (E1) and (E2). Now, consider the row-column vector pairs \( (R^+ + \epsilon, C^1) \) and \( (R^- - \epsilon, -C^2) \). Since \( \epsilon \) 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 \( \epsilon \) satisfies (E2), for both of the row-column vector pairs, sum of the rows equals 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.
$c_1 \quad c_2 \quad \cdots \quad c_{n_1} \geq 0$

\[
\begin{array}{c}
(r_1^+ + \epsilon_1) \\
(r_2^+ + \epsilon_2) \\
\vdots \\
(r_m^+ + \epsilon_m) \\
c_{n_1+1} < 0 \quad \cdots \quad c_n
\end{array}
\]

$A^+$

$A^-$

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 preferences on $X$. First, we show that there is a \textbf{signed weight function} $\lambda : \mathcal{P} \rightarrow [-1, 1]$ 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 $\lambda$ can assign negative weights to preferences. Once we obtain this signed weight function $\lambda$, let $\succ$ be the collection of preferences that receive positive weights, and let $\succ'$ be the collection of preferences that receive negative weights. Let $\mathcal{D}'$ be the collection of the inverse of the preferences in $\succ'$. Finally, let $\lambda^*$ be the weight function obtained from $\lambda$ by assigning the absolute value of the weights assigned by $\lambda$. It directly follows that $p$ is prudential with respect to the RPM $(\succ, \mathcal{D}', \lambda^*)$. We first introduce some notation and present crucial observations to construct the desired signed weight function $\lambda$.

Let $p$ be a given RCF and Let $q : X \times \Omega \rightarrow \mathbb{R}$ 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 Möbius inversion.\textsuperscript{18}

\textsuperscript{18}See Stanley (1997) Section 3.7. See also Fiorini (2004) who makes the same observation.
Lemma 6 For each choice set $S \in \Omega$, and alternative $a \in S$,

$$p(a, S) = \sum_{S \subseteq T \subseteq X} q(a, T)$$ (10)

if and only if

$$q(a, S) = \sum_{S \subseteq T \subseteq X} (-1)^{|T| - |S|} p(a, T)$$ (11)

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 Möbius inversion, we get the conclusion. ■

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

$$\sum_{a \in X} |q(a, S)| \leq 2^k$$ (12)

Proof. First, note that (12) can be written as follows:

$$\sum_{a \in S} |q(a, S)| + \sum_{b \notin S} | - q(b, S)| \leq 2^k$$ (13)

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 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 (13) 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

$$q(a, S) = \sum_{S \subseteq T \subseteq X} (-1)^{|T| - |S|} p(a, T) \quad \text{and} \quad q(b, S) = \sum_{S \subseteq T \subseteq X} (-1)^{|T| - |S| - 1} p(b, T)$$ (14)

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 (14) that

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

(15)

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 (15), 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 $\binom{m}{n-k}$ 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 (15) reaches its maximum (minimum) when the negative (positive) terms are 0 and positive (negative) ones are $1(-1)$. Thus, we get

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

It follows from the binomial theorem that both leftmost and rightmost sums are equal to $2^{k-1}$. This combined with (15) 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 $\sum_{a \in X} |q(a, S)| \leq 2^k$. \hfill \Box

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 preferences $\succ_1, \succ_2 \in \mathcal{P}$, $\succ_1$ 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 preferences, such that each set (equivalence class) contains preferences 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 \( \Omega_k \), Similarly, if \( \lambda \) is a signed weight function over \( \mathcal{P} \), then let \( \lambda^k \) be the restriction of \( \lambda \) 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 \( \lambda \) represents \( p \) if and only if for each \( k \in \{1, \ldots, n\} \), \( \lambda^k \) represents \( p_k \). In what follows, we inductively show that for each \( k \in \{1, \ldots, n\} \), there is a signed weight function \( \lambda^k \) over \( \mathcal{P}^k \) which represents \( p_k \). For \( k = n \) we obtain the desired \( \lambda \).

For \( k = 1 \), \( \Omega_1 = \{X\} \) and \( \mathcal{P}^1 \) consists of \( n \)-many equivalence classes such that each class contains all the preferences that top rank the same alternative, irrespective of whether they are chosen with positive probability or not. That is, if \( X = \{x_1, \ldots, x_n\} \), then we have \( \mathcal{P}^1 = \{[\succ^{x_1}], \ldots, [\succ^{x_n}]\} \), where for each \( i \in \{1, \ldots, n\} \) and preference \( \succ_i \in [\succ^{x_i}] \), \( \max(X, \succ_i) = x_i \). Now for each \( x_i \in X \), define \( \lambda^1([\succ^{x_i}]) = p(x_i, X) \). It directly follows that \( \lambda^1 \) 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 preferences that top rank the same two alternatives. Now for each \([\succ^2] \in \mathcal{P}^2\) such that \( x_{i_1} \) is the first-ranked alternative and \( x_{i_2} \) is the second-ranked alternative, define \( \lambda^2([\succ^2]) = p(x_{i_2}, X \setminus \{x_{i_1}\}) - p(x_{i_2}, X) \). It directly follows that \( \lambda^2 \) 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 \( \lambda^k \) over \( \mathcal{P}^k \) that represents \( p_k \). Next, we show that we can construct \( \lambda^{k+1} \) 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 \( \{[\succ^{k+1}_1], \ldots, [\succ^{k+1}_{n-k}]\} \subset \mathcal{P}^{k+1} \). Given \( \lambda^k \), we require \( \lambda^{k+1} \) satisfy for each \([\succ^k] \in \mathcal{P}^k\),

\[
\lambda^k([\succ^k]) = \sum_{j=1}^{n-k} \lambda^{k+1}([\succ^{k+1}_j]) \tag{16}
\]

If \( \lambda^{k+1} \) satisfies (16), then since \( \lambda^k \) represents \( p_k \) by the induction hypothesis, we get for each \( S \in \Omega_k \) and \( x \in S \), \( p(x, S) = \lambda^{k+1}([\{\succ_j \in \mathcal{P}^{k+1} : x = \max(S, \succ_j)\}] \).

28
Next, we show that $\lambda^{k+1}$ can be constructed such that (16) holds, and for each $S \in \Omega_{k+1}\setminus\Omega_k$, $\lambda^{k+1}$ represents $p_{k+1}(S)$. To see this, pick any $S \in \Omega_{k+1}\setminus\Omega_k$. It follows that $|S| = n - k$. Let $S = \{x_1, \ldots, x_{n-k}\}$ and $X \setminus S = \{y_1, y_2, \ldots, y_k\}$. 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 $\{[\succ^k_1], \ldots, [\succ^k_{k!}]\}$ be the collection of all such classes. Each preference 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 $\succ^k_{ij}$ linearly orders $X \setminus S$ as in $\succ^k_i$ and ranks $x_j$ in the $k + 1$th position. Consider the associated equivalence class $[\succ^k_{ij+1}]$. Next, we specify $\lambda^{k+1}([\succ^k_{ij+1}])$, the signed weight of $[\succ^k_{ij+1}]$, such that the resulting $\lambda^{k+1}$ represents $p_{k+1}$. To see this we proceed in two steps.

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

$$\sum_{j \in J} \lambda^{k+1}_{ij} = \lambda^k([\succ^k_i]) \quad \text{(RS)}$$

$$\sum_{i \in I} \lambda^{k+1}_{ij} = q(x_j, S) \quad \text{(CS)}$$

, then $\lambda^{k+1}$ represents $p_{k+1}(S)$. For each $S \in \Omega$ and $x_j \in S$, $q(x_j, S)$ is as defined in (11) by using the given RCF $p$.

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

$$p(x_j, S) = \sum_{[\succ^{k+1}] \in B^{k+1}(x_j, S)} \lambda^{k+1}([\succ^{k+1}]) \quad \text{(17)}$$

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 preferences such that $W(\succ, a)$ is exactly $S \cup \{a\}$ and for each $k \in \mathbb{N}$ such that $n - k \leq |S|$, $Q^{k+1}(a, S)$ 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 $Q^{k+1}(a, S) = \{[\succ^{k+1}] \in \mathcal{P}^{k+1} : [\succ^{k+1}] \subset Q(a, S)\}$. Note that, for each $x_j \in S$, we have $Q(x_j, S) = \bigcup_{i \in I} [\succ^{k+1}]$. Moreover, it directly follows from the definitions of $Q(x_j, \cdot)$ and $B(x_j, \cdot)$ that

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

(18)

It follows from this observation that the right-hand side of (17) can be written as

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

(19)

i. Since (CS) holds, we have

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

(20)

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

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

(22)

Since by the induction hypothesis, $\lambda^k$ represents $p_k$, we have

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

(23)

Next suppose that we substitute (23) into (22). 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

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

(24)
Since (RS) holds, we have
\[
\sum_{\{x \succ^k \in Q(x, T)\}} \lambda^k([x \succ^k]) = \sum_{\{x \succ^{k+1} \in Q^{k+1}(x, T)\}} \lambda^{k+1}([x \succ^{k+1}])
\] (25)
Thus, if we combine (22)-(25), then we obtain that (21) holds.

Now, (19) combined with (20) and (21) imply that the right-hand side of (17) equals to \(\sum_{S \subseteq T} q(x, T)\). Finally, it follows from Lemma 6 that
\[
p(x, S) = \sum_{S \subseteq T} q(x, T)
\] (26)
Thus, we obtain that (17) holds.

In what follows we show that for each \(S \in \Omega_{k+1} \setminus \Omega_k\), 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 \(r_i = \lambda^k([x \succ^k])\) and for each \(j \in J\) let \(c_j = q(x, S)\). 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.
\[
\sum_{j \in J} q(x, S) = \sum_{i \in I} \lambda^k[x \succ^k]
\] (27)
First, if we substitute (11) for each \(q(x, S)\), then we get
\[
\sum_{j \in J} q(x, S) = 1 - \sum_{j \in J} \sum_{S \subseteq T} (-1)^{|T| - |S|} p(x, T)
\] (28)
Now, let \(F(x)\) be the collection of preferences \(\succ\) such that there exists \(T \in \Omega\) such that \(S \subseteq T\) and \(x\) is the \(\succ\)-best alternative in \(T\), i.e. \(F(x) = \{\succ \in P : \max(T, \succ) = x \text{ for some } S \subseteq T\}\). For each \(k \in \mathbb{N}\) such that \(n - k \leq |S|\), let \(F(x)\) be the set of associated equivalence classes in \(P^k\). Next we show that for each \(x \in S\),
\[
\sum_{S \subseteq T} (-1)^{|T| - |S|} p(x, T) = \sum_{\{x \succ^k \in F(x)\}} \lambda^k([x \succ^k])
\] (29)
To see this, first, since by the induction hypothesis, \(\lambda^k\) represents \(p_k\), we can replace each \(p(x, T)\) with \(\sum_{\{x \succ^k \in F(x)\}} \lambda^k([x \succ^k])\). Next, consider the set collection
\{B(x_j, T)\}_{S \subseteq T}$. Note that if we apply the principle of inclusion-exclusion to this set collection, then we obtain $F(x_j)$. It follows that (29) holds.

Next, substitute (29) in (28). Then, since, by the induction hypothesis, $\lambda^k$ 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 $[\succ^k] \notin \cup_{j \in J} F(x_j)$ if and only if $\{\succ^k\} \cap S = \emptyset$. This means $\mathcal{P}^k \setminus \cup_{j \in J} F(x_j) = \{\succ^k\}_{i \in I}$. Then, it directly follows that (27) 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 $S \in \Omega_{k+1} \setminus \Omega_k$

$$
\sum_{i \in I} |\lambda^k([\succ^k_i])| + \sum_{j \in J} |g(x_j, S)| \leq 2k! \quad (30)
$$

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 $k$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

$$
\sum_{i \in I} |\lambda^k([\succ^k_i])| = \sum_{y \notin S} \sum_{i \in I(y)} |\lambda^k([\succ^k_i])| \quad (31)
$$

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

$$
\sum_{i \in I(y)} |\lambda^k([\succ^k_i])| = \sum_{[\succ^k_i] \in \mathcal{Q}^k(y, T)} |\lambda^k([\succ^k_i])| \quad (32)
$$

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

$$
q(y, T) = \sum_{[\succ^k_i] \in \mathcal{Q}^k(y, T)} \lambda^k([\succ^k_i]) \quad (33)
$$

Next, consider the construction of the values $\{\lambda^k([\succ^k_i])\}_{i \in I(y)}$ values 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 \( \{[\succ k_i]\} = Q^k(y, T) \). Therefore, we directly have
\[ |\lambda^k([\succ k_i])| = |q(y, T)|. \]
For \( k > 2 \), we construct \( \lambda^k \) by applying Lemma 5. It follows from iii of Lemma 5 that
\[ \sum_{[\succ k_i] \in Q^k(y, T)} |\lambda^k([\succ k_i])| \leq |q(y, T)| + \frac{(k - 1)!}{n - k + 1} \quad (34) \]
Now if we sum (34) over \( y \notin S \), we get
\[ \sum_{y \notin S} \sum_{[\succ k_i] \in Q^k(y, S \cup y)} |\lambda^k([\succ k_i])| \leq \left( \sum_{y \notin S} |q(y, S \cup y)| \right) + \frac{k!}{n - k + 1} \quad (35) \]
Recall that by definition we have \( Q^k(y, S \cup y) = Q^k(y, S) \) 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 (35), then we get
\[ \sum_{i \in I} |\lambda^k([\succ k_i])| + \sum_{j \in J} |q(x_j, S)| \leq \sum_{x \in X} |q(x, S)| + \frac{k!}{n - k + 1} \quad (36) \]
Since by Lemma 7, \( \sum_{x \in X} |q(x, S)| \leq 2^k \), we get
\[ \sum_{i \in I} |\lambda^k([\succ k_i])| + \sum_{j \in J} |q(x_j, S)| \leq 2^k + \frac{k!}{n - k + 1} \quad (37) \]
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 (37) implies that (30) holds. Thus, we complete the inductive construction of the desired signed weight function \( \lambda \). This completes the proof.
References


Falmagne, J.-C. (1978), ‘A representation theorem for finite random scale systems’, *Journal of Mathematical Psychology* 18(1), 52–72. 6, 7, 13, 15


Marschak, J. et al. (1959), Binary choice constraints on random utility indicators, Technical report, Cowles Foundation for Research in Economics, Yale University. 5

McFadden, D. (1978), ‘Modeling the choice of residential location’, *Transportation Research Record* (673). 5, 6

Natenzon, P. (2012), Random choice and learning, in ‘34º Meeting of the Brazilian Econometric Society’. 12


Tversky, A. (1972), ‘Elimination by aspects: A theory of choice.’, *Psychological review* 79(4), 281. 6, 10