OPTIMAL RATIONALIZATION OF CHOICE AND THE MEASUREMENT OF RATIONALITY

JOSE APESTEGUIA[†] AND MIGUEL A. BALLESTER[‡]

Whenever IIA does not hold, any single rationale that could be used to explain individual choices would inevitably lead to mistakes in the rationalization of choice. However, using a rationale or another to explain choices yet matters. Different rationales explain different choices, and lead to mistakes of a very different nature. Obviously, it is of our interest to find the rationale that *best explains* choice behavior.

Suppose that a decision maker (hereafter DM) has cardinal utilities over a collection of alternatives X. Utilities are derived from a random process, where the utility of each alternative follows a cumulative distribution function (cdf) F. Given the probability distribution of utility values, a rationale immediately generates a collection of mistakes (choice problems in which the chosen element does not correspond with the top element within the set). Moreover, a magnitude of the mistake can be derived from the pair rationale-choice problem. Basically, we may value the importance of a mistake endogenously, by the expected difference in the utility values derived from actual choice behavior, and the top element of the set.

More formally, let $X = \{1, \ldots, n\}$ be the set of alternatives with typical element denoted by $i \in X$. $\mathcal{X}$ is the collection of all subsets of X. c denotes a choice function on $\mathcal{X}$ that assigns to every $A \in \mathcal{X}$, a unique element $c(A) \in A$. A binary relation $P \subseteq X \times X$ is a linear order if it is antisymmetric, transitive, and complete. $P(A)$ is the maximal element in $A \in \mathcal{X}$ according to P. Denote by $\mathcal{P}$ the space of all possible linear orders on $X \times X$. For a linear order P define $P_{(i)} = \{x : |\{y : xPy\} \cup \{x\}| = i\}$. That is, $P_{(i)}$ is that alternative that has $i-1$ alternatives below in the linear order P. Then, denote by $e_i(P, c)$ the number of sets $A \in \mathcal{X}$ such that $c(A) = P_{(i)}$ and let $e(P, c) = (e_1(P, c), \ldots, e_n(P, c))$. We write $e(P, c) \trianglelefteq e(P', c)$ whenever $e(P', c)$ majorizes $e(P, c)$.

Given a linear order $P \in \mathcal{P}$ and the cdf F, we may evaluate the expected utility difference between the top element of a set and the chosen element according to c.

Date: March, 2008.

[†]Universitat Pompeu Fabra. http://www.econ.upf.es/~apesteguia/.

[‡]Universitat Autonoma de Barcelona. http://pareto.uab.es/mballester/.

Consider the collection of random variables $U_{(1)} \leq U_{(2)} \leq \cdots \leq U_{(n)}$, where $U_{(i)}$ denotes the i-th order statistic, representing the utility value of that alternative that is better than $i-1$ alternatives and worse than $n-i$ alternatives. The expected value of the i-th order statistic is denoted by $\mathbb{E}[U_{(i)}]$. A linear order P assigns to every alternative in X an order statistic. Finally, the problem arises on which is the linear order P that best explains choices, i.e., with the lowest expected error:

$$r(c) = \max_{P \in \mathcal{P}} r(P, c) = \max_{P \in \mathcal{P}} \mathbb{E}[\sum_{A \subseteq X} u(c(A)) - u(P(A))]$$

As a first result, we provide the following theorem:

Theorem 0.1. *For every choice function c:*

- $e(P, c) \trianglelefteq e(P', c) \Rightarrow r(P, c) \geq r(P', c)$.
- *The optimal linear order* P^* *is such that* $e_i(P^*, c) \leq e_{i+1}(P^*, c)$ *for every* $i = 1, \ldots, n-1$.
- $r(c) = [\sum_{i=1}^{n} e(P^*, c) - 2^{i-1}]\mathbb{E}[U_{(i)}]$.

The above result establishes that to find the optimal rationale in our setting is as simple as finding the rationale that provides an increasing order in the number of problem sets that are rationalized by each alternative. Further, this can be used to characterize the utility loss of actual choice as compared with the one that would imply optimal choice according to P^*.