Minimal Rationalizations

Igor Kopylov*

February 28, 2019

Abstract

Any path-independent choice function Φ can be *rationalized* by some set Θ of total orders over N consumption alternatives. I refine this model of Aizerman and Malishevsky (1981) and construct a *minimal rationalization* Θ such that any set of smaller size does not rationalize Φ. Given that $|\Theta| \leq k$, this construction takes polynomial time $O(N^{k+1})$.

Any *binary* rationalization $|\Theta| \leq 2$ is minimal. For at least 95% of choice functions Φ that have binary rationalizations Θ, such Θ is determined uniquely. This proportion converges to 100% for large N. For $k \geq 3$, uniqueness holds asymptotically when N is large and minimal rationalizations Θ are restricted to a fixed number of elements $n < N$.

Starting from choice data in M menus, I define a *minimax rationalization* Θ that is minimal for the maximal path-independent Φ that is consistent with M available observations. Given that $|\Theta| \leq k$, the minimax rationalization can be found in time $O(MN^{k+2})$ that is polynomial with respect to N and linear with respect to M.

1 Introduction

In the classic model of *utility maximization*, all observed choices $\Phi(A)$ in all *menus*—subsets A of a finite consumption space Z—are rationalized by maximizing some utility function U in A. This model requires the weak axiom of revealed preference (WARP) or some equivalent conditions as in Arrow [4] or Sen [29].

One general critique of utility maximization and WARP is that choices can be produced by *heterogeneous preferences*. This heterogeneity is natural across subjects, but is also common in observations within subjects. In particular, Hey

*Department of Economics, University of California, Irvine. E-mail: ikopylov@uci.edu. I thank Larry Epstein, Itzhak Gilboa, Michael Richter, Asen Kochov, Bart Lipman, Chris Chambers, Federico Echenique, Don Saari, Luca Anderlini for their suggestions. I received valuable feedback at seminars in Boston University, Caltech, Paris School of Economics, Georgetown Unirversity, and IMBS. The draft is preliminary.

[16] and Agranov and Ortoleva [1] provide evidence that individual preferences can vary persistently even over short periods of time.

There are several models that identify heterogenous preferences from choices among a finite number $N = |Z|$ of consumption alternatives. Kalai, Rubinstein, and Spiegler [19] (henceforth, KRS) start with single-valued choices $c(A)$ observed in some menus A and seek the *minimal number* of rationalizations $R_1, \ldots, R_k$ (total orders on Z) such that each observed choice $c(A)$ maximizes some R_i in A. This approach has some problematic issues. First, the empirical meaning of each R_i is unclear because it is not specified when a maximizer of R_i in a menu A should be chosen in this menu. Second, the minimal system $R_1, \ldots, R_k$ is not unique,[1] which hinders further interpretations of R_i. Third, the KRS model cannot be identified in polynomial time (Demuynck [10]) and hence, requires some ad hoc estimates in applications. For example, Crawford and Pendakur [8] use the KRS approach to fit 500 household consumption decisions over different kinds of milk by five distinct rankings. They obtain these rankings by randomly ordering the data points, and then constructing lower and upper bounds for the number of rankings that are required to fit their data.

On the other hand, the *random utility model* (RUM) in Block and Marschak [5] starts with stochastic choice data—probabilities $p(z, A)$ of selecting an element $z \in A$ in a menu A. This model does not minimize the number of rationalizations, but identifies a probability distribution ν over all R such that

$$p(z, A) = \sum_R \nu(R) p_R(z, A) \tag{1}$$

where $p_R(z, A) = 1$ if z maximizes R in A and $p_R(z, A) = 0$ otherwise. This identification strategy can be problematic as well. First, the necessary and sufficient conditions for the existence of ν—the axiom of revealed stochastic preference (ARSP) in McFadden and Richter [26] and the non-negativity of Block-Marschak polynomials in Falmagne [13]—have exponential complexity and require observations in all menus $A \subset Z$. It is unrealistic to get such data unless N is very small. Second, the probability measure ν is not identified uniquely, and degrees of freedom in ν increase with N (see Sher, Fox, Kim and Bajari [30]). Third, the regularity axiom that is required for RUM satisfies regularity axioms, $p(x, A) \geq p(x, B)$ whenever $A \subset B$. This condition can be easily violated by decoys (Huber, Payne, and Puto [18]) and other context effects. For example, if $p(x, \{x, y, z\}) > p(x, \{x, y\})$ is likely if x and y are respectively large and small portions of French fries, and z is a medium portion that is offered at the same price as large. As N increases, regularity becomes even more fragile due to measurement errors. Formally, if $N \geq 2n + 1$ and p satisfies RUM, then there is another p^* that differs from p in a single menu

[1] Liang [22, Proposition 5.1] shows that uniqueness is impossible in the KRS model, with the exception of the singleton case $k = 1$ and another very special situation where $k = 2$, and R_1 and R_2 are two antagonistic orders such that xR_1y if and only if yR_2x.

A such that

$$|p(z, A) - p^*(z, A)| < \tfrac{1}{4^n} \quad \text{for all } z \in A.$$

For example, if $N > 10$, it is enough to perturb choices in a single menu by 0.1% to violate RUM.

In this paper, I identify heterogenous preferences over N alternatives via a new method that

- guarantees polynomial complexity with respect to N,
- provides asymptotic uniqueness as N increases,
- applies when choices are observed only for some menus rather than the entire domain 2^Z.

Moreover, this method can be robust under small perturbations of stochastic choice data.

1.1 Main Results

My main results are formulated for path-independent choice functions Φ that are defined for all menus $A \subset Z$. Aizerman and Malishevski [2] (henceforth, AM) establish that Φ has a rationalization

$$\Theta = \{R_1, \ldots, R_n\}$$

such that a consumption alternative z is chosen in a menu A if and only if z is maximal in A for some $R \in \Theta$. This rationalization notion is stronger than in KRS: not only each choice should maximize some R in Θ, but each maximizer of any R in Θ should be chosen.

AM's proof is constructive (see also Moulin [27, Theorem 5])) and identifies the largest possible rationalization Θ_{max} for Φ. Yet, AM's construction has an exponential complexity and requires that all selections $\Phi(A)$ in all menus A must be observed. There are no uniqueness claims either. The maximal rationalization Θ_{max} is typically large and non-transparent even if a small Θ can be found. For instance, the set Θ_{max} can have the size $2^{|Z|-1}$ even if Φ has a binary rationalization $|\Theta| = 2$ (see Example 2.1 below).

I adopt the KRS minimality criterion for path-independent choice functions Φ. and construct *minimal rationalizations* Θ such that any smaller set $|\Theta^*| < |\Theta|$ does not rationalize Φ. My main results (Theorems 1 and 2) deliver an explicit *focal algorithm* that derives Θ from the function Φ via an endogenously defined family $\mathcal{F}$ of focal menus.

In Theorem 4, I show that for any given $k \geq 1$, the focal algorithm takes polynomial time $O(N^{k+1})$ to find a minimal rationalization Θ such that

$$|\Theta| \leq k \tag{2}$$

or establish that no such Θ exists. Moreover, this procedure requires only choices in menus that have at most $k+1$ elements. The regularization constraint (2) is required because in general, the size of minimal rationalizations can grow exponentially with N (see Example 3.3).

The same constraint $|\Theta| \leq k$ delivers some uniqueness claims about minimal rationalizations. Binary rationalizations $|\Theta| \leq 2$ are always minimal. They are determined uniquely for at least 95% of all choice functions Φ that have a binary rationalization (see Theorem 5). The 95% bound can be further improved to $1 - \frac{5}{|Z|^2}$, which guarantees uniqueness asymptotically as $|Z| \to \infty$. The uniqueness of generic binary rationalizations $\Theta = \{R_1, R_2\}$ allows to model behavioral effects separately for these two rankings. Note that both R_1 and R_2 can be identified in terms of choices in *triples*—menus A such that $|A| \leq 3$—even when choices in larger menus are distorted by inattention or not observed at all.

When $k \geq 3$, minimal rationalizations are rarely unique, but their restrictions to any given menu $B \subset Z$ are asymptotically unique for almost all choice functions that have rationalizations $|\Theta| \leq k$. This claim is shown in Theorem 6.

1.2 Minimax Rationalizations and Stochastic Choice

To apply the above results in empirical settings, one can define path-independent functions Φ endogenously in terms of stochastic choice data or other primitives.

Suppose that distributions $p(\cdot, A)$ are observed only for M menus $A_1, \ldots, A_M$. Fix some threshold $\varepsilon > 0$ and consider the set Θ_{max} of all orders R such that for any A_i and $z \in A_i$ that maximizes R in A_i,

$$p(z, A_i) \geq \varepsilon.$$

In other words, Θ_{max} consists of all rankings R such that the maximization of R with probability at least ε is consistent with all available observations. Thus the maximal path-independent choice function Φ such that for all A_i,

$$\Phi(A_i) \subset \{z \in A_i : p(z, A_i) \geq \varepsilon\} \tag{3}$$

is rationalized by Θ_{max}. If Θ_{max} is empty, then no such Φ exists.

Theorems 1–6 can deliver *minimax rationalizations* Θ that are minimal for the maximal Φ that satisfies (3). Given that $|\Theta| \leq k$, this construction requires time $O(MN^{k+2})$ that is polynomial for N and linear for the number of observations M.

For example, suppose that $Z = \{a, b, c, d\}$ and p is observed for two menus $A_1 = \{a, b, c\}$ and $A_2 = \{c, d\}$. Let

$$p(a, A_1) = 0.6 \qquad p(b, A_1) = 0.3 \qquad p(c, A_1) = 0.1$$
$$p(c, A_2) = 0.9 \qquad p(d, A_2) = 0.1.$$

Take any $\varepsilon \in (0.1, 0.3)$. Then there is a unique minimax rationalization that consists of two rankings R_1 and R_2 such that $aR_1cR_1dR_1b$ and $bR_2cR_2dR_2a$. This rationalization is unchanged even if subjects exhibit some context effects. Indeed, let choices in the menu Z have probabilities $p(a, Z) = 0.3$, $p(b, Z) = 0.2$, $p(c, Z) = 0.48$, and $p(d, Z) = 0.02$. As $p(c, Z) > p(c, A_1)$, then RUM does not hold here. The popularity of c in Z can be motivated by a decoy effect where d can attract attention to c because c dominates d, while a and b do not. The minimax rationalization for the augmented dataset and $\varepsilon \in (0.1, 0.2)$ still consists of R_1 and R_2. This robustness holds because $c \notin \Phi(Z)$.

Next, consider the red bus/blue bus example of Luce [23], where the choices are between a taxi (a), a blue bus (b), and a red bus (c). Let $Z = \{a, b, c\}$ and assume

$$p(a, \{a, b\}) = p(a, \{a, c\}) = p(a, Z) = 0.5 > p(b, Z) > p(c, Z) > 0. \tag{4}$$

Note that some variation in b and c is assumed here (e.g. let b be cheaper than c). Then the unique minimax rationalization for ε in the interval $(p(c, Z), p(b, Z)]$ consists of two rankings aR_1bR_1c and bR_2cR_2a. These rankings capture the natural intuition for (4), where the taxi is more convenient in one scenario, but buses are better in the other.[2] This construction is unchanged if all observed probabilities are perturbed a little. However, the RUM model would not even hold if $p(a, \{a, b\})$ or $p(a, \{a, c\})$ is slightly perturbed to be smaller than $p(a, Z)$.

Moreover, the computational efficiency of the focal algorithm allows to vary ε, try other consistency criteria, or impose strict dominance principles on hypothetical rationalizations R. To do so, one can add N menus to the sequence $A_1, \ldots, A_M$ so that for each $z_i \in Z$, A_{M+i} consists of z_i and all elements that are strictly dominated by z_i. Dominance requires that $p(z_i, A_{M+i}) = 1$. Then one can find a minimax rationalization for the augmented data that has size $M + N$. Section 4 provides further discussion of minimax rationalizations.

1.3 Related Literature

Eliaz, Richter, and Rubinstein [12] (henceforth, ERR) characterize the binary rationalization model and provide another algorithm that identifies R_1 and R_2. They comment on p. 219 that their method does not extend to the non-binary case. Moreover, they do not show the uniqueness properties of R_1 and R_2, and they do not discuss how binary rationalizations can be applied to incomplete data.

Danilov and Koshevoy [9] introduce *Plottizations*—maximal path-independent functions that are consistent with primitive choices. They do not model minimal rationalizations or stochastic choice. Applications of path-independent choice functions to matching mechanisms are studied by Chambers and Yenmez [6].

[2]If $\varepsilon > p(b, Z)$ or $\varepsilon \leq p(c, Z)$, then one gets distinct minimax rationalizations with two and three rankings accordingly. The main case $\varepsilon \in (p(c, Z), p(b, Z)]$ corresponds to the most accurate minimax rationalization of size two.

My Theorem 1 has an analogue for subjective state spaces derived from preferences over menus (Kopylov [21, Theorem 3]), but the two frameworks differ both conceptually and technically.

Apesteguia, Ballester, and Lu [3] characterize a single-crossing RUM (SCRUM) where the support of ν in (1) satisfies the single-crossing property with respect to some exogenous ranking $\succ$ on Z. This model is uniquely identified by stochastic choices in binary menus, and this procedure is robust to small variations or absence of data in larger menus.[3] However, the exogenous ranking $\succ$ is crucial for this regularization. If $\succ$ is not given a priori, then the estimation of $\succ$ appears to be an exponentially hard problem. There are $N!$ candidates, and for each of them, a distinct SCRUM representation is derived from data in binary menus. The problem can be further complicated if choice data is not available for some binary menus.

The classic model of random choice by Luce [23] explains each probability measure $p(\cdot, A)$ in each menu A by a single cardinal utility function $u : Z \to \mathbb{R}_{++}$,

$$p(z, A) = \frac{u(z)}{\sum_{x \in A} u(x)}.$$

Block and Marschak [5] show that the Luce model is a special case of RUM. Similarly to Apesteigua et al., the Luce model can be uniquely identified from choices in pairs.[4] However, the Luce model and its main assumption (IIA) are easily violated by heterogenous preferences, such as the bus example (4) where the choices in binary menus imply $p(a, \{a, b, c\}) = \frac{1}{3}$.

Gul and Pesendorfer [15] characterize the RUM model for choices of lotteries via expected utility rankings R. Their identifications rely on small variations in menus of lotteries and do not have direct analogues in finite settings.

RUM is also prominent in econometric literature, both in parametric (e.g. McFadden [25]) and non-parametric forms (e.g. Kitamura and Stoye [20], Matzkin [24]).

2 Preliminaries

Let $Z = \{a, b, c, x, y, z \dots\}$ be a consumption domain of finite size $N \geq 2$.

Let $\mathcal{A} = \{A, B, C, \dots\}$ be the set of all *menus*—non-empty subsets of Z. The size of any menu $A \in \mathcal{A}$ is denoted as $|A|$. With some abuse of notation, small menus are listed without curly brackets and commas: singleton menus $\{x\}$ are written as x and menus like $\{a, b, c\}$ as abc.

A family $\mathcal{C} \subset \mathcal{A}$ is called a *chain* if for all $A, B \in \mathcal{C}$, either $A \supset B$ or $B \supset A$. If $\mathcal{C}$ is a chain, then $|\mathcal{C}| \leq N$ because for all $A, B \in \mathcal{C}$, either $A = B$ or $|A| \neq |B|$. A chain $\mathcal{C}$ is called *total* if $|\mathcal{C}| = N$.

[3]The perfect fit for binary menus does not guarantee accuracy for choices in larger menus.

[4]Fix any $z \in Z$. Let $u(y) = \frac{p(y,\{y,z\})}{p(z,\{y,z\})}$.

Let $\mathcal{R} = \{R, \dots\}$ be the set of all complete and transitive rankings of Z. For any $R \in \mathcal{R}$, write its symmetric parts as I. Say that $R \in \mathcal{R}$ is *total* (or *linear*) if for all $x, y \in Z$, xIy implies $x = y$.

Let $\mathcal{T}$ be the set of all total orders $R \in \mathcal{R}$. The pair of brackets $\langle \dots]$ is used to list rankings $R \in \mathcal{T}$ in a brief symbolic form. For example, $\langle abcd]$ denotes the total order $aRbRcRd$ on $Z = abcd$.

For any $R \in \mathcal{T}$, let $\mathcal{L}(R)$ be the family of its *contour sets* that have the form

$$L(R, y) = \{z \in Z : yRz\}$$

for some $y \in Z$. The family $\mathcal{L}(R)$ is a total chain because $R \in \mathcal{T}$. Conversely, if $\mathcal{C}$ is a total chain, then there is a unique $R_{\mathcal{C}} \in \mathcal{T}$ such that $\mathcal{C} = \mathcal{L}(R_{\mathcal{C}})$. The total order $\mathcal{R}_{\mathcal{C}}$ is represented over all $z \in Z$ by the utility function

$$U_{\mathcal{C}}(z) = |\{A \in \mathcal{C} : z \notin A\}| \tag{5}$$

that counts the number of menus in the chain $\mathcal{C}$ that do not contain z. For example, $\mathcal{C} = \{a, ac, acd, Z\}$ is a total chain. Thus $\mathcal{C} = \mathcal{L}(R_{\mathcal{C}})$, where $R_{\mathcal{C}} = \langle bdca]$ is represented by $U_{\mathcal{C}}$ such that $U_{\mathcal{C}}(b) = 4$, $U_{\mathcal{C}}(d) = 3$, $U_{\mathcal{C}}(c) = 2$ and $U_{\mathcal{C}}(a) = 1$.

2.1 Choice functions and Rationalizations

Adapt the classic framework of Uzawa [31] and Arrow [4].

Say that $\Phi : \mathcal{A} \to 2^Z$ is a *choice function* if $\Phi(A)$ is not empty and $\Phi(A) \subset A$ for all $A \in \mathcal{A}$. Consider two standard axioms.

Axiom (α)**.** *For all $A, B \in \mathcal{A}$ and $z \in A$, if $z \in \Phi(A \cup B)$, then $z \in \Phi(A)$.*

This axiom is named Property α by Sen [29]; it appears earlier in Chernoff [7].

Axiom (Path-Independence)**.** *For all $A, B \in \mathcal{A}$,*

$$\Phi(A \cup B) = \Phi(\Phi(A) \cup B).$$

This condition is proposed by Plott [28]. Path-independence implies Axiom α,

$$z \notin \Phi(A) \quad \Rightarrow \quad z \notin \Phi(A \cup B) = \Phi(\Phi(A) \cup (B \setminus A)).$$

For any ranking $R \in \mathcal{R}$, non-empty set $\Theta \subset \mathcal{R}$, and menu $A \in \mathcal{A}$, write

$$\begin{aligned}
\Phi_R(A) &= \{x \in A : x \text{ maximizes } R\} \\
\Phi_\Theta(A) &= \bigcup_{R \in \Theta} \Phi_R(A) = \{x \in A : x \text{ maximizes some } R \in \Theta\}.
\end{aligned}$$

Say that the choice functions Φ_R and Φ_Θ are *rationalized* by R and Θ respectively.

Uzawa [31, Theorem 4] characterizes all functions Φ_R that can be rationalized by some $R \in \mathcal{T}$. Say that Φ is *single-valued* if $|\Phi(A)| = 1$ for all $A \in \mathcal{A}$.

Theorem (Uzawa). *Φ is single-valued and obeys Axiom α if and only if $\Phi = \Phi_R$ for some $R \in \mathcal{T}$. Moreover, such $R \in \mathcal{T}$ is unique.*

The required ranking R is easily derived from choices in *pairs* (two-element menus): for all $x, y \in X$,

$$xRy \quad \Leftrightarrow \quad x \in \Phi(xy). \tag{6}$$

Single-valuedness and Axiom α imply that such R is a total order. Without single-valuedness, one can WARP or Arrow's [4] equivalent condition to establish that R is complete and transitive. Note that WARP implies path-independence.

AM characterize the class of choice functions Φ_Θ that are rationalized by some $\Theta \subset \mathcal{R}$. Without loss in generality, such rationalizations can be restricted to the class of total orders so that $\Theta \subset \mathcal{T}$.

Let $\mathbb{P}$ be the class of all path-independent choice functions.

Theorem (Aizerman–Malishevsky). *For any choice function Φ,*

$$\Phi \in \mathbb{P} \quad \Leftrightarrow \quad \Phi = \Phi_\Theta \textit{ for some } \Theta \subset \mathcal{T} \quad \Leftrightarrow \quad \Phi = \Phi_\Theta \textit{ for some } \Theta \subset \mathcal{R}.$$

AM's proof is constructive (see also Moulin [27, Theorem 5]). They show that any $\Phi \in \mathbb{P}$ is rationalized by the set

$$\Theta_{max} = \{R \in \mathcal{T} \, : \, \Phi_R(A) \subset \Phi(A) \quad \text{for all } A \in \mathcal{A}\}. \tag{7}$$

Note that Θ_{max} contains all sets Θ such that $\Phi = \Phi_\Theta$, but Φ may have rationalizations Θ that are much smaller than Θ_{max}. As illustrated next, the size of Θ_{max} can grow exponentially with N (e.g. $|\Theta_{max}| = 2^{N-1}$) even when Φ allows small rationalizations Θ (e.g. $|\Theta| = 2$).

Say that I is an *interval* if

$$I = [m, n] = \{i \in \mathbb{N} : m \leq i \leq n\}$$

for some $m, n \in \mathbb{N}$ such that $1 \leq m \leq n \leq N$.

Example 2.1. *Let $Z = [1, N]$. Let $\Phi = \Phi_\Theta$ where $\Theta = \{R_1, R_2\}$ and $R_1, R_2 \in \mathcal{T}$ are represented by $U_1(i) = i$ and $U_2(i) = N - i$ respectively. Then the rationalization Θ_{max} consists of all $R \in \mathcal{T}$ such that $\mathcal{L}(R)$ is a total chain of intervals*

$$[1, N] = I_N \supsetneq I_{N-1} \supsetneq I_{N-2} \supsetneq \cdots \supsetneq I_1.$$

There are 2^{N-1} chains of this sort because for each $k \in [2, N]$, there are two ways to reduce an interval I_k of size k to $I_{k-1} \subset I_k$ of size $k - 1$. Thus $|\Theta_{max}| = 2^{N-1}$.

The exponential complexity of Θ_{max} is problematic. First, it is hard to interpret the multitude of its elements in any practical way. Second, it is unclear when Θ_{max} can be reduced to some convenient small rationalizations $\Theta \subset \Theta_{max}$. Third, identification (7) requires exponential time and exponential amount of choice data: it must retrieve and analyze all choices $\Phi(A)$ in $2^N - 1$ menus A.

3 Main Results

Say that $\Theta \subset \mathcal{T}$ is *minimal* if for all $\Theta^* \subset \mathcal{T}$,

$$\Phi_\Theta = \Phi_{\Theta^*} \quad \Rightarrow \quad |\Theta^*| \geq |\Theta|,$$

so that Φ_Θ cannot be rationalized by any set that has fewer elements than Θ.

Let $\mathbb{M}$ be the set of all minimal sets $\Theta \subset \mathcal{T}$. Any $\Theta \in \mathbb{M}$ is called a *minimal rationalization* for Φ_Θ.

Take any path-independent $\Phi \in \mathbb{P}$. By AM's Theorem, Φ has a minimal rationalization $\Theta \in \mathbb{M}$. One can find Θ by *brute force*: check all subsets $\Theta \subset \Theta_{max}$ and take the smallest one such that $\Phi = \Phi_\Theta$. However, this method is very slow because it must check the equalities $\Phi(A) = \Phi_\Theta(A)$ across exponentially large numbers of menus $A \in \mathcal{A}$ and subsets $\Theta \subset \Theta_{max}$. Moreover, the search by brute force does not provide any insights about the size, composition, and uniqueness of minimal rationalizations.

To construct minimal rationalizations in a faster and more informative way, define an endogenous domain of *focal menus.* For each $y \in Z$, let

$$\mathcal{B}_y = \{B \in \mathcal{A} : y \in \Phi(B \cup y) \text{ and } y \notin B\}$$

be the family of all menus $B \in \mathcal{A}$ such that $y \notin B$ is chosen in $B \cup y$. The sets $\mathcal{B}_y$ are sufficient to recover selections $\Phi(A)$ in all menus A,

$$\Phi(A) = \{y \in A : A \setminus y \in \mathcal{B}_y\}.$$

Take any $y \in Z$. Say that $F \in \mathcal{B}_y$ is *focal for* y if F is maximal in $\mathcal{B}_y$. The family of all menus that are focal for y is written as

$$\mathcal{F}_y = \{F \in \mathcal{B}_y : y \notin \Phi(B) \text{ for all } B \supsetneq F \cup y\}.$$

A menu $F \in \mathcal{A}$ is called *focal* if F is focal for some $y \in Z$. Let

$$\mathcal{F} = \bigcup_{y \in Z} \mathcal{F}_y$$

be the set of all focal menus. This family is called the *focus* of Φ.

A menu $A \in \mathcal{A}$ is called *central* if A is the intersection of some focal menus. All focal menus are central. By convention, $A = Z$ is central, but not focal.[5] Let

$$\pi(\mathcal{F}) = \{A \in \mathcal{A} : A = \bigcap_{F \in \mathcal{F} : F \supset A} F\}$$

be the set of all central menus.

Chains $\mathcal{E} \subset \pi(\mathcal{F})$ are called *central.*

[5] Z is the intersection of an empty family of focal menus.

Theorem 1. *For any* $\Theta \subset \mathcal{T}$ *and* $\Phi \in \mathbb{P}$*,*

$$\Phi = \Phi_\Theta \qquad \Leftrightarrow \qquad \mathcal{F} \subset \bigcup_{R \in \Theta} \mathcal{L}(R) \subset \pi(\mathcal{F}). \tag{8}$$

Moreover, any central chain $\mathcal{D}$ *can be extended to a total central chain* $\mathcal{E} \supset \mathcal{D}$*.*

The proof is in the appendix.

Theorem 1 characterizes all rationalizations for any path-independent choice function $\Phi \in \mathbb{P}$ in term of its focus $\mathcal{F}$. It also suggests an algorithm for finding minimal rationalizations that is effective both computationally and conceptually. In particular, it implies some *uniqueness* claims about minimal rationalizations.

Define the *width* W of Φ as the smallest number of chains that cover the focus $\mathcal{F}$. Formally, there are W chains $\mathcal{C}_1, \dots, \mathcal{C}_W$ such that

$$\mathcal{F} \subset \mathcal{C}_1 \cup \dots \cup \mathcal{C}_W, \tag{9}$$

but there are no $W-1$ chains $\mathcal{C}_1, \dots, \mathcal{C}_{W-1}$ such that $\mathcal{F} \subset \mathcal{C}_1 \cup \dots \cup \mathcal{C}_{W-1}$.

Note that any cover of the focus $\mathcal{F}$ by W chains can be rewritten as a partition of $\mathcal{F}$ into W disjoint chains,

$$\mathcal{F} = \mathcal{D}_1 \cup \dots \cup \mathcal{D}_W, \tag{10}$$

where $\mathcal{D}_i = (\mathcal{F} \cap \mathcal{C}_i) \setminus \cup_{n=1}^{i-1} \mathcal{C}_n$ for all $i \leq W$. Say that partition (10) is *unique* if for any other partition of $\mathcal{F}$ into W disjoint chains $\mathcal{D}_1^*, \dots, \mathcal{D}_W^*$, the families $\{\mathcal{D}_i\}_{i=1}^W$ and $\{\mathcal{D}_i^*\}_{i=1}^W$ are equal.

Theorem 2. *For any* $\Theta \in \mathbb{M}$ *and* $\Phi \in \mathbb{P}$*,*

$$\Phi = \Phi_\Theta \qquad \Leftrightarrow \qquad \mathcal{F} \subset \bigcup_{R \in \Theta} \mathcal{L}(R) \subset \pi(\mathcal{F}) \quad \text{and} \quad |\Theta| = W. \tag{11}$$

Moreover, if the focus $\mathcal{F}$ *has a unique partition into* W *chains* $\mathcal{D}_1, \dots, \mathcal{D}_W$*, and each chain* $\mathcal{D}_i$ *has a unique extension to a total central chain* $\mathcal{E}_i \supset \mathcal{D}_i$*, then the minimal rationalization* $\Theta \in \mathbb{M}$ *is unique.*

Thus for any path-independent choice function $\Phi \in \mathbb{P}$, the size of its minimal rationalization Θ can be computed as the width W of the focus $\mathcal{F}$.

The well-known combinatorial result of Dilworth [11] provides another interpretation for W. A family $\mathcal{D} \subset \mathcal{A}$ is called an *antichain* if for all $A, B \in \mathcal{D}$, $A \subset B$ implies $A = B$. Dilworth's Theorem asserts that

$$W = \text{the maximal size of an antichain } \mathcal{D} \subset \mathcal{F}.$$

For example, the family $\mathcal{F}_y$ is an antichain for any $y \in Z$ because $\mathcal{F}_y$ consists of maximal elements in $\mathcal{B}_y$. Thus

$$W \geq |\mathcal{F}_y|. \tag{12}$$

Theorem 2 is derived from Theorem 1 in a constructive way.

Take any choice function $\Phi \in \mathbb{P}$. Consider the *focal algorithm* that runs in three broad steps.

Step 1. *Find the focus $\mathcal{F}$.*

Step 2. *Find the width W and partition $\mathcal{F}$ into W disjoint chains $\mathcal{D}_1, \ldots, \mathcal{D}_W$,*

Step 3. *Extend each $\mathcal{D}_i \subset \mathcal{F}$ to a total central chain $\mathcal{E}_i \supset \mathcal{D}_i$. Let*

$$\Theta = \{\mathcal{R}_{\mathcal{E}_1}, \ldots, \mathcal{R}_{\mathcal{E}_W}\} \tag{13}$$

where $\mathcal{L}(R_{\mathcal{E}_i}) = \mathcal{E}_i$ for all $i \leq W$. Note that the rankings $\mathcal{R}_{\mathcal{E}_i}$ are constructed together with their utility representations $U_{\mathcal{E}_i}$.

By Theorem 1, the focal algorithm is guaranteed to produce some rationalization Θ. Indeed, $\mathcal{F}$ can be partitioned into W chains as in (10). Any chain $\mathcal{D}_i \subset \mathcal{F}$ is central and hence, can be extended to a total central chain $\mathcal{E}_i \supset \mathcal{D}_i$. Then

$$\mathcal{F} \subset \mathcal{D}_1 \cup \cdots \cup \mathcal{D}_W \subset \mathcal{E}_1 \cup \cdots \cup \mathcal{E}_W \subset \pi(\mathcal{F}).$$

By (11), Θ rationalizes Φ. On the other hand, for any rationalization Θ^*, condition (11) implies that $\mathcal{F}$ is covered by the chains $\mathcal{L}(R)$ for $R \in \Theta^*$. Thus $|\Theta^*| \geq W$.

Thus the rationalization Θ produced by the focal algorithm is minimal and satisfies $|\Theta| = W$.

Turn to the uniqueness part. Let $\mathcal{D}_1, \ldots, \mathcal{D}_W$ be the unique partition of the focus $\mathcal{F}$ into W chains. Let Θ satisfy (13) for the unique total central chains $\mathcal{E}_i \supset \mathcal{D}_i$. Let Θ^* be any other minimal rationalization for Φ. Then $\Theta^* = \{R_1^*, \ldots, R_W^*\}$ for some distinct $R_i^* \in \mathcal{T}$. Let $\mathcal{C}_i^* = \mathcal{L}(R_i^*)$ for all i. By Theorem 2, $\mathcal{F}$ is covered by W chains $\mathcal{C}_i^*$ and hence, partitioned by W chains $\mathcal{D}_i^* = (\mathcal{F} \cap \mathcal{C}_i^*) \setminus \cup_{n=1}^{i-1} \mathcal{C}_n^*$. Thus the families $\{\mathcal{D}_i\}_{i=1}^W$ and $\{\mathcal{D}_i^*\}_{i=1}^W$ are equal. Each chain $\mathcal{D}_i^*$ has a unique extension to a total central chain $\mathcal{E}_i^* \supset \mathcal{D}_i^*$. Note that $\mathcal{C}_i^* \supset \mathcal{D}_i^*$ is a total central chain and hence, $\mathcal{C}_i^* = \mathcal{E}_i^*$. Thus the families $\{\mathcal{E}_i\}_{i=1}^W$, $\{\mathcal{E}_i^*\}_{i=1}^W$ and $\{\mathcal{C}_i^*\}_{i=1}^W$ are all equal, and $\Theta^* = \Theta$.

The uniqueness condition in Theorem 2 holds in some settings (e.g. Example 3.1 below), but is rare in general. In Sections 3.4, I obtain weaker asymptotic statements about uniqueness.

3.1 Examples

To illustrate Theorem 2 and the focal algorithm, consider several examples.

Example 3.1. *Let $Z = abcd$. Take Φ such that*

$\Phi(abcd) = ac$ $\Phi(abc) = ac$ $\Phi(acd) = ac$ $\Phi(abd) = ad$ $\Phi(bcd) = bc$

$\Phi(ab) = a$ $\Phi(ac) = ac$ $\Phi(ad) = ad$ $\Phi(bc) = bc$ $\Phi(bd) = bd$ $\Phi(cd) = cd.$

The maximal elements in the families

$$\mathcal{B}_a = \{\mathbf{bcd}, bc, bd, cd, b, c, d\} \qquad \mathcal{B}(c) = \{\mathbf{abd}, ab, ad, bd, a, b, d\}$$
$$\mathcal{B}_b = \{\mathbf{cd}, c, d\} \qquad \mathcal{B}_d = \{\mathbf{ab}, a, b, \mathbf{c}\}$$

are marked in bold. They form the focus $\mathcal{F} = \{bcd, abd, cd, ab, c\}$ *of width* $W = 2$*. Indeed,* $\mathcal{F}$ *is not a chain, and has a unique partition into two chains*

$$\mathcal{F} = \mathcal{D}_1 \cup \mathcal{D}_2 = \{bcd, cd, c\} \cup \{abd, ab\}.$$

The chains $\mathcal{D}_1$ *and* $\mathcal{D}_2$ *have unique extensions to total central chains*

$$\mathcal{E}_1 = \{Z, bcd, cd, c\} \supset \mathcal{D}_1$$
$$\mathcal{E}_2 = \{Z, abd, ab, b = bcd \cap ab\} \supset \mathcal{D}_2$$

Thus $\Theta = \{\langle abdc], \langle cdab]\}$ *is the unique minimal rationalization for* Φ*, providing that* Φ *is path-independent. Instead of checking* $\Phi \in \mathbb{P}$*, one can check* $\Phi = \Phi_\Theta$ *directly.*

Example 3.2. *Let* $Z = abcd$*. Take* Φ *such that*

$$\begin{array}{llllll} \Phi(Z) = ab & \Phi(abc) = ab & \Phi(acd) = a & \Phi(abd) = ab & \Phi(bcd) = b & \\ \Phi(ab) = ab & \Phi(ac) = a & \Phi(ad) = a & \Phi(bc) = b & \Phi(bd) = b & \Phi(cd) = cd. \end{array}$$

The maximal elements in the families

$$\begin{array}{ll} \mathcal{B}_a = \{\mathbf{bcd}, bc, bd, cd, b, c, d\} & \mathcal{B}_c = \{\mathbf{d}\} \\ \mathcal{B}_b = \{\mathbf{acd}, ac, ad, cd, a, c, d\} & \mathcal{B}_d = \{\mathbf{c}\} \end{array}$$

form the focus $\mathcal{F} = \{bcd, acd, c, d\}$ *of width* $W = 2$*. There are two distinct ways to cover* $\mathcal{F}$ *by total central chains:*

$$\mathcal{F} = \{Z, bcd, cd, c\} \cup \{Z, acd, cd, d\} = \{Z, bcd, cd, d\} \cup \{Z, acd, cd, c\}$$

Accordingly, Φ *has two minimal rationalizations* $\Theta = \{\langle abcd], \langle badc]\}$ *and* $\Theta^* = \{\langle abdc], \langle bacd]\}$*. Check that* $\Phi = \Phi_\Theta = \Phi_{\Theta^*}$*.*

Example 3.3. *Let* $Z = [1, N]$*. Take any* $q \in [1, N]$*. For all* $A \in \mathcal{A}$*, let*

$$\Phi^q(A) = \{x \in A : |A \cap [x, N]| \leq q\}$$

consist of top q *numbers in the menu* A*. If* $|A| \leq q$*, then* $\Phi^q(A) = A$*. Note that* Φ^q *is path-independent because for any* $x \in \Phi^q(A \cup B)$*, either* $x \in \Phi^q(A)$ *or* $x \in B$*.*

The class of all focal menus for 1 *is*

$$\mathcal{F}_1 = \{A \in [2, N] : |A| = q - 1\}.$$

There are $\binom{N-1}{q-1}$ *of such menus. The family* $\mathcal{F}_1$ *is an antichain because all* $A \in \mathcal{F}_1$ *have the same size* $q - 1$*. Thus* $W \geq \binom{N-1}{q-1}$*. Note that* Φ^q *has a rationalization* Θ *such that* $|\Theta| = \binom{N-1}{q-1}$*. For any* $A \in \mathcal{F}_1$*, let* $R_A \in \mathcal{T}$ *be represented by*

$$U_A(x) = \begin{cases} x & \text{if } x \in A \\ x + N & \text{if } x \notin A. \end{cases}$$

Then Φ^q *is rationalized by* $\Theta = \{R_A : A \in \mathcal{F}_1\}$*. Theorem 2 asserts that* $W \leq |\Theta| = \binom{N-1}{q-1}$*. Thus* $W = \binom{N-1}{q-1}$*, and* Θ *is a minimal rationalization for* Φ^q*.*

Example 3.3 shows that minimal rationalizations need not be small. In particular, rationalizing the "top two" choice function Φ^2 requires at least $N-1 = \binom{N-1}{1}$ total orders (this observation is due to ERR). If $q = \lfloor N/2 \rfloor$, then $W = \binom{N-1}{q-1}$ is the central binomial coefficient that grows at the exponential rate $O\left(\frac{2^N}{\sqrt{0.5\pi N}}\right)$. For example, if $N = 20$ and $q = 10$, then $W = \binom{19}{9} = 92378$.

Therefore, some choice functions $\Phi \in \mathbb{P}$ can be rationalized only by very large sets Θ, but still modeled in some parsimonious form (e.g. the top q representation). Another lesson is that in the worst-case scenario, finding a minimal rationalization requires exponential time with respect to N because W can grow exponentially with N.

3.2 Computational Complexity

To reduce the computational complexity of the focal algorithm, constrain the size $|\Theta|$ by some integer k. This exogenous constraint can be also used to make uniqueness claims about Θ. It can be motivated by a priori considerations about heterogenous preferences. For example, $k = 2$ can be natural in models of household choice.

Take any $k \in \mathbb{N}$. Let

$$\mathcal{A}_k = \{A \in \mathcal{A} : |A| \leq k\}$$
$$\mathbb{T}_k = \{\Theta \subset \mathcal{T} : |\Theta| \leq k\}$$

be the families of all menus $A \in \mathcal{A}$ and sets $\Theta \subset \mathcal{T}$ respectively that contain at most k elements. Let

$$\mathbb{P}_k = \{\Phi \in \mathbb{P} : \Phi = \Phi_\Theta \text{ for some } \Theta \in \mathbb{T}_k\}.$$

In other words, $\mathbb{P}_k$ is the class of all choice functions for which there are k total orders $R_1, \ldots, R_k \in \mathcal{T}$ (not necessarily distinct) such that

$$\Phi(A) = \Phi_{R_1}(A) \cup \Phi_{R_2}(A) \cup \cdots \cup \Phi_{R_k}(A) \qquad \text{for all } A \in \mathcal{A}.$$

Obviously, $\mathbb{P}_1 \subset \mathbb{P}_2 \subset \mathbb{P}_3 \subset \cdots \subset \mathbb{P} = \bigcup_{k=1}^{\infty} \mathbb{P}_k$.

Take any $\Phi \in \mathbb{P}$ with focus $\mathcal{F}$ and width W.

Corollary 3. *For any $k \in \mathbb{N}$, the following statements are equivalent:*

(i) $\Phi \in \mathbb{P}_k$,

(ii) $W \leq k$,

(iii) the focus $\mathcal{F}$ can be covered by k chains $\mathcal{D}_1, \ldots, \mathcal{D}_k$,

(iv) for any $k+1$ focal menus $F_1, \ldots, F_{k+1} \in \mathcal{F}$, there are $i \neq j$ such that $F_i \subset F_j$.

Indeed, the equivalence between (i) and (ii) follows from Theorem 2. It can be written as a formula

$$W = \min\{k \in \mathbb{N} : \Phi \in \mathbb{P}_k\}.$$

The equivalence of (ii), (iii), and (iv) follows from the definition of the width W and Dilworth's Theorem.

To find minimal rationalizations, one needs to use operations of several types

- given a menu $A \in \mathcal{A}$, observe the choice selection $\Phi(A)$,
- given a menu $A \in \mathcal{A}$ and an order $R \in \mathcal{T}$, find the maximizer $\Phi_R(A)$,
- given any menus $A, B \in \mathcal{A}$, perform a standard set operation (e.g. intersection $A \cap B$) or check whether A and B are nested.

Call such operations *basic tasks.*

Theorem 4. *For any $k \geq 1$ and choice function $\Phi \in \mathbb{P}$, one can use the focal algorithm to test the hypothesis $\Phi \in \mathbb{P}_k$ and find a minimal rationalization $\Theta \in \mathbb{T}_k$ if $\Phi \in \mathbb{P}_k$ is true. This algorithm requires*

(i) at most $\frac{2}{k!}N^{k+1} + O(N^k)$ basic tasks, and

(ii) observations $\Phi(A)$ only in menus $A \in \mathcal{A}_{k+1}$.

The proof is in the appendix.

This result suggests that it is computationally feasible to find minimal rationalizations for $k = 10$ and $N \approx 20$, or for $k = 4$ and $N \approx 1000$.

3.3 Uniqueness for binary rationalizations

Rationalizations $\Theta \in \mathbb{T}_2$ are called *binary.* Note that any binary rationalization is minimal. If $|\Theta| = 1$, then $\Theta \in \mathbb{M}$ is trivial. If $\Theta = \{R_1, R_2\}$ with two distinct rankings $R_1, R_2 \in \mathcal{T}$, then there are x, y such that xR_1y, but not yR_2x. Then $\Phi_\Theta(xy) = xy$ and hence, $\Phi \notin \mathbb{P}_1$. Thus $\Phi \in \mathbb{M}$.

For any $\Phi \in \mathbb{P}_2$, a binary rationalization $\Theta \in \mathbb{T}_2$ can be found in $5N^2$ basic tasks. (See the proof of Theorem 4 in the appendix.) It takes another N^3 operations to test $\Phi \in \mathbb{T}_2$ if that is not guaranteed a priori.

The uniqueness of a binary rationalization $\Theta \in \mathbb{T}_2$ h can be established via a more accurate condition than the one in Theorem 2.

A set $\Theta = \{R_1, R_2\} \in \mathbb{T}_2$ is called *dissonant* if for any $A \in \mathcal{L}(\mathcal{R}_1) \cap \mathcal{L}(R_2)$, either the restrictions of R_1 and R_2 are equal on A, or they are equal on $Z \setminus A$. In particular, dissonance must hold if R_1 and R_2 do not share any lower contour set A of size $|A| \in [2, N-2]$. Obviously, if A or $Z \setminus A$ is a singleton, then R_1 and R_2 have the same vacuous restriction to A or $Z \setminus A$ respectively.

Theorem 5. *A set $\Theta \in \mathbb{T}_2$ is dissonant if and only if Θ is the unique binary rationalization for the choice function Φ_Θ. Moreover,*

$$\frac{|\Phi \in \mathbb{P}_2 : \ \Phi \textit{ has a unique binary rationalization}|}{|\mathbb{P}_2|} \geq \max\left\{\tfrac{95}{100}, 1 - \tfrac{5}{N^2}\right\}. \qquad (14)$$

For example, $\Theta = \{\langle abdc], \langle cdab]\}$ is dissonant in Example 3.1, but $\Theta = \{\langle abcd], \langle badc]\}$ is not dissonant in Example 3.2 where uniqueness does not hold.

The dissonance property holds for most sets $\Theta \in \mathbb{T}_2$. Accordingly, the proportion of choice functions $\Phi \in \mathbb{P}_2$ that have unique binary rationalizations exceeds 95% for any Z and converges to 1 as $N \to \infty$. Example 3.2 presents a rare situation where binary rationalizations are not unique. If $N = 4$, then the set $\mathbb{P}_2$ has 294 elements, and 288 choice functions $\Phi \in \mathbb{P}_2$ have unique binary rationalizations.

3.4 Uniqueness in $\mathbb{P}_k$

Take any $k \geq 3$.

Minimal rationalizations for choice functions $\Phi \in \mathbb{P}_k$ can be unique even if they are not binary (as in Example A.1), but such uniqueness is rare.[6] To provide some intuition, consider the generic case where N is large, and take some random triple $\Theta = \{R_1, R_2, R_3\} \in \mathbb{T}_3$. With probability close to one, $\Theta \in \mathbb{M}$ is minimal because the top elements in R_1, R_2, R_3 are all distinct. The three focal sets of size $N - 1$ form an antichain. Take the worst two elements a, b for R_1. Assume aR_1b. Then with probability 50% the two orders R_2 and R_3 should have distinct comparisons between a and b, such as aR_2b and bR_3a. Replace R_1 with R_1^* that switches only the two bottom elements a and b so that bR_1^*a and $R_1^* = R_1$ for all other comparisons. Let $\Theta^* = \{R_1^*, R_2, R_3\}$. Then $\Phi_\Theta = \Phi_{\Theta^*}$ and Θ^* is another minimal rationalization. The same argument can be applied to the pairs of the worst elements in R_2 and R_3. With probability close to 1, all of these pairs will be distinct. Thus one can expect that at most $\frac{1}{8}$ of all triples $\Theta = \{R_1, R_2, R_3\} \in \mathbb{T}_3$ can produce unique minimal rationalizations for Φ_Θ.

To make a positive statement about uniqueness of minimal rationalizations $\Theta \in \mathbb{T}_k$ for $\Phi \in \mathbb{P}_k$, restrict Θ to a fixed set $B \subset Z$ and vary the domain Z.

Say that two rationalizations Θ, Θ^* are *equivalent* on a menu $B \in \mathcal{A}$ if there is a bijection $\beta : \Theta \to \Theta^*$ such that for all $R \in \Theta$, the two orders $\beta(R) \in \mathcal{T}$ and $R \in \mathcal{T}$ are the same when restricted to B. For example, the sets

$$\Theta = \{\langle abc], \langle acb], \langle cba]\}$$
$$\Theta^* = \{\langle cab], \langle bca], \langle acb]\}$$

are equivalent on $B = ab$ because the ranking aRb occurs twice and bRa occurs once in both Θ and Θ^*. Yet Θ and Θ^* are not equivalent on $B = ac$ because Θ

[6] In the KRS model it is impossible beyond a special case of binary rationalizations. See Liang [22, Proposition 5.1].,

contains two rankings aRc and one ranking cRa, while Θ^* has two rankings cRa and one ranking aRc.

For any menu $B \in \mathcal{A}$, let $\mathbb{P}_{k,B}$ be the class of all choice functions $\Phi \in \mathbb{P}_k$ such that any of its rationalizations $\Theta, \Theta^* \in \mathbb{T}_k$ are equivalent on B. A fortiori, if $\Phi \in \mathbb{P}_{k,B}$, then all of its minimal rationalizations Θ are equivalent on B because $\Theta \in \mathbb{T}_k$.

Theorem 6. *For any $k \geq 3$, any set B, and consumption domain $Z \supset B$,*

$$\lim_{N\to\infty} \frac{|\mathbb{P}_{k,B}|}{|\mathbb{P}_k|} = 1. \tag{15}$$

Moreover, for any $\alpha > 0$, there is $M_\alpha \in \mathbb{N}$ such that

$$\frac{|\Phi \in \mathbb{P}_k : \ \Phi \textit{ has at most } M_\alpha \textit{ minimal rationalizations}|}{|\mathbb{P}_k|} \geq 1 - \alpha. \tag{16}$$

The proof is in the appendix. Roughly speaking, if k and B are fixed, then one should expect that if Z is large relative to B, then minimal rationalizations for a choice function $\Phi \in \mathbb{P}_k$ should be all equivalent on B.

Example 3.4. *Let $Z = abcde$. Take Φ such that*

$\Phi(abcde) = ad$	$\Phi(abcd) = ad$	$\Phi(abce) = ac$	$\Phi(acde) = ad$	$\Phi(bcde) = bd$
$\Phi(abc) = ac$	$\Phi(abd) = ad$	$\Phi(abe) = ab$	$\Phi(acd) = ad$	$\Phi(ace) = ac$
$\Phi(ade) = ad$	$\Phi(bcd) = bd$	$\Phi(bce) = bce$	$\Phi(bde) = bd$	$\Phi(cde) = cde$
$\Phi(ab) = ab$	$\Phi(ac) = ac$	$\Phi(ad) = ad$	$\Phi(ae) = a$	$\Phi(bc) = bc$
$\Phi(bd) = bd$	$\Phi(be) = be$	$\Phi(cd) = cd$	$\Phi(ce) = ce$	$\Phi(de) = de$

The maximal elements in the families

$$\begin{aligned}
\mathcal{B}_a &= \{\mathbf{bcde},\ \textit{all menus } A \subset bcde\} \\
\mathcal{B}_d &= \{\mathbf{abce},\ \textit{all menus } A \subset abce\} \\
\mathcal{B}_b &= \{\mathbf{ae}, \mathbf{cde}, cd, de, ce, a, c, d, e\} \\
\mathcal{B}_c &= \{\mathbf{abe}, \mathbf{de}, ab, ae, be, a, b, d, e\} \\
\mathcal{B}_e &= \{\mathbf{bdc}, bd, bc, dc, b, d, c\}
\end{aligned}$$

form the focus $\mathcal{F} = \{abce, bcde, cde, abe, bcd, ae, de\}$ that has width $W = 3$. It can be partitioned into three chains in two different ways:

$$\mathcal{F} = \{abce, abe, ae\} \cup \{bcde, cde, de\} \cup \{bcd\} = \{abce, abe, ae\} \cup \{cde, de\} \cup \{bcde, bcd\}.$$

The first chain has a unique extension to $\{abce, abe, ae, e\}$ and produces the order $R_1 = \langle dcbae]$. The second chain produces R_2 that is either $\langle abcde]$ or $\langle abced]$. The third chain produces R_3 that is either $\langle aebdc]$, or $\langle aebcd]$, or $\langle aedbc]$, or $\langle aedcb]$. Note that the restriction of $\Theta = \{R_1, R_2, R_3\}$ to $B = \{a, b, e\}$ is uniquely determined as $\langle bae]$, $\langle abe]$, and $\langle aeb]$. Obviously, if one observes the choice function Φ only on subsets of B, then such identification is impossible because the ranking $\langle abe]$ becomes superfluous.

4 Minimax Rationalizations

When Z is not very small, it is hardly possible to observe choices in all pairs $|A| \leq 2$ and a fortiori, in all larger menus. Moreover, it is unrealistic to expect that big datasets should fit some path-independent choice function precisely.

Yet minimal rationalizations can be still found for path-independent choice functions that are defined endogenously to *approximate* primitive data.

To introduce such approximations, consider a primitive *data sequence* $\mathcal{S}$

$$(A_1, C_1), (A_2, C_2), \ldots, (A_M, C_M)$$

that consists of M pairs $(A_i, C_i) \in \mathcal{A} \times \mathcal{A}$ such that $C_i \subset A_i$. Each menu $C_i \subset A_i$ is interpreted as the set of all choices in A_i that are consistent with the available observations in A_i.

For example, given a distribution $p(\cdot, A_i)$ and some threshold $\varepsilon > 0$, let

$$C_i = \{z \in A_i : p(z, A_i) \geq \varepsilon\}.$$

Say that a total order $R \in \mathbb{T}$ is *consistent* with $\mathcal{S}$ if $\Phi_R(A_i) \in C_i$ for all i. In other words, $R \in \mathbb{T}$ is rejected as inconsistent based on a single mismatch $\Phi_R(A) \notin C_i$ for some i.

Let Θ_{max} be the collection of all such R,

$$\Theta_{max} = \{R \in \mathbb{T} : \Phi_R(A) \in C_i \quad \text{for all } i.\}.$$

If Θ_{max} is empty, then the above consistency notion is too strong to model the data sequence $\mathcal{S}$.

Suppose that Θ_{max} is not empty. Let $\Pi \in \mathbb{P}$ be rationalized by Θ_{max}. Then

$$\Pi(A_i) \subset C_i \quad \text{for all } i. \tag{17}$$

Equivalently, Π is the maximal path-independent choice function that satisfies inclusions (17). Indeed, any $\Pi^* \in \mathbb{P}$ that satisfies (17) can be rationalized by some $\Theta \subset \Theta_{max}$ and hence, satisfies $\Pi^*(A_i) \subset \Pi(A_i)$ for all i.

Say that $\Theta \subset \mathbb{T}$ is a *minimax rationalization* for the data sequence $\mathcal{S}$ if Θ is a minimal rationalization for Π.

Let $\mathbb{S}_k$ be the set of all data sequences such that the corresponding function Π belongs to $\mathbb{P}_k$.

Theorem 7. *For any $k \geq 1$ and $\mathcal{S}$, it requires $O(MN^{k+2})$ basic operations to establish whether $\mathcal{S} \in \mathbb{S}_k$ and if so, find a minimax rationalization $\Theta \in \mathbb{T}_k$.*

The minimax rationalization is convenient in models of *stochastic choice*, where the primitive data is given via probabilities $p(z, A)$ of selecting an element $z \in A$ in a menu A. Consider representation

$$p(z, A) = \sum_R \nu(R, A) p_R(z, A), \tag{18}$$

where the probability distribution $\nu(\cdot, A)$ is defined over $\mathcal{T}$ and can vary with the menu A.

Take some threshold $\varepsilon > 0$ and assume that there is a minimal set Θ of total orders such that for all A and $R \in \Theta$,

$$\nu(R, A) > \varepsilon \geq 1 - \sum_{R \in \Theta} \nu(R, A). \tag{19}$$

A APPENDIX: PROOFS

I start with a lemma.

Take any path-independent $\Phi \in \mathbb{P}$. Let $\mathcal{F}$ be its focus.

Lemma A.1. *For any $F \in \mathcal{F}$ and $y \in Z$,*

$$y \notin F \quad \Rightarrow \quad y \in \Phi(F \cup y), \tag{20}$$

$$F \in \mathcal{F}_y \quad \Rightarrow \quad F \cup y = \bigcap_{F^* \in \mathcal{F} : F^* \supsetneq F} F^*. \tag{21}$$

Proof. Indeed, if F is focal for y, then $y \in \Phi(F \cup y)$ by definition. If F is focal for some some $z \neq y$, then $z \in \Phi(F \cup z)$, but $z \notin \Phi(F \cup z \cup y)$. If $y \notin \Phi(F \cup y)$, then by path-independence,

$$\Phi(F \cup z \cup y) = \Phi(\Phi(F \cup y) \cup z)$$

includes z because $\Phi(F \cup y) \cup z \subset F \cup z$. This contradiction proves (20).

Take any menu $F \in \mathcal{F}_y$ that is focal for y. Take any $F^* \in \mathcal{F}$ such that $F^* \supsetneq F$. Then F^* is focal for some $x \neq y$. Suppose that $y \notin F^*$. Then $x \in \Phi(F^* \cup x)$, but $y \notin \Phi(F^* \cup y)$ because F is maximal in $\mathcal{B}_y$. Thus

$$\Phi(F^* \cup x \cup y) = \Phi(\Phi(F^* \cup y) \cup x)$$

contains x because $\Phi(F^* \cup y) \cup x \subset F^* \cup x$. As $F^* \cup y \in \mathcal{B}_x$, then $F^* \cup y = F^*$ because F^* is maximal in $\mathcal{B}_x$. Thus $y \in F^*$, and

$$F \cup y \subset \bigcap_{F^* \in \mathcal{F} : F^* \supsetneq F} F^*.$$

To show the opposite inclusion, take any $x \notin F \cup y$. Then $x \in \Phi(F \cup y \cup x)$. Otherwise,

$$\Phi(F \cup y \cup x) = \Phi(\Phi(F \cup y \cup x) \cup y).$$

and hence, $y \in \Phi(F \cup y \cup x)$ because $y \in \Phi(F \cup y)$ and $\Phi(F \cup y \cup x) \cup y \subset F \cup y$. As $x \in \Phi(F \cup y \cup x)$, then $F \cup y \in \mathcal{B}_x$. Thus there is a menu $F_x \supset F \cup y$ that is focal for x. Thus $F \cup y \supset \bigcap_{x \notin F \cup y} F_x$, and (21) holds. ∎

A.1 Proof of Theorem 1

Take any $\Theta \subset \mathcal{T}$. Assume that $\Phi = \Phi_\Theta$. Show the inclusions

$$\mathcal{F} \subset \bigcup_{R \in \Theta} \mathcal{L}(R) \subset \pi(\mathcal{F}). \tag{22}$$

Take any $R \in \Theta$. Take any $B \in \mathcal{L}(R)$. Then $B = \{z \in Z : xRz\}$ for some $x \in Z$. Take any $y \notin B$. Then $y = \Phi_R(B \cup y) \in \Phi_\Theta(B \cup y) = \Phi(B \cup y)$. As $B \in \mathcal{B}_y$, then there is $F_y \in \mathcal{F}_y$ such that $F_y \supset B$. As $F_y \in \mathcal{F}_y$, then $y \notin F_y$. Thus $B = \bigcap_{y \notin B} F_y$ is central.

Take any focal menu $F \in \mathcal{F}$. Then F is maximal in $\mathcal{B}_y$ for some $y \in Z$. As $y \in \Phi_\Theta(F \cup y)$, then there is $R \in \Theta$ such that $y \in \Phi_R(F \cup y)$. Take the contour set $B = \{z \in Z : yRz\}$. As $y \in \Phi_R(F \cup y)$, then $F \cup y \subset B$. Suppose that $F \cup y \neq B$. Then there is $x \in B$ such that $x \neq y$ and $x \notin F$. As $\Phi_R(x \cup y \cup F) = y$, then F is not maximal in $\mathcal{B}_y$. By contradiction, $F \cup y = B$. Let $y^* = \Phi_R(B \setminus y)$. Then

$$F = B \setminus y = \{z \in Z : y^* R z\} \in \mathcal{L}(R)$$

and hence, (22) holds.

Assume instead that (22) holds. Take any non-singleton menu $A \in \mathcal{A}$ and show that $\Phi(A) = \Phi_\Theta(A)$. Take any $y \in \Phi(A)$. As $A \setminus y \in \mathcal{B}_y$, then there is a menu $F \supset A \setminus y$ that is focal for y. Take $R \in \Theta$ such that $F \in \mathcal{L}(R)$. Then $y = \Phi_R(F \cup y)$. As $A \subset F \cup y$ and $y \in A$, then $y \in \Phi_R(A) \subset \Phi_\Theta(A)$.

Conversely, take any $y \in \Phi_\Theta(A)$. Then $y = \Phi_R(A)$ for some $R \in \Theta$. Let $x = \Phi_R(A \setminus y)$ and $B = \{z \in Z : xRz\}$. Then $B \in \mathcal{L}(R)$ and $y = A \setminus B$. As B is central and $y \notin B$, then there is some focal menu $F \supset B$ such that $y \notin F$. By (20), $y \in \Phi(F \cup y)$ and hence, $y \in \Phi(A)$.

Finally, take any central chain $\mathcal{D} \subset \pi(\mathcal{F})$ and extend it to a total central chain $\mathcal{E} \supset \mathcal{D}$. Wlog $Z \in \mathcal{D}$. If $|\mathcal{D}| = N$, then $\mathcal{D}$ is total. Suppose that $|\mathcal{D}| < N$. As $\mathcal{D}$ is a chain, then for all $A, B \in \mathcal{D}$, either $A = B$, or $|A| \neq |B|$. Thus there is $B \in \mathcal{D}$ such that $|B| > 1$, but there is no $A \in \mathcal{D}$ such that $|A| = |B| - 1$.

Let A be the menu of the largest size in $\mathcal{D}$ such that $|A| < |B|$. If no such menu exist, let $A = \emptyset$. I claim that $\Phi(B) \setminus A \neq \emptyset$. If $A = \emptyset$, then it is true. Suppose $A \neq \emptyset$. Then $A \in \mathcal{D}$ is central. Then there is a focal menu $F \supset A$ such that $F \supset B$ does not hold. (Otherwise, $A = B$.) Let $y \in B \setminus F$. By (20), $y \in \Phi(A \cup y)$. If $\Phi(B) \subset A$, then by path-independence $\Phi(B) = \Phi(\Phi(B) \cup y)$ contains y because $\Phi(B) \cup y \subset A \cup y$. By contradiction, $\Phi(B) \setminus A \neq \emptyset$.

Take any $y \in \Phi(B) \setminus A$. Take a menu $F \supset B \setminus y$ that is focal for y. Then

$$A \subset B \cap F = B \setminus y. \tag{23}$$

As $|B \setminus y| = |B| - 1$, then $\mathcal{E}^* = \mathcal{D} \cup \{B \cap F\}$ is a central chain such that $\mathcal{E}^* \supset \mathcal{D}$ and $|\mathcal{E}^*| = |\mathcal{D}| + 1$. By induction with respect to the size of the chain $\mathcal{D}$, there is a total central chain $\mathcal{E} \supset \mathcal{D}$.

A.2 The proof of Theorem 4

Its main steps are sketched next.

First, let $k = 1$. In this case, one needs at most $\frac{N(N-1)}{2}$ observations to check whether $\Phi(A)$ is single-valued for all two-element menus $A \in \mathcal{A}_2$. If so, then $\Phi \in \mathbb{P}_1$ and its minimal rationalization Θ consists of the ranking R such that for all $x, y \in Z$, xRy if and only if $x \in \Phi(xy)$. The case $k = 1$ can be analyzed even faster in $O(N \log N)$ time by running a standard sorting algorithm, such as mergesort or heapsort, on the ranking R.

Second, let $k \geq 2$. Assume that $\Phi \in \mathbb{P}_k$. In the appendix, I show that the focal algorithm takes $O(N^k)$ basic tasks to obtain a minimal rationalization $\Theta \in \mathbb{T}_k$. If the algorithm fails at this stage, then $\Phi \in \mathbb{P}_k$ is rejected. To make the test conclusive, one needs to check further that

$$\Phi(A) = \Phi_\Theta(A) \quad \text{for all } A \in \mathcal{A}_{k+1}. \tag{24}$$

It takes up to $2k + 1$ basic tasks[7] to check (24) for one menu A, and there are $\frac{1}{(k+1)!} N^{k+1} + O(N^k)$ menus in $\mathcal{A}_{k+1}$. The product does not exceed

$$\frac{2k+1}{(k+1)!} N^{k+1} + O(N^k) < \frac{2}{k!} N^{k+1} + O(N^k).$$

Thus finding Θ at the preliminary stage is asymptotically faster than checking equalities (24).

Corollary 8. *For any $k \geq 1$ and choice function $\Phi \in \mathbb{P}_k$, one can find a minimal rationalization $\Theta \in \mathbb{T}_k$ in $O(N^k)$ basic tasks.*

Note that for $k \geq 3$, the focus $\mathcal{F}$ is partitioned into the minimal number of chains via an algorithm from Hopcroft and Karp [17]. As $\mathcal{F}$ is assumed to be covered by k chains, then $|\mathcal{F}| \leq kN$. Therefore, Hopcroft and Karp's algorithm finds a minimal partition (10) of the focus $\mathcal{F}$ in $O\left(N^{5/2}\right)$ basic tasks. See the appendix for further details.

Take any $k \geq 2$ and $\Phi \in \mathbb{P}_k$. I argue that the focus $\mathcal{F}$ can be found in $O(N^k)$ basic tasks that require observations $\Phi(A)$ only for $A \in \mathcal{A}_{k+1}$.

Step I. I show how the focus $\mathcal{F}$ can be found in WN^{W+1} basic operations. Roughly speaking, all focal menus are obtained via a bottom-up procedure rather than the top-down approach in AM's proof. Then a minimal rationalization Θ can be identified in terms of the focus $\mathcal{F}$ via standard algorithms from discrete mathematics.

Write all elements in Z in some order $Z = \{z_1, z_2, \ldots, z_N\}$.

[7] Find $\Phi_R(A)$ for $R \in \Theta$, take the union of these elements, observe $\Phi(A)$, and check the equality (24).

Take any $a \in Z$ and let $\mathcal{F}_a$ be the set of all menus that are focal for a. By definition, $\mathcal{F}_a$ is an antichain because all of its elements are maximal in $\mathcal{B}_a$ and hence, cannot be nested. Thus $|\mathcal{F}_a| \leq W$.

Note that for any $A \in \mathcal{B}_a$, one can find a focal menu $F \in \mathcal{F}_a$ such that $F \supset A$ in at most N basic operations. To do so, let $A_0 = A$ and for each $i = 1, \dots, N$, let

$$A_i = \begin{cases} A_{i-1} \cup z_i & \text{if } A_{i-1} \cup z_i \in \mathcal{B}_a \\ A_{i-1} & \text{if } A_{i-1} \cup z_i \notin \mathcal{B}_a. \end{cases}$$

Let $F = A_N$. Then $F \in \mathcal{F}_a$ because for any $x \notin F \cup a$, $x \cup A_{i-1}$ does not belong to $\mathcal{B}_a$. By α, $x \cup F$ does not belong to $\mathcal{B}_a$ either. Thus F is maximal in $\mathcal{B}_a$.

Step II. This step is a well-known problem in combinatorics. Hopcroft and Karp [17] formulate an algorithm that finds the width and a minimal partition of any partial order (e.g. a family of menus) of size n in $O\left(n^{5/2}\right)$ time. Here

Focal menus can be found in polynomial time with respect to the size of the set Z and the width of the focus $W(\mathcal{F})$.

Felsner, Raghavan, and Spinrad [14] provide faster algorithms that recognize whether the width of a given partial order does not exceed k. For $k = 1, 2, 3$ the recognition can be done in linear time.

Step III.

I claim that the set $\mathcal{F}_a$ can be found in $O(WN^W)$ basic operations.

For any menu $A \in \mathcal{A}$, it takes two basic operations $a \in \Phi(A \cup a)$ and $a \notin A$ to establish whether $A \in \mathcal{B}_a$. Find the minimal i such that $z_i \in \mathcal{B}_a$. If no such i exists, then $\mathcal{F}_a$ is empty.

check whether A For any elements $a, x \in Z$ and $A \in \mathcal{A}$, say that x is acceptable for (a, A) if $a \in \Phi(x \cup a \cup A)$.

To do so,

Example A.1. *Let $Z = abcde$. Take Φ such that*

$\Phi(abcde) = ae$	$\Phi(abcd) = ab$	$\Phi(abce) = ae$	$\Phi(acde) = ae$	$\Phi(bcde) = bce$
$\Phi(abc) = ab$	$\Phi(abd) = ab$	$\Phi(abe) = ae$	$\Phi(acd) = ac$	$\Phi(ace) = ae$
$\Phi(ade) = ae$	$\Phi(bcd) = bc$	$\Phi(bce) = bce$	$\Phi(bde) = be$	$\Phi(cde) = ce$
$\Phi(ab) = ab$	$\Phi(ac) = ac$	$\Phi(ad) = ad$	$\Phi(ae) = ae$	$\Phi(bc) = bc$
$\Phi(bd) = bd$	$\Phi(be) = be$	$\Phi(cd) = c$	$\Phi(ce) = ce$	$\Phi(de) = de$

The maximal elements in the families

$$\mathcal{B}_a = \{\mathbf{bcde},\ \textit{all menus } A \subset bcde\}$$
$$\mathcal{B}_e = \{\mathbf{abcd},\ \textit{all menus } A \subset abcd\}$$
$$\mathcal{B}_b = \{\mathbf{acd}, \mathbf{cde}, ac, ad, cd, de, ce, a, c, d, e\}$$
$$\mathcal{B}_c = \{\mathbf{bde}, \mathbf{ad}, bd, be, de, a, b, d, e\}$$
$$\mathcal{B}_d = \{\mathbf{a}, \mathbf{b}, \mathbf{e}\}$$

form the focus $\mathcal{F} = \{abcd, bcde, acd, cde, bde, ad, a, b, e\}$ that has width $W = 3$. Indeed, $\mathcal{F}$ contains a maximal antichain acd, cde, bde, and can be partitioned into three chains in two different ways[8]

$$\mathcal{F} = \mathcal{D}_1 \cup \mathcal{D}_2 \cup \mathcal{D}_3 = \{bcde, cde, e\} \cup \{bde, b\} \cup \{abcd, acd, ad, a\} = \\ \{cde, e\} \cup \{bcde, bde, b\} \cup \{abcd, acd, ad, a\} = \mathcal{D}'_1 \cup \mathcal{D}'_2 \cup \mathcal{D}'_3.$$

Moreover, each $\mathcal{D}_i$ has a unique extension to a total central chain,

$$\mathcal{E}_1 = \{Z, bcde, cde, de = cde \cap bde, e\} \supset \{bcde, cde, e\} = \mathcal{D}_1$$
$$\mathcal{E}_2 = \{Z, bcde, bde, bd = bde \cap abcd, b\} \supset \{bde, b\} = \mathcal{D}_2$$
$$\mathcal{E}_3 = \{Z, abcd, acd, ad, a\} \supset \{abcd, acd, a\} = \mathcal{D}_3.$$

The chains $\mathcal{D}'_i$ are uniquely extended by $\mathcal{E}_i$ as well. Thus $\Theta = \{\langle abcde], \langle acedb], \langle ebcda]\}$ is the unique minimal rationalization Φ. Check that $\Phi = \Phi_\Theta$.

A.3 Proofs of Theorems 5

Wlog let $Z = [1, N]$. If Z consists of some other elements, then relabel them by the numbers $1, \ldots, N$.

Take any $\Phi \in \mathbb{P}_2$. Let $\mathcal{F}_i$ be the set of all menus that are focal for $i \in [1, N]$. By (12), $|\mathcal{F}_i| \le 2$.

To find $\mathcal{F}_i$, construct two sequences of sets

$$A^0_{i1} \subset A^1_{i1} \subset \cdots \subset A^N_{i1} \subset Z \quad \text{and} \quad A^0_{i2} \subset A^1_{i2} \subset \cdots \subset A^N_{i2} \subset Z$$

Start with singleton menus $A^0_{i1} = i$ and $C^0 = i$. For each $k \ge 1$, consider two possible cases.

(1) If $i \in \Phi(C^{k-1} \cup k)$, let $A^k_{i1} = A^{k-1}_{i1} \cup k$ and $C^k = \Phi(C^{k-1} \cup k)$.

(2) If $i \notin \Phi(C^{k-1} \cup k)$, keep $A^k_{i1} = A^{k-1}_{i1}$ and $C^k = C^{k-1}$.

Note that $|C^{k-1} \cup k| \le 3$ because $|C^{k-1}| \le 2$. By induction, $|C^k| \le 2$ because $\Phi \in \mathbb{P}_2$, and the choice selection $\Phi(C^{k-1} \cup k)$ contains at most two elements.

If $A^N_{i1} = i$, then $\mathcal{F}_i = \emptyset$.

If $A^N_{i1} \ne i$, then the menu $F_{i1} = A^N_{i1} \setminus i$ is focal for i. Let $m \in [1, N]$ be the minimal index such that $m \notin A^N_{i1}$ and $i \in \Phi(i \cup m)$. If no such m exists, then $\mathcal{F}_i = \{F_{i1}\}$.

Suppose that m exists. Let $A^0_{i2} = C^0 = i \cup m$. For each $k \ge 1$, consider two possible cases.

[8] The sets *cde*, *bde*, *acd* form an antichain and hence, belong to distinct chains. The set *acd* is the only one that can be in one chain with *ad*, or *a*, or *abcd*. The set *bde* is the only one that can be in one chain with *b*. The set *cde* is the only one that can be in one chain with *e* once *b* is added to the chain with *bde*. The set *bcde* can be in one chain with either *cde* or *bde*. But it does not affect the minimal rationalization.

(1) If $i \in \Phi(C^{k-1} \cup k)$, let $A^k_{i2} = A^{k-1}_{i2} \cup k$ and $C^k = \Phi(C^{k-1} \cup k)$.

(2) If $i \notin \Phi(C^{k-1} \cup k)$, keep $A^k_{i2} = A^{k-1}_{i2}$ and $C^k = C^{k-1}$.

The menu $F_{i2} = A^N_{i2} \setminus i$ is the second focal menu for i. Thus $\mathcal{F}_i = \{F_{i1}, F_{i2}\}$.

This construction is illustrated in Example A.2.

Note that each set $\mathcal{F}_i$ is obtained in at most $3N$ basic operations: at most N operations to find all A^k_{i1}, at most N operations to find m or establish that it does not exist, and at most N operations to find all A^k_{i2}. Thus the entire focus $\mathcal{F} = \cup_{i=1}^N \mathcal{F}_i$ can be found in at most $3N^2$ basic operations. Moreover, for each $n \in [1, N-1]$, one can populate the list $\mathcal{F}(n)$ of all focal menus F such that $|F| = n$. As $\mathcal{F}(n)$ is an antichain, then $|\mathcal{F}(n)| \leq 2$.

Next, cover the focus $\mathcal{F}$ by two total central chains $\mathcal{E}_1, \mathcal{E}_2$,

$$Z = E^N_1 \supsetneq E^N_1 - 1 \supsetneq \cdots \supsetneq E^1_1 \supsetneq \emptyset$$
$$Z = E^N_2 \supsetneq E^{N-1}_2 \supsetneq \cdots \supsetneq E^1_2 \supsetneq \emptyset.$$

Note that $|E^k_1| = |E^k_2| = k$ for all $k \in [1, N]$. For each $n \in [1, N]$, let

$$\mathcal{E}(n) = \{E^n_1, E^n_2\}.$$

Note that $\mathcal{F}(n) \subset \mathcal{E}(n)$ is required.

Proceed inductively starting from $k = N$. For each $k < N$, assume that the central menus E^{k+1}_1 and E^{k+1}_2 are given. Assume that the families $\mathcal{E}(n)$ for $n > k$ are uniquely defined. Construct E^k_1 and E^k_2 and argue that the family $\mathcal{E}(k) = \{E^k_1, E^k_2\} \supset \mathcal{F}(k)$ is determined uniquely as well.

Assume that $\mathcal{E}(k+1)$ consists of one menu $E = E^{k+1}_1 = E^{k+1}_2$. Then E is a subset of all focal menus in $\mathcal{F}(n)$ for $n > k$. Thus each intersection of focal menus F such that $|F| \leq k+1$ must have at least $k+1$ elements. Thus to continue the chains $\mathcal{E}_1$ and $\mathcal{E}_2$ by k-element central menus, one can use only focal menus $F \in \mathcal{F}(k)$. If there is only one such menu $F \in \mathcal{F}(k)$, let $E^k_1 = E^k_2 = F$. If there are two distinct menus $F_1, F_2 \in \mathcal{F}(k)$, let $E^k_1 = F_1$ and $E^k_2 = F_2$. In either case, the family $\mathcal{E}(k)$ is defined uniquely as $\mathcal{F}(k)$.

Assume that $\mathcal{E}(k+1)$ consists of two distinct menus $E_1 = E^{k+1}_1$ and $E_2 = E^{k+1}_2$. Take any focal menu $F \in \mathcal{F}(k)$. Then exactly one of the inclusions $F \subset E_1$ or $F \subset E_2$ must hold. Indeed, if these inclusions do not hold, then the focus $\mathcal{F}$ cannot be covered by two central chains because $\mathcal{E}_1$ and $\mathcal{E}_2$ are identified uniquely for all sizes in $[k+1, N]$ and cannot be continued to size k. If both of these inclusions, hold, then $F = E_1 \cap E_2$ because $E_1 \neq E_2$. By (21), the focal menu F cannot be written as the intersection of larger focal menus. Thus each focal menu of size k can be uniquely matched with either E_1 or E_2.

Consider several cases.

- $\mathcal{F}(k)$ contains two distinct focal menus $F_1, F_2 \in \mathcal{F}$ of size k. Wlog $F_1 \subset E_1$ and $F_2 \subset E_2$. Let $E^k_1 = F_1$ and $E^k_2 = F_2$. The family $\mathcal{E}(k) = \mathcal{F}(k)$ is determined uniquely.

- $\mathcal{F}(k)$ contains one focal menu F. Wlog $F \subset E_1$. Find the largest $m \geq k+1$ such that $|E_1^m \cap E_2| = k$. Such m exists because $|E_1^N \cap E_2| = k+1$, and $|E_1^{k+1} \cap E_2| \leq k$. To find such m, one needs to check at most N intersections $E_1^m \cap E_2$. Let $E_1^k = F_1$ and $E_2^k = E_1^m \cap E_2$. The family $\mathcal{E}(k)$ is unique because there is exactly one central menu $E_2^k \subset E_2$ such that $|E_2^k| = k$.
- $\mathcal{F}(k)$ is empty. Find the largest $m \geq k+1$ such that $|E_1^m \cap E_2| = k$. Let $E_2^k = E_1^m \cap E_2$. Then find the largest $m \geq k+1$ such that $|E_2^m \cap E_1| = k$. Let $E_1^k = E_2^m \cap E_1$. Again the family $\mathcal{E}(k)$ is unique.

Each of these cases requires at most $2N$ basic operations. There $N-1$ steps. So the construction of the chains $\mathcal{E}_1, \mathcal{E}_2$ takes at most $2N^2$ operations. Assigning the utility functions $U_{\mathcal{E}_1}$ and $U_{\mathcal{E}_2}$ can be done at no extra cost when $\mathcal{E}_1^k$ and $\mathcal{E}_2^k$ are found. Moreover, in all of the above three cases, the menus E_1^k and E_2^k are determined uniquely together with their order. So continuation of the chains $\mathcal{E}_1$ and $\mathcal{E}_2$ need not be unique only when $\mathcal{E}(k+1)$ consist of one menu, and there are two distinct focal menus F_1, F_2 of size k. Moreover, if $k = N-1$, then one can take $\mathcal{E}_1^{N-1} = F_1$ by convention.

Consider another example.

Example A.2. *Let $Z = abcdef$. Take Φ such that*

$\Phi(abc) = ac$	$\Phi(abd) = ab$	$\Phi(abe) = ab$	$\Phi(abf) = ab$	$\Phi(acd) = ac$
$\Phi(ace) = ac$	$\Phi(acf) = ac$	$\Phi(ade) = ae$	$\Phi(adf) = ad$	$\Phi(aef) = ae$
$\Phi(bcd) = bc$	$\Phi(bce) = ce$	$\Phi(bcf) = bc$	$\Phi(bde) = be$	$\Phi(bdf) = b$
$\Phi(bef) = be$	$\Phi(cde) = ce$	$\Phi(cdf) = c$	$\Phi(cef) = ce$	$\Phi(def) = e$
$\Phi(ab) = ab$	$\Phi(ac) = ac$	$\Phi(ad) = ad$	$\Phi(ae) = ae$	$\Phi(af) = a$
$\Phi(bc) = bc$	$\Phi(bd) = b$	$\Phi(be) = be$	$\Phi(bf) = b$	$\Phi(cd) = c$
$\Phi(ce) = ce$	$\Phi(cf) = c$	$\Phi(de) = e$	$\Phi(df) = df$	$\Phi(ef) = e.$

The fast algorithm for finding the focus $\mathcal{F}$ has the following sketch where choices are marked in bold.

$\mathbf{a} \to \mathbf{ab} \to \mathbf{a}b\mathbf{c} \to \mathbf{a}b\mathbf{c}d \to \mathbf{a}b\mathbf{c}de \to \mathbf{a}b\mathbf{c}def$	$F_{a1} = bcdef$
$\mathbf{b} \to \mathbf{ba} \to \mathbf{ba}d \to \mathbf{ba}de \to \mathbf{ba}def$	$F_{b1} = adef$
$\mathbf{b} \to \mathbf{bc} \to \mathbf{bc}d \to \mathbf{bc}df$	$F_{b2} = cdf$
$\mathbf{c} \to \mathbf{ca} \to \mathbf{ca}b \to \mathbf{ca}bd \to \mathbf{ca}bde \to \mathbf{ca}bdef$	$F_{c1} = abdef$
$\mathbf{d} \to \mathbf{da} \to \mathbf{da}f$	$F_{d1} = af$
$\mathbf{e} \to \mathbf{ea} \to \mathbf{ea}d \to \mathbf{ea}df$	$F_{e1} = adf$
$\mathbf{e} \to \mathbf{eb} \to \mathbf{e}b\mathbf{c} \to \mathbf{e}b\mathbf{c}d \to \mathbf{e}b\mathbf{c}df$	$F_{e2} = bcdf$
$\mathbf{f} \to \mathbf{df}$	$F_{f1} = d.$

Note that $\mathcal{F}_a = \{bcdef\}$, $\mathcal{F}_b = \{adef, cdf\}$, $\mathcal{F}_c = \{abdef\}$, $\mathcal{F}_d = \{af\}$, $\mathcal{F}_e = \{adf, bcdf\}$, *and* $\mathcal{F}_f = \{d\}$. *See details Appendix A.3.*

Construct two total central chains $\mathcal{E}_1$ *and* $\mathcal{E}_2$ *as follows.*

$\mathcal{E}_1$	$\mathcal{E}_2$	
Z	Z	Z *is a central menu of size* 6
$bcdef$	$abdef$	*two focal menus of size* 5
$bcdf$	$adef$	*two focal menus of size* 4
cdf	adf	*two focal menus of size* 3
$df = cdf \cap abdef$	af	*one focal menu of size* 2 *and a central filling*
d	$f = af \cap bcdef$	*one focal menu of size* 1 *and a central filling.*

Thus if $\Phi \in \mathbb{P}_2$, *then* $\Theta = \{\langle aebcfd, cbedaf]\}$ *is its binary rationalization. It is unique because* Θ *is dissonant.*

Turn to Theorem 5. Suppose that $\Theta \in \mathbb{T}_2$ is dissonant. Let $\Phi = \Phi_\Theta$. Assume that the above algorithm does not have a unique continuation at some step $k < N-1$. Let $\Theta^* = \{R_1^*, R_2^*\}$ be a rationalization produced by the algorithm. By construction, there is a common lower contour set C of size $k+1$ for the rankings R_1^* and R_2^*. As $\Theta = \{R_1, R_2\}$ is dissonant, then C cannot be a lower contour set for both R_1 and R_2. Wlog C is not a lower contour set for R_1. Thus there is $a \notin C$ such that $a \neq \Phi_{R_1}(C \cup a)$. However,

$$\Phi_{R_1^*}(C \cup a) = \Phi_{R_2^*}(C \cup a) = a.$$

Thus $\Phi_{\Theta^*}(C \cup a) = a \neq \Phi_\Theta(C \cup a)$. This contradiction shows that the above algorithm will have a unique continuation at each step $k < N-1$ when $\Phi = \Phi_\Theta$ and Θ is dissonant.

Let $Z = [1, N]$. Note that

$$|\mathbb{T}_2| = N! + \frac{N!(N!-1)}{2} = \frac{N!(N!+1)}{2}.$$

The cases $N \leq 6$ are checked directly.

- If $N \leq 3$, then all binary rationalizations $\Theta \in \mathbb{T}_2$ are unique.
- Let $N = 4$. Then there are 300 distinct binary rationalizations. Only 12 of them are not unique. Accordingly, there are 6 choice functions $\Phi \in \mathbb{P}_2$ that do not have unique binary rationalization. Thus 288 of binary rationalizations are unique. Note that $\frac{288}{288+6} > 97\%$.
- Let $N = 5$. Then there are 7260 distinct binary rationalizations. Only $540 = \frac{120*9}{2}$ of them are not unique (for each ranking R, there are nine R' such that $\{R, R'\}$ is not unique). Accordingly, there are 270 choice functions $\Phi \in \mathbb{P}_2$ that do not have unique rationalizations, and $6720 = 7260 - 540$ that have unique ones. Note that $\frac{6720}{6720+270} \approx 96.1\%$.

- For $N = 6$ and $N = 7$, the ratio of choice functions $\Phi \in \mathbb{P}_2$ with unique binary rationalizations still exceeds 95%.

Let $N \geq 8$. Take any binary rationalization $\Theta \in \mathbb{T}_2$.

Suppose that $\Theta = \{R, R'\}$ where R and R' are distinct. For each $k \in [2, N-1]$, let

$$r_k = k!(N-k)!$$

be the number of rankings $R' \in \mathcal{T}$ such that $L(R,k) = L(R',k)$. Note that

$$r_2 = r_{N-2} > r_3 = r_{N-3} \geq r_i$$

for any $i \in [3, N-3]$. Thus the total number of R' for which $\{R, R'\}$ need not be unique does not exceed

$$Q_N = (N-2)! + \sum_{k=3}^{N-3} r_k + (N-2)!.$$

Note that if R' and R have the same order on the bottom two elements, then uniqueness is not rejected. Thus $r_2 = 2(N-2)!$ is reduced to $(N-2)!$ in the above sum. Similarly for r_{N-2}. Thus there are at most $0.5Q_N N!$ of sets $\Theta \in \mathbb{T}_2$ that are not unique binary rationalizations. The number of choice functions $\Phi \in \mathbb{P}_2$ that do not have unique binary rationalizations is at most $0.25Q_N N!$. There are at least $|\mathbb{T}_2| - 0.5Q_N N!$ of choice functions $\Phi \in \mathbb{P}_2$ for which uniqueness holds.

$$\alpha = \frac{0.25Q_N N!}{|\mathbb{T}_2| - 0.5Q_N N! + 0.25Q_N N!} \leq \frac{Q_N}{2N! - Q_N}$$

Check that $\alpha < 5\%$ for $N = 8, 9$. Note that

$$Q_N \leq 2(N-2)! + (N-4) * 6 * (N-3)! \leq 8(N-2)!$$

and conclude that

$$\alpha \leq \frac{4}{N(N-1)-4} < \frac{5}{N^2}$$

for all $N \geq 10$.

References

[1] M. Agranov and P. Ortoleva. Stochastic choice and preferences for randomization. *Journal of Political Economy*, 125:40–68, 2017.

[2] M. Aizerman and A. Malishevski. General theory of best variants choice. *IEEE Trans. Automat. Control*, pages 1030–1041, 1981.

[3] J. Apesteguia, M. Ballester, and J. Lu. Single-crossing random utility models. *Econometrica*, 85:661–674, 2017.

[4] K. Arrow. Rational choice functions and orderings. *Econometrica*, 26:121–127, 1959.

[5] H. Block and J. Marschak. Random orderings and stochastic theories of response. In Olkin, Ghurye, Hoeffding, Madow, and Mann, editors, *Contributions to Probability and Statistics*, pages 97–132. Stanford University Press, Palo Alto, 1960.

[6] C. Chambers and M. B. Yenmez. Choice and matching. *American Economic Journal: Microeconomics*, 9:126–147, 2017.

[7] H. Chernoff. Rational selection of decision functions. *Econometrica*, 22:422–443, 1954.

[8] I. Crawford and K. Pendakur. How many types are there? *Economic Journal*, 123:77–95, 2012.

[9] V. Danilov and G. Koshevoy. Mathematics of plott choice functions. *Mathematical Social Sciences*, 49:245–272, 2005.

[10] T. Demuynck. The computational complexity of rationalizing boundedly rational choice behavior. *Journal of Mathematical Economics*, 47:425–433, 2011.

[11] R. P. Dilworth. A decomposition theorem for partially ordered sets. *Annals of Mathematics*, 51:161166, 1950.

[12] K. Eliaz, M. Richter, and A. Rubinstein. Choosing the two finalists. *Economic Theory*, 46:211–219, 2011.

[13] J. Falmagne. A representation theorem for finite random scale systems. *Journal of Mathematical Psychology*, 18:52–72, 1978.

[14] S. Felsner, V. Raghavan, and J. Spinrad. Recognition algorithms for orders of small width and graphs of small dilworth number. *Order*, 20:351–364, 2003.

[15] F. Gul and W. Pesendorfer. Random expected utility. *Econometrica*, 74:121–146, 2006.

[16] J. Hey. Does repetition improve consistency?,. *Experimental Economics*, 4:5–54, 2001.

[17] J. Hopcroft and R. Karp. An $n^{5/2}$ algorighm for maximum matchings in bipartite graphs. *SIAM Journal on Computing*, 2:225–231, 1973.

[18] J. Huber, J. Payne, and C. Puto. Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. *Journal of Consumer Research*, 9:90–98, 1982.

[19] G. Kalai, A. Rubinstein, and R. Spiegler. Rationalizing choice functions by multiple rationales. *Econometrica*, 70:2481–2488, 2002.

[20] Y. Kitamura and J. Stoye. Nonparametric analysis of random utility models. *Econometrica*, 86:1883–1909, 2018.

[21] I. Kopylov. Combinatorial models of subjective state spaces. Mimeo, UC Irvine, 2018.

[22] A. Liang. Inference of preference heterogeneity from choice data. *Journal of Economic Theory*, 179:275–311, 2019.

[23] D. Luce. *Individual Choice Behavior: A Theoretical Analysis*. Wiley, New York, 1959.

[24] R. Matzkin. Nonparametric estimation of nonadditive random functions. *Econometrica*, 71:1339–75, 2003.

[25] D. McFadden. Conditional logit analysis of qualitative choice behavior. In P. Zarembka, editor, *Frontiers in Economics*. Academic Press, New York, 1973.

[26] D. McFadden and K. Richter. Stochastic rationality and revealed stochastic preference. In J. Chipman, D. McFadden, and K. Richter, editors, *Preferences, Uncertainty and Rationality*, pages 161–186. Westview Press, Boulder, 1991.

[27] H. Moulin. Choice functions over a finite set: A summary. *Social Choice and Welfare*, 2:147–160, 1985.

[28] C. Plott. Path independence, rationality and social choice. *Econometrica*, 41:1075–1091, 1973.

[29] A. Sen. Quasi-transitivity, rational choice and collective decisions. *Review of Economic Studies*, 36:381–393, 1969.

[30] I. Sher, J. Fox, K. Kim, and P. Bajari. Partial identification of heterogeneity in preference orderings. NBER Working paper 17346, 2017.

[31] H. Uzawa. Note on preferences and axioms of choice. *Annals of the Institute of Statistical Mathematics*, 8:35–40, 1956.