Inference of Choice Correspondences*

Narayanaswamy Balakrishnan[†] Efe A. Ok[‡] Pietro Ortoleva[§]

December 7, 2022

Abstract

Despite being the fundamental primitive of the study of decision-making in economics, choice correspondences are not observable: even for a single menu of options, we observe at most one choice of an individual at a given point in time, as opposed to the *set* of all choices she deems most desirable in that menu. However, it may be possible to observe what a person chooses from a feasible menu at various times, repeatedly. We propose a method of inferring the choice correspondence of an individual from this sort of choice data. First, we derive our method axiomatically, assuming an ideal dataset. Next, we develop statistical techniques to implement this method for real-world situations where the sample is often small. A special case of this methodology allows for the estimation of individual preferences from repeated pairwise choice data. To demonstrate the applicability of the method, we use it on the data of a famous experiment (Tversky, 1969) on transitivity of preferences. We find that the conclusions this data leed to are more nuanced than the original ones.

JEL Classification: C81, D11, D12, D81.

Keywords: Choice Correspondences, Estimation, Multiple-Hypotheses Testing, Stochastic Choice Functions, Transitivity of Preferences

---

*A previous version circulated under the title "Inferential Choice Theory." We thank Paul Feldman, José Luis Montiel Olea and Gerelt Tserenjigmid for their valuable observations and comments on an earlier draft of this paper, and Mauricio Almeida Couri Riberio for his superb research assistance. In addition, Ok gratefully acknowledges the support of NSF grant NSF-1851610 and Ortoleva that of NSF grant SES- 2048947.

[†]Department of Mathematics and Statistics, McMaster University. E-mail: bala@mcmaster.ca.

[‡]Department of Economics and the Courant Institute of Applied Mathematics, New York University. E-mail: efe.ok@nyu.edu.

[§]Department of Economics and SPIA, Princeton University. E-mail: pietro.ortoleva@princeton.edu.

1 Introduction

At the core of revealed preference theory is the idea that, while preferences and utilities are unobservable constructs, choices are *observable*. An advantage is thus given to theories derived from choice, and the main primitive of this approach is the notion of *choice correspondence*, a function that maps each feasible menu of options to the set of choices from that menu. Most textbooks use choice correspondences as the starting point of microeconomic theory from which preferences, then utility functions, and then the entirety of economic analysis, are derived. Over the last century, choice correspondences proved to be exceptionally useful for the development of the theory of rational decision-making as well as its boundedly rational alternatives.

On closer scrutiny, however, one has to concede that choice correspondences are unobservable as well. After all, correspondences assigns to any given menu a *set* of choices, but it is practically impossible to observe such a set. The choice correspondence from the set $\{x, y\}$ may well be $\{x, y\}$, e.g., if the individual is indifferent or preferences are incomplete; but, at each trial, we can only observe her choose x or y, not both. Indeed, while the use of choice correspondences is often motivated in first-year Ph.D. courses as based on observability, this often leads to embarrassing questions—in our experience, invariably asked by some alert student—about how such set-valued functions can really be observed, and how their properties can actually be tested. Many prominent textbooks either avoid the discussion of this issue (thereby choosing to treat choice correspondences as theoretical abstractions), or simply assume it away by working only with single-valued choice correspondences.[1] A few suggest informal ways of thinking about choice correspondences as observable entities, but this never goes beyond offering a few passing sentences to this effect.[2]

The present paper introduces a general method of inferring choice correspondences using the data obtained through *repeated observations of choices* made by individuals. More

[1]Simplifying as it is, this latter approach not only fails solving the unobservability problem, it serves rather poorly as a foundation for even the most basic economic models of decision making. After all, if a choice correspondence is single-valued, it can never be rationalized by a preference relation that allows for indifference, thereby ruling out the standard model of consumer choice, as well as all non-degenerate cases of expected utility theory under risk, among others. In addition, it is known that single-valuedness hypothesis cannot be satisfied by *continuous* choice correspondences, unless one makes severe assumptions on the grand space of alternatives (cf. Nishimura and Ok, 2014).

[2]For instance, Mas-Colell, Whinston and Green (1995) says that the set of choices of an individual from a menu B "can be thought of as containing those alternatives that we would actually see chosen if the decision maker were repeatedly to face the problem of choosing an alternative from set B." In many ways, our work can be thought of as trying to extract formal content from this intuitive statement in a manner that allows one infer choice correspondences in practice.

precisely, we aim to "compute" a choice correspondence from the number of times each option is chosen by a person from a given menu when asked repeatedly. As such, our starting point is that what is observable comes in the form of a vector of relative frequencies (of choices), that is, in the form of stochastic choice data.[3] This data may originate from experiments—below, we show how our procedure can be applied to existing data—or from market behavior, exploiting the granular observation of choices that marketing firms have access to in the digital age.

Our Approach. We approach the problem in two stages. These are distinct from each other both procedurally and conceptually, and are primed to capture different aspects of the matter. In the first stage, we examine methods of constructing a choice correspondence if the analyst had access to the actual relative choice frequencies of an individual with perfect accuracy. In the second stage, we address the issue that real data includes only a finite (and often small) sample of observations, and suggest a statistical procedure to take this into account to infer an (empirical) choice correspondence for that person. We summarize what we actually do in these stages next.

Choice Imputation with Ideal Data. At the outset, we look at the problem at hand theoretically, assuming that the analyst has access to the probability $\mathbb{P}(x, S)$ with which each option x is chosen by a subject from a given menu S. In other words, we study the functions that map any given stochastic choice function $\mathbb{P}$ to a choice correspondence. We refer to any such function as a *choice imputation* (provided that it never declares an option with zero probability of being selected in a menu as a "choice" in that menu).

There are many interesting types of choice imputations. For example, we may declare a feasible option in S as a "choice" if that option has a positive probability of being chosen in S. But this is likely to be too permissive. After all, if the probability of x being chosen from $\{x, y\}$ is negligibly small, it may be reasonable to think of it as a "mistake," instead of a bona fide "choice." We may also go to the opposite extreme, and consider only the options with maximum likelihood of being chosen in a menu. But, obviously, this may well be too restrictive; for instance, it would not consider x as a choice from $\{x, y\}$ even if the probability of x being chosen is as high as .49. And, of course, there are many intermediate imputations that possess a less extreme makeup. No imputation is likely to be suitable in

[3]In the literature, the term *stochastic choice* is used to indicate the relative frequency of the choices of an individual in repeated trials, as well as across individuals when each person is observed only once. In this paper, we exclusively focus on multiple choices made by the same individual.

all contexts; there does not appear to be a reason *a priori* to work with any one specific imputation.

To address this issue systematically, we adopt an axiomatic approach, and consider some basic properties that characterize an interesting one-parameter family of choice imputations. A special element of this family corresponds to the idea mentioned in footnote 2, and maps any given stochastic choice function $\mathbb{P}$ to the choice correspondence that chooses in a menu S all alternatives x with $\mathbb{P}(x, S) > 0$. For any other element of this family, there exists a constant $\lambda \in (0, 1]$ such that any $\mathbb{P}$ is mapped to the choice correspondence that declares

$$\left\{x \in S : \mathbb{P}(x, S) \geq \lambda \max_{y \in S} \mathbb{P}(y, S)\right\}$$

as the set of "choices" from any menu S. We refer to any one of these choice imputations as a *Fishburn imputation*.[4] The family of Fishburn imputations includes the two examples discussed above, but it allows for many intermediate cases: For any $\lambda \in (0, 1)$, the associated imputation declares x as a choice in a menu S when its probability of being selected is higher than a factor (namely, λ) of the choice probability of any other option in S. It is worth noting that, for menus with at least three options, this is not the same as focusing on alternatives chosen with probability higher than a certain threshold.[5]

The value of λ here determines how exclusive the associated Fishburn imputation really is; we thus call it the *level of selectivity*. Low values of λ corresponds to inclusive imputations in which any choice with even a small choice probability is considered as a choice. The largest value $\lambda = 1$ corresponds to maximally exclusive imputations where only those alternatives with maximum likelihood of being chosen in a menu are qualified as "choices" in that menu. The decision of which level of selectivity λ to be adopted in a particular empirical application belongs to the analyst and should best be tailored to that application. (This is very much reminiscent of statistical parameters used in complex hypothesis testing models.) For example, in highly noisy environments—e.g., when choices are made under time pressure, under disturbance, or with low incentives—the analyst may adopt a higher λ, discarding options that have intermediate probability of being chosen. In other contexts, a lower level of selectivity may be more reasonable to adopt.

[4]This map was first introduced in Fishburn (1978), although with a different goal; see below.

[5]For any $\theta \in (0, 1)$, consider the function that maps any stochastic choice function $\mathbb{P}$ to the choice correspondence that declares $\{x \in S : \mathbb{P}(x, S) \geq \theta\}$ as the set of "choices" from any menu S. Easy examples show that this is distinct from any Fishburn imputation (unless the choice domain consists only of pairwise problems).

Choice Imputation with Real Data. In reality, of course, we do not observe $\mathbb{P}(\cdot, S)$ directly, but rather get information about it through finitely many observations. This brings us to the second stage of our construction: Even if we have decided to use a Fishburn imputation with a particular level of selectivity λ, our ultimate elicitation problem requires us decide whether or not to include an option x in the set of "choices" from S, given the empirical distribution of observed choices. This leads to the multiple-hypotheses testing problem whose null hypotheses are

$$H_x : \mathbb{P}(x, S) \geq \lambda \max_{y \in S} \mathbb{P}(y, S), \qquad x \in S.$$

At this junction, we adopt the standard assumption that choice trials are independent and the probability of choice of any alternative in S is the same in each trial. In addition to λ, the associated test procedure depends, of course, on the sample size n (i.e., the total number of times we see the agent choose from S), the number of times each x is chosen in S, and the level of control α for the family-wise error rate (to be chosen by the analyst). Given these parameters, we develop a statistical method to compute the set of all potential choices of the individual from the menu S as

$$\{x \in S : H_x \text{ is not rejected at the control level } \alpha\}$$

by building on the Benjamini-Hochberg multiple testing procedure. This method extends to inferring the values of the choice correspondence of the decision maker across multiple menus. As such, it assigns to every data set pertaining to repeated choice trials a particular choice correspondence that depends on the observed data, the number of repetitions, as well as two parameters chosen by the analyst, namely, the level of control α and the level of selectivity λ.

Inferring Preferences. While choice correspondences have a more foundational standing in economic theory, a very large majority of theoretical and experimental work on decision theory and behavioral economics are couched in terms of preference relations. Yet, from the empirical viewpoint, focusing on preferences are not less problematic than working with choice correspondences. If we have observed a person choose x over y in pairwise comparisons a certain number of times, and y over x in others, we need a rule to justify one of the following four possible conclusions: (i) "the subject prefers x over y strictly," (ii) "the subject prefers y over x strictly,", (iii) "the subject is indifferent between x and y," and

(iv) "the subject is unable to compare x and y." The methods developed in Section 2 and 3 provide a rule that is directly applicable to this sort of situations.

We develop our methodology below in terms of choice correspondences only because this is more comprehensive than focusing on individual preferences. The entirety of the present work, its theoretical as well as statistical parts, remains applicable if one is interested instead in inferring a person's preferences over a finite set of alternatives. For, this case is none other than determining one's choice correspondence on the domain of *pairwise* choice problems alone. In fact, for such a restricted domain, the statistical tests that we provide in Section 3 have closed-form descriptions, and possess a fair bit of (statistical) power.

Stochastic vs. Deterministic Choice. Before proceeding, we should directly address a potential doubt. If data comes in the form of an empirical stochastic choice function, one may be tempted to altogether abandon deterministic choice theory and focus exclusively on stochastic choice as a primitive, making use of the rich nature of this data. Indeed, this is studied by a vast and vibrant literature, a major avenue of research in decision-making.

Yet, this does not invalidate the need to properly construct deterministic choice. With rare exceptions, the *entirety* of economic analysis is built on deterministic choices, often summarized as preferences or utilities.[6] Unless we are ready to leave it all aside and focus solely on stochastic counterparts, an interesting but rather radical suggestion, one would have to derive deterministic behavior. Take, for example, Expected Utility Theory and Prospect Theory. Both are built on preference relations—deterministic constructs directly derived from choice correspondences (e.g., from doubleton sets). To test if behavior follows either model—a classic quest in the experimental literature—we need to construct the preference relation, as this is the language in which these theories are defined. Similarly, we need to derive preferences any time we want to test their transitivity (as we do in Section 4), stationary and exponential discounting for time preferences, altruism in social preferences, ambiguity aversion; if we want to test if behavior abides by standard notions in game theory—e.g., the Nash Equilibrium is defined on deterministic preferences; and choice correspondences are needed to test for rationality in the form of the Weak Axiom of Revealed Preferences, or one of the many models of bounded rationality, from satisficing to

[6]This is different than adopting models of stochastic choice to study the choices of heterogenous agents, where the stochasticity in the data derives from (unobserved) heterogeneity in preferences. The nature of the latter type of models is in fact more in line with deterministic choice theory. In those models, each agent makes a deterministic choice, but the analyst sees the choice data only in the aggregate, and hence evaluates it "as if" it is stochastic.

salience to reference-dependence.

In almost any such work, the properties to be tested are defined for deterministic preferences or choice correspondences, but insofar as the data is collected by observing the choices repeatedly, it comes in the form of an empirical stochastic choice function. By the nature of the problem, therefore, one needs to convert this data into a deterministic preference relation or choice correspondence, and the methodology we develop below is primed to address this need.

Moreover, summarizing a stochastic choice function in the form of a choice correspondence may help one better understand the structure of that function. The idea behind this is no different than using summary statistics to describe probability distributions, or inequality indices to study income distributions. In all these cases, using a "statistic" causes loss of information, but this is more than compensated by the transparency one gains about a particular aspect of the given distribution. In our case too "reducing" a stochastic choice function to a choice correpondence causes loss of information, but it may well help us uncover a hidden structure of that function. For instance, two empirical stochastic choice functions whose inferred choice correspondences are the same must share certain characteristics that may not at all be evident from raw data. Inferring choice correspondences may thus help classify repeated choice data at large.

A Case Study. To demonstrate how our method can be easily applied on real data, and that one does not need an unrealistic set of observations for it to work, we use it to reevaluate the choice behaviors observed in Experiment 1 of Tversky (1969), a classical and eminent example of stochastic choice data from doubleton menus.

As we discuss below, many papers in the literature on boundedly rational deterministic choice cites Tversky (1969) as providing evidence for *nontransitive preferences*, that is, cyclic *deterministic* choice behavior over doubleton menus. However, Tversky (1969) only reports relative choice frequencies—an (empirical) stochastic choice function—and uses them to test for violations of what is called Weak Stochastic Transitivity. He does not even discuss the issue of the potential acyclicity of (deterministic) imputed preference relations.

We apply our method to impute deterministic choice correspondences and test whether they are transitive. We find that for reasonable choices of parameters (like $\alpha = .05$ and $\lambda = .5$), the majority of imputed preference relations (about 62% of them) are actually transitive. Indeed, it is only with rather extreme parameters (very high levels of λ) that a majority of subjects is classified as non-transitive. At the very least, this suggests caution against using Tversky's data to motivate violations of WARP in deterministic environments.

Related Literature. While the issues of observability of choice correspondences are well-known, only a few papers have attempted to elicit them from data, and to our knowledge all have done so introducing novel experimental procedures instead of using standard (repeated) choice data. For instance, Bouacida (2019) asks subjects to choose from a set of alternatives, but allows them to choose multiple options and give them an additional (small) payment if they do so; in that case, the agent receives one of her choices randomly. Other papers use unincentivized additional questions after the choice to elicit the strength of preferences, and to identify indifferences and/or incomparabilities. Some of these papers, notably, Costa-Gomes et al. (2021), use this information to construct choice correspondences. Whether one believes these procedures to be effective, or that they in fact introduce additional confounds, it is plain that they are tailored for particular experiments, and hence cannot be applied to typical choice data. In particular, none of these procedures is applicable to data collected in past experiments with repeated choice trials, such as those we use in our empirical application below.

We should also note that some authors have used different experimental methods to elicit multiple choices, such as allowing subjects to use randomization devices[7] or choice deferrals.[8] As is well known, however, these methods provide information that is markedly different from that needed for deriving a choice correspondence.[9]

Finally, we emphasize that the literature on revealed preference theory based on choice from budget sets, which started with the seminal work of Afriat (1967), does take into account the issue of limited observability of choices and allows for the possibility of (unobserved) choice correspondences. The papers that belong to this strand treat each observation as a selection from the demand set of the agent at a given price configuration, and do not attempt to construct the entire demand set at the associated budget. (See Chambers and Echenique, 2016 for a review.) As such, they implicitly treat each observation as equally informative of one's demand correspondence, which relates it to the special case of Fishburn

[7]See Cohen, et al. (1985, 1987), Rubinstein (2002), Kircher, et al. (2013), Agranov and Ortoleva (2017, 2021), Dwenger et al. (2018), Miao and Zhong (2018), Cettolin and Riedl (2019), and Feldman and Rehbeck (2020).

[8]See Danan and Ziegelmeyer (2006), Sautua (2017), Costa-Gomes et al. (2021), Gerasimou (2021).

[9]For example, Agranov and Ortoleva (2017, 2020) observe that many subjects are willing to randomize between two alternatives x and y. This is conceptually distinct from saying that both alternatives belong to these subjects' choice sets from the menu $\{x, y\}$; instead, it suggests merely that some agents may prefer a particular randomization over the given options (as it would be the case, for instance, if they possessed quasiconcave utility functions over lotteries). See Cerreia-Vioglio, et al. (2019) for more on this.

imputations with $\lambda = 0$, where each element chosen, however infrequently, is considered to be a choice.

2 Imputation of Choice Correspondences

Throughout the paper, X stands for an arbitrarily fixed nonempty finite set with $|X| \geq 3$. We denote by $\mathfrak{X}$ the collection of all nonempty subsets of X, and by $\mathfrak{X}_2$ the collection of all subsets of X that contain exactly two elements.

2.1 Choice Imputations

Choice Correspondences. By a **choice correspondence** on $\mathfrak{X}$, we mean a set-valued map $C : \mathfrak{X} \to \mathfrak{X}$ such that $C(S) \subseteq S$. The standard (if a bit ambiguous) interpretation is that $C(S)$ includes all feasible alternatives that the individual deems worth choosing. (As we explained in the Introduction, giving empirical content to this interpretation is one of the main objectives of the present paper.) We denote the collection of all choice correspondences on $\mathfrak{X}$ by $\mathsf{cc}(X)$.

Stochastic Choice Functions. By a **stochastic choice function** on $\mathfrak{X}$, we mean a function $\mathbb{P} : X \times \mathfrak{X} \to [0, 1]$ such that

$$\sum_{x \in S} \mathbb{P}(x, S) = 1 \quad \text{and} \quad \mathbb{P}(y, S) = 0$$

for every $S \in \mathfrak{X}$ and $y \in X \backslash S$. The collection of all such functions is denoted by $\mathsf{scf}(X)$.

For any $\mathbb{P} \in \mathsf{scf}(X)$, the map $x \mapsto \mathbb{P}(x, S)$ defines a probability distribution on S. From an individualistic perspective, we interpret this distribution by imagining that a decision maker has been observed making choices from the feasible menu S multiple times, and the relative frequency of the times x is chosen from S is $\mathbb{P}(x, S)$ in the limit, as the number of observations tends to infinity. Thus, $\mathbb{P}(x, S)$ is not an observable quantity, just like the probability of getting heads in a particular coin toss is not observable. Instead, from the viewpoint of an outside observer, it is a random entity. Put informally, it is *approximately* observable in the sense that any choice experiment that tracks the choices of the agent from S repeatedly provides a sample wherein the empirical value of the relative frequency of the times x is chosen in S is a strongly consistent estimator of $\mathbb{P}(x, S)$.[10] As we discussed in the introduction, we will first consider below how to derive choices from $\mathbb{P}$ as if this function is known, and only later account for the unobservability of $\mathbb{P}$.

Before we get to work, here is one extra bit of notation: Given a stochastic choice function $\mathbb{P}$ and $S \in \mathfrak{X}$, we put

$$M_{\mathbb{P}}(S) := \max_{z \in S} \mathbb{P}(z, S) \quad \text{and} \quad m_{\mathbb{P}}(S) := \min_{z \in S} \mathbb{P}(z, S)$$

That is, $M_{\mathbb{P}}(S)$ and $m_{\mathbb{P}}(S)$ are the choice probabilities of items in S with the maximum and the minimum likelihood, respectively.

Choice Imputations. At the center of our analysis is a map that assigns a choice correspondence to any stochastic choice function $\mathbb{P}$, that is, a map of the form

$$\Psi : \mathsf{scf}(X) \to \mathsf{cc}(X).$$

The only condition that we impose on this map at the outset is that

$$\mathbb{P}(x, S) = 0 \quad \text{implies} \quad x \notin \Psi(\mathbb{P})(S) \tag{1}$$

for any $S \in \mathfrak{X}$ and $\mathbb{P} \in \mathsf{scf}(X)$. This condition forbids designating an item that is never chosen in S as a choice from S. We refer to any Ψ that satisfies this property as a **choice imputation**. In words, a choice imputation Ψ is a method of transforming the behavior of an individual represented by a stochastic choice function into a deterministic choice correspondence. Loosely speaking, we wish this method to associate a choice correspondence $C_{\mathbb{P}}$ to $\mathbb{P}$ in such a way that, for any menu S, the set $C_{\mathbb{P}}(S)$ consists of all items in S that have a "significant" probability of being chosen in S, eliminating, for instance, items that are chosen by mistake, in a rush, etc..

Example 1. An interesting choice imputation is one that includes anything chosen with positive probability in the associated choice set. Formally, this imputation, which we denote by Ψ_0, maps any $\mathbb{P} \in \mathsf{scf}(X)$ to the choice correspondence $C_{\mathbb{P},0}$ on $\mathfrak{X}$ defined by

$$C_{\mathbb{P},0}(S) := \{x \in S : \mathbb{P}(x, S) > 0\}.$$

[10] This statement is readily formalized by means of the Glivenko-Cantelli Theorem.

The choice imputation of Example 1 provides a natural starting point, and indeed, it is implicitly suggested by Mas-Colell, Whinston and Green (1995); see footnote 2. Nevertheless, it appears too inclusive. It is arguable that if x is chosen from $\{x, y\}$ with probability 0.001, it should probably not be included in the choice set of the agent at the menu $\{x, y\}$.

Example 2. For any menu S, one may wish to include in the set of all choices only those options in S with the maximum probability of being chosen. This method is captured by the choice imputation Ψ_1 which maps any $\mathbb{P} \in \mathsf{scf}(X)$ to the choice correspondence $C_{\mathbb{P},1}$ on $\mathfrak{X}$ defined by

$$C_{\mathbb{P},1}(S) := \{x \in S : \mathbb{P}(x, S) \geq \mathbb{P}(y, S) \text{ for all } y \in S\}.$$

The choice imputations Ψ_0 and Ψ_1 are extreme members of an interesting one-parameter family.

Example 3. For any $\mathbb{P} \in \mathsf{scf}(X)$ and $\lambda \in (0, 1]$, define the map $C_{\mathbb{P},\lambda} : \mathfrak{X} \to \mathfrak{X}$ by

$$C_{\mathbb{P},\lambda}(S) := \{x \in S : \mathbb{P}(x, S) \geq \lambda M_{\mathbb{P}}(S)\}.$$

In words, $C_{\mathbb{P},\lambda}(S)$ contains a feasible alternative $x \in S$ iff there are no alternatives in S that are chosen at least $\frac{1}{\lambda}$ times more frequently than x. For λ close to 1, it seems unexceptionable that we qualify the members of $C_{\mathbb{P},\lambda}(S)$ as potential "choices" of the agent from S; conversely, for λ close to 0, it makes sense to think of the members of $S \backslash C_{\mathbb{P},\lambda}(S)$ as objects that are chosen due to occasional mistakes. For any given $\lambda \in [0, 1]$, the map Ψ_λ defined on $\mathsf{scf}(X)$ by $\Psi_\lambda(\mathbb{P}) := C_{\mathbb{P},\lambda}$ is a choice imputation (where $C_{\mathbb{P},0}$ is defined in Example 1). We refer to this as a **Fishburn imputation** with level of selectivity λ.[11]

When a set includes only two items, say x and y, it is easy to see that $x \in C_{\mathbb{P},\lambda}\{x, y\}$ iff $\mathbb{P}(x, \{x, y\}) \geq \frac{\lambda}{1+\lambda}$. That is, an item is deemed as a "choice" in a doubleton menu iff it is chosen with a probability above a fixed threshold. For larger sets, however, Fishburn imputations are more complex. Whether or not x is included in $C_{\mathbb{P},\lambda}(S)$ depends not only on the probability $\mathbb{P}(x, S)$, but also on the highest choice probability in S, namely, $M_{\mathbb{P}}(S)$. For a given probability of choosing x in S, that option will be included in the set of choices from S only when the maximum probability of choice is not too high. Thus, the criterion is

[11] The correspondences $C_{\mathbb{P},\lambda}$ were first considered by Fishburn (1978) who sought the characterization of $\mathbb{P}$ such that for every $\lambda \in [0, 1]$, there is a (utility) function $u_\lambda : X \to \mathbb{R}$ with $C_{\mathbb{P},\lambda}(S) = \arg\max u_\lambda(S)$ for every $S \in \mathfrak{X}$. In turn, they were recently used by Ok and Tserenjigmid (2020, 2022) to produce rationality criteria for stochastic choice rules.

more selective for sets in which some option is chosen with a very high chance, less selective for sets in which all options are chosen with low probability.[12]

The extent of selectivity of a Fishburn imputation depends on the value of λ, which is to be chosen by the analyst for the problem at hand. It may be reasonable to pick higher values of λ—that is, a more selective criterion—for environments in which there is noise, or more generally, when mistakes are expected. In those cases one may wish to disregard options that are not chosen with sufficiently high probability. On the other hand, lower values of λ may be appropriate when there is reason to consider objects chosen with low probability as genuine selections as well.

Remark 1. It may be of theoretical interest to compute the Fishburn imputations of some well-known models of stochastic choice. To illustrate, take any two real injective maps on X and any map $\theta : \mathfrak{X} \to (0,1)$. The stochastic choice function $\mathbb{P}$ on X where

$$\mathbb{P}(x,S) := \theta(S)1_{\arg\max u(S)}(x) + (1-\theta(S))1_{\arg\max v(S)}(x)$$

for any $x \in S$ and $S \in \mathfrak{X}$, is said to be a **dual random utility model**; this model has been nicely characterized by Manzini and Mariotti (2018). For any $S \in \mathfrak{X}$ and any injection $f : X \to \mathbb{R}$, let $x(f,S)$ stand for the (unique) maximizer of f in S. Then, the Fishburn imputation of $\mathbb{P}$ (with level of selectivity λ) is readily computed. Assuming wlog. $\theta(S) \geq \frac{1}{2}$, then $\Psi_\lambda(\mathbb{P})(S) = \{x(u,S)\}$ when $\frac{1-\theta(S)}{\theta(S)} < \lambda$ and $\Psi_\lambda(\mathbb{P})(S) = \{x(u,S), x(v,S)\}$ otherwise. □

In this paper we mostly work with Fishburn imputations, but there are several other types of choice imputations that may be useful in empirical work. For good measure, we next present a selection of such alternatives.

Example 4. Consider the map $\Psi : \mathsf{scf}(X) \to \mathsf{cc}(X)$ with

$$\Psi(\mathbb{P})(S) := \{x \in S : \mathbb{P}(x,S) \geq \min\{\theta, M_{\mathbb{P}}(S)\}\}, \qquad S \in \mathfrak{X}$$

for some $\theta \in (0,1)$. In any menu S, the map Ψ includes an item as a "choice" if either that item is chosen with a probability above a threshold θ, or if it is the item with the maximum

[12]It is thus easy to see that this criterion, like most models of stochastic choice, is not immune to adding duplicates into a menu. For example, consider x, y, y' where y and y' are duplicates, and suppose $\mathbb{P}(x,\{x,y\}) = 0.2$, $\mathbb{P}(y,\{x,y\}) = 0.8$, while $\mathbb{P}(x,\{x,y,y'\}) = 0.2$, $\mathbb{P}(y,\{x,y,y'\}) = 0.4$, and $\mathbb{P}(y',\{x,y,y'\}) = 0.4$ (which may be natural since y and y' are duplicates). Then for certain values of λ we have $x \in C_{\mathbb{P},\lambda}(\{x,y\})$ but $x \notin C_{\mathbb{P},\lambda}(\{x,y,y'\})$, that is, adding duplicates affects the selection. Depending on one's view of mistakes, this may be a feature or a concern. In either case, this is shared by most models on stochastic choice data, and simply calls for the careful identification of choice items.

likelihood of being chosen in S. Ψ is then a choice imputation, but it is not a Fishburn imputation.

Example 5. Define $c_{\mathbb{P}}(S) := C_{\mathbb{P},1}(S) \cup C_{\mathbb{P},1}(S \backslash C_{\mathbb{P},1}(S))$ (with the convention that $C_{\mathbb{P},1}(\varnothing) := \varnothing$) for any $S \in \mathfrak{X}$ and $\mathbb{P} \in \mathsf{scf}(X)$. (Thus, $c_{\mathbb{P}}(S)$ contains all alternatives in S that are the most probably chosen in S as well as those that are second most probably chosen.) The map Ψ defined on $\mathsf{scf}(X)$ by $\Psi(\mathbb{P}) := c_{\mathbb{P}}$ is a choice imputation (but again, it is not a Fishburn imputation).

Example 6. Let $\lambda : \mathbb{N} \to (0, 1]$ be any non-constant decreasing function, and consider the map $\Psi : \mathsf{scf}(X) \to \mathsf{cc}(X)$ with

$$\Psi(\mathbb{P})(S) := \{x \in S : \mathbb{P}(x, S) \geq \lambda(|S|) M_{\mathbb{P}}(S)\}, \qquad S \in \mathfrak{X}.$$

This is a choice imputation that acts over menus with the same size just like a Fishburn imputation, but it may use different factors over menus with different cardinalities. It may be useful if one subscribes to the view that it gets harder to achieve a given fraction of the maximum likelihood in larger menus.[13]

2.2 Foundations

Recall that in the first step of our analysis we presume that the actual stochastic choice function $\mathbb{P}$ of a given individual is known. (Or, equivalently, in this first step we suppose that the data is perfectly informative about the true stochastic choice function of the decision maker.) As such, our problem is to decide which sort of choice imputation to use in order to transform $\mathbb{P}$ into a deterministic choice correspondence.

It is plain that every choice imputation has its advantages and disadvantages. We thus start our analysis by looking at axiomatic ways of evaluating such procedures in the abstract. The postulates below are imposed on an arbitrarily given choice imputation Ψ; for ease of notation, we denote the value of Ψ at $\mathbb{P}$ by $C_{\mathbb{P}}$.

[13] Suppose we set $\lambda = \frac{1}{2}$, and consider the two menus $S := \{x, y\}$ and $T := \{x, z_1, ..., z_6, y\}$. If $\mathbb{P}(x, S) = \frac{1}{3}$ while $\mathbb{P}(x, T) = \mathbb{P}(z_1, T) = \cdots = \mathbb{P}(z_6, T) = 0.1$, we have $\mathbb{P}(x, S) \geq \frac{1}{2}\mathbb{P}(y, S)$ and $\mathbb{P}(x, T) < \frac{1}{2}\mathbb{P}(y, T)$, so the Fishburn imputation $\Psi_{1/2}$ deems x as a choice from S, but not from T, while one may argue that the latter conclusion is not acceptable, and choose to use a factor less than $\frac{1}{3}$ to work with in the context of menus that contain more than two alternatives.

Anonymity. We begin by positing that if the choice probability distributions of two individuals with stochastic choice functions $\mathbb{P}$ and $\mathbb{Q}$ over a menu S are identical, then the value of the choice correspondences we attribute to them should be equal to each other at S. This seems like a reasonable property to impose on a choice imputation.

A. Anonymity. For every $\mathbb{P}, \mathbb{Q} \in \mathrm{scf}(X)$ and $S \in \mathfrak{X}$,

$$\mathbb{P}(\cdot, S) = \mathbb{Q}(\cdot, S) \quad \text{implies} \quad C_{\mathbb{P}}(S) = C_{\mathbb{Q}}(S).$$

A choice imputation that satisfies this property uses only the information about the choice behavior of a person in a menu S to infer her set of choices from that menu. While it constitutes a natural starting point, it would not be suitable for a method of imputation that looks at the choice behavior of an individual across all menus to decide how low a low a choice probability in S should be to count as a "mistake."

Imputations for Pairwise Choices. Suppose that, for a stochastic choice function $\mathbb{P}$ on X, we have somehow deemed the choice probability $\mathbb{P}(x, \{x, y\})$ large enough to include x in the choice set $C_{\mathbb{P}}\{x, y\}$. Let $\{z, w\}$ be another menu, and suppose z is primed to be chosen from $\{z, w\}$ even more frequently than x is from $\{x, y\}$, that is, $\mathbb{P}(z, \{z, w\}) \geq \mathbb{P}(x, \{x, y\})$. Since we have declared $x \in C_{\mathbb{P}}\{x, y\}$, consistency demands that we also declare $z \in C_{\mathbb{P}}\{z, w\}$. That is:

B. Monotonicity across Pairwise Choice Data. For every $\mathbb{P} \in \mathrm{scf}(X)$ and $S, T \in \mathfrak{X}_2$,

$$x \in C_{\mathbb{P}}(S) \text{ and } \mathbb{P}(z, T) \geq \mathbb{P}(x, S) \quad \text{imply} \quad z \in C_{\mathbb{P}}(T).$$

We next consider a fairly weak form of continuity.

C. Continuity on Pairwise Menus. For every $\mathbb{P}, \mathbb{P}_1, \mathbb{P}_2, \ldots \in \mathrm{scf}(X)$ and $S \in \mathfrak{X}_2$ with $S = C_{\mathbb{P}_k}(S)$ for each $k = 1, 2, \ldots$,

$$m_{\mathbb{P}_k}(S) \to m_{\mathbb{P}}(S) > 0 \quad \text{implies} \quad S = C_{\mathbb{P}}(S).$$

In words, given a doubleton menu S, if the choice sets imputed from each term of a sequence of stochastic choice functions include both elements of S, and if the associated smallest choice probabilities converge to a strictly positive value, then both elements of S should be included in the set of choices from S in the limit as well.[14]

The properties A, B and C are so basic that it seems unlikely they would impose much discipline on choice imputations. However, our first result shows that these conditions pin down the structure of a choice imputation on primitive choice problems up to a single parameter.

Theorem 1. Let $\Psi : \mathrm{scf}(X) \to \mathrm{cc}(X)$ be a choice imputation that satisfies the properties A, B, and C. Then, there exists a (unique) $\lambda \in [0, 1]$ such that

$$\Psi(\mathbb{P})(S) = C_{\mathbb{P},\lambda}(S) \quad \text{for every } S \in \mathfrak{X}_2 \text{ and } \mathbb{P} \in \mathrm{scf}(X).$$

In other words, every choice imputation that satisfies properties A, B and C must act like a Fishburn imputation on pairwise choice problems. This is noteworthy, because a great majority of choice experiments in the literature are about individual preferences and thus present the subjects only with pairwise choice problems. Theorem 1 gives a fairly strong reason for using Fishburn imputations for inferring the preferences of the subjects in such experiments.

Independence of Irrelevant Alternatives. While the properties A, B and C restrain the behavior of a choice imputation with respect to pairwise choice situations, the following property controls that behavior on arbitrary menus, and it does this by forcing the imputation be consistent with that used for pairwise choice problems.

To illustrate, suppose a person chooses x from a menu S with probability 0.2, while the item she chooses from S with the maximum likelihood is y, and that with probability 0.6. Suppose we also know that this individual chooses x against y in pairwise comparisons 25 percent of the time, that is, $\mathbb{P}(x, \{x, y\}) = .25$ and $\mathbb{P}(y, \{x, y\}) = .75$. How should then $C_{\mathbb{P}}(S)$ relate to $C_{\mathbb{P}}\{x, y\}$? Observe that we have chosen our numbers so that the relative probability of choosing x against y in the set S is the same as that in $\{x, y\}$ (for $\frac{.2}{.6} = \frac{.25}{.75}$). Thus, if we wish to abide by the principle of *Independence of Irrelevant Alternatives* (as formulated by, say, Luce, 1959), it would be natural to include x in $C_{\mathbb{P}}(S)$ if, and only if, $x \in C_{\mathbb{P}}\{x, y\}$. This principle maintains that an analyst may or may not find $\mathbb{P}(x, S)$ too small to include x in

[14] The condition $m_{\mathbb{P}}(S) > 0$ is essential in the formulation of this property. For instance, if (x_k) and (y_k) are two sequences in X such that $\mathbb{P}_k(x, \{x, y\}) = \frac{1}{k}$ for each k, and $\mathbb{P}(x, \{x, y\}) = 0$, we do not wish $C_{\mathbb{P}}\{x, y\}$ to include x (otherwise $\mathbb{P} \mapsto C_{\mathbb{P}}$ would not define a choice imputation; recall (1)).

$C_{\mathbb{P}}(S)$, but whatever is her decision in this regard, it should be the same in the context of $C_{\mathbb{P}}\{x, y\}$.

D. *Independence of Irrelevant Alternatives.* For every $\mathbb{P} \in \mathsf{scf}(X)$, $S \in \mathfrak{X}$, and $x, y \in S$ such that $\mathbb{P}(y, S) = M_{\mathbb{P}}(S)$ and

$$\frac{\mathbb{P}(x, \{x, y\})}{\mathbb{P}(y, \{x, y\})} = \frac{\mathbb{P}(x, S)}{\mathbb{P}(y, S)},$$

we have $x \in C_{\mathbb{P}}(S)$ iff $x \in C_{\mathbb{P}}\{x, y\}$.

Independence of Irrelevant Alternatives type axioms (such as Luce's Choice Axiom) are much discussed in the literature on stochastic choice. They typically lead to ratio-scale representations and easy-to-use formulae in making probabilistic computations. They rule out various ways in which choices may be menu-dependent and are well-known to be sensitive to the presence of perfectly substitutable alternatives. Having said that, it seems desirable to explore the consequences of Property D before entertaining its relaxations that allow for menu dependent methods of choice imputation.

Remark 2. The four postulates we have considered above are logically independent. In particular, any of the choice imputations presented in Examples 4, 5 and 6 satisfies properties A, B and C, but not D. □

A Characterization Theorem. We are now ready to state our main characterization result.

Theorem 2. A choice imputation $\Psi : \mathsf{scf}(X) \to \mathsf{cc}(X)$ satisfies the properties A, B, C, and D if, and only if, it is a Fishburn imputation, that is, there exists a (unique) $\lambda \in [0, 1]$ such that $\Psi(\mathbb{P}) = C_{\mathbb{P},\lambda}$ for every $\mathbb{P} \in \mathsf{scf}(X)$.

Our postulates thus characterize the family of Fishburn imputations completely. For any choice imputation Ψ that satisfies these properties, there exists a unique $\lambda \in [0, 1]$, which we dub the *level of selectivity*, such that an item x is included in $C_{\mathbb{P}}(S)$ iff it is chosen with a probability at least as high as λ times the choice probability of an option in S with the maximum likelihood of being chosen, and this, for any menu S. As discussed above, different levels of selectivity λ make the imputation more or less inclusive, and should be chosen by the analyst according to the problem at hand.

3 Imputation of Choice from Sample Data

We have so far acted as if $\mathbb{P}$ is known, and looked at methods of deducing a choice correspondence from $\mathbb{P}$. In reality, however, all we have is a data set that reports a person's choice frequencies per menu for a finite, and often fairly small, number of observations. We must thus account for sampling errors when inferring one's choice correspondence from her *empirical* stochastic choice function. This section is devoted to this issue.

We begin with setting up the general statistical inference problem at hand. We then confine our attention to inferring the behavior of a choice correspondence over *pairwise* choice situations by means of a suitable hypothesis-testing procedure. We pay heed to this special case because almost all repeated choice experiments in the literature use doubleton menus, and the idea behind Fishburn imputations is particularly agreeable in that case (Theorem 1). Moreover, we shall find that in this case our procedure has a closed-form, and it is quite easy to implement. In the latter part of the section, we turn to the case of menus of arbitrary size, and observe that much of what we are able to do with doubleton menus can be extended to that case. However, by necessity, the associated test procedure is then more complicated.

3.1 The General Hypothesis Testing Problem

To explain the nature of the statistical problem at hand formally, let $S_1, ..., S_m$ be m many menus in $\mathfrak{X}$. For each i, suppose we have observed an individual make a choice from the menu S_i, n_i many times. Then, the data at hand is in the form of *realizations* of the random variables

$$L_{n_i}(x, S_i) := \text{the number of times } x \text{ is chosen in } S_i \text{ in } n_i \text{ observations}$$

where $x \in S_i$ and $i = 1, ..., m$. Our problem is to decide which elements x of each S_i should be included among the choices from that menu, given the realization of these random variables. For each i, let $\ell_{n_i}(\cdot, S_i) := n_i^{-1} L_{n_i}(\cdot, S_i)$, and note that every realization of $\ell_{n_i}(\cdot, S_i)$ is a probability distribution over the contents of S_i, but this distribution may be a poor representative of $\mathbb{P}(\cdot, S_i)$, especially when the sample size n_i is small. Thus, even if we have decided which choice imputation Ψ to employ if $\mathbb{P}$ were known, approaching the problem of eliciting the choices of the individual from the menu S_i simply by applying Ψ at the given realization of $\ell_{n_i}(\cdot, S_i)$ may not properly account for sampling errors.

Let us enumerate S_i as $\{x_{i,1}, ..., x_{i,k_i}\}$ for each $i = 1, ..., m$. In the abstract, once a certain choice imputation Ψ is agreed upon, we are confronted with a general multiple-hypotheses testing problem whose null hypotheses are

$$\begin{array}{ccc} H_{1,1} : x_{1,1} \in C_{\mathbb{P}}(S_1) & \cdots & H_{1,k_1} : x_{1,k_1} \in C_{\mathbb{P}}(S_1) \\ \cdots & \cdots & \cdots \\ H_{m,1} : x_{m,1} \in C_{\mathbb{P}}(S_m) & \cdots & H_{m,k_m} : x_{m,k_m} \in C_{\mathbb{P}}(S_m) \end{array} \tag{2}$$

where, of course, $C_{\mathbb{P}} = \Psi(\mathbb{P})(\cdot)$.[15] Clearly, we need to make a sampling assumption to deal with this problem. We adopt the following standard postulate in this regard, which is maintained in the remainder of the paper.

> ***Assumption 1.*** The choice trials are independent and the probability of choice of any alternative in a menu is the same in each trial.

In addition, we need to specify the level of control we wish to impose on the *family-wise error rate* (FWER) or the *false discovery rate* (FDR) for the problem (2).[16] Let us denote this level by α, which is usually set as .05 or .1. (When there is only one hypothesis, α is none other than the *level of significance* of the test.) Once this is done, we have all the ingredients we need, namely, a realization of each $L_{n_i}(\cdot, S_i)$ (the choice data from each menu in the sample), Ψ (the method of imputation) and α (the level of control), to impute the value of the choice correspondence at each S_i statistically. Put precisely, given these ingredients and the choice data, the set of choices of the person from each menu S_i is determined simply as the set of all $x_{i,j}$ (where $j = 1, ..., k_i$) for which $H_{i,j}$ is not rejected at the control level α.

In view of the analysis presented in Section 2, we study this problem in more concrete terms by choosing Ψ to be a Fishburn imputation. Thus, for a fixed level of selectivity $\lambda \in [0, 1]$ to be chosen by the analyst, our multiple-hypotheses testing problem is

$$\begin{array}{ccc} H_{1,1} : \mathbb{P}(x_{1,1}, S_1) \geq \lambda M_{\mathbb{P}}(S_1), & \cdots & H_{1,k_1} : \mathbb{P}(x_{1,k_1}, S_1) \geq \lambda M_{\mathbb{P}}(S_1) \\ \cdots & \cdots & \cdots \\ H_{m,1} : \mathbb{P}(x_{m,1}, S_m) \geq \lambda M_{\mathbb{P}}(S_m), & \cdots & H_{m,k_m} : \mathbb{P}(x_{m,k_m}, S_m) \geq \lambda M_{\mathbb{P}}(S_m). \end{array} \tag{3}$$

Its simultaneous nature makes this testing problem nontrivial. One can of course disregard

[15] Depending on the size of the menus, and the context, one may be able to reduce the number of hypotheses here to enhance the testing procedure. See, for instance, Section 3.2.3 below.

[16] All terms we use in this paper that pertain to multiple hypotheses testing are defined and explained in detail in Lehmann and Romano (2005, Chapter 9).

the issue of multiplicity, and instead test each hypothesis in (3) in isolation at the significance level α. Unfortunately, as is well known, under this method the probability of false rejections rapidly increase with the total number of the tests (i.e., $k_1 + \cdots + k_m$).[17] Perhaps the simplest resolution of this issue is by using the so-called *Bonferroni correction*, which means we test each hypothesis at the significance level $\alpha/(k_1 + \cdots + k_m)$. While this procedure ensures that FWER is at most α, its ability to reject any of the null hypotheses is low, so it is rarely used. Instead, it is common to adopt a step-down testing approach for this purpose, such as the *Holm procedure* (of Holm, 1979) which also controls FWER at level α, or the Benjamini-Hochberg procedure (of Benjamini and Hochberg, 1995) which controls FDR (but not necessarily FWER) at level α. In what follows, we shall adopt the latter procedure due to its superior ability of rejecting null hypotheses.

Regardless of the procedure, however, one always runs into unacceptably high rates of acceptance of the null hypotheses when $k_1 + \cdots + k_m$ is large. This is an empirical issue and there is not much one can do about it at a theoretical level. Having said this, in the present context, there is at least one simplification that would reduce the multiplicity of the problem (3) to some extent:

> ***Assumption 2.*** For each i, any element of $\arg\max L_{n_i}(\cdot, S_i)$ is included in the choice set at S, and hence all hypotheses in (3) corresponding to these elements are deleted from the system.

We thus always qualify a most frequently chosen alternative in a menu as a "choice" from that menu. Obviously, if one wishes to infer nonempty-valued choice correspondences in an experiment, this assumption is necessary. Even when the researcher may be content with designating $\varnothing$ as the value of the choice correspondence at a menu (perhaps because the observed choices do not support any particular alternative), Assumption 2 is basically unexceptionable for menus of small size (which is, again, the relevant case for choice experiments in practice). For instance, if S is a menu with two elements, and an alternative y in S is chosen at least 50% of the time, throughout the observation period, it would be absurd not to include y as a choice in S. Similarly, if S contains three elements, and y is chosen most frequently in S, it is only natural to include y in the choice set at S. We will

[17]For instance, if we have two independent tests, the upper bound for (i.e., the control of) the probability of at least one false rejection is $2\alpha - \alpha^2$, for three independent tests this bound is $3\alpha - 3\alpha^2 + \alpha^3$ and so on. To get a better sense of the problem, set $\alpha = .05$. Then, if we have 5 independent tests to perform, the probability of at least one false rejection would be controlled at .23. With 50 independent tests, this number increases to .92. See Lehmann and Romano (2005, pp. 348-349).

see below that in addition to its straightforward appeal, Assumption 2 may at times achieve a substantial simplification of the test procedure. In particular, it allows us to reduce the multiplicity of the tests in (3) by at least m hypotheses.[18]

The next two subsections are devoted to the analysis of the tests (3) under the Assumptions 1 and 2. But before we move to this analysis, it may be useful to take stock. The overall approach we propose in this paper decomposes the elicitation of a choice correspondence (from data) into two stages. First, a decision is made as to which choice imputation one would use if $\mathbb{P}$ were observable. In particular, if one is set on using a Fishburn imputation, a value for the factor λ is chosen (but of course one may choose to carry out the procedure in terms of several choices for λ). Second, one deals with issues that arise due to the finiteness of data sets by means of a statistical procedure that accounts for sampling errors. These two stages, and hence the choices for λ and α are kept separate as they pertain to different domains and are conceptually distinct.

An alternative approach would be to use the realization of ℓ_n itself as a stochastic choice function to compute $\Psi(\ell_n)(S)$ directly for a suitable Ψ. Indeed, one can show that this is precisely what the maximum likelihood estimation would entail, at least in the case of pairwise choice situations. However, while the simplicity of this approach is surely an advantage, it is simply too coarse to account for sampling errors properly.

3.2 The Case of Pairwise Choice Situations

We now examine the multiple testing problem (3) in the special, but empirically common, case of pairwise choice situations. As we have discussed earlier, this case is particularly important, because it is the one that captures the *estimation of preferences*.

3.2.1 Inferring the Set of Choices from a Single Menu

Let us begin with the simplest case of inferring the value of a choice correspondence at a single doubleton menu. We thus fix an arbitrary $S \in \mathfrak{X}_2$, which we enumerate as $\{x, y\}$, and suppose λ and α are chosen. Our problem is to decide whether or not we should declare either x or y as a choice from the menu S given the realizations of $L_n(x, S)$ and $L_n(y, S)$ that pertains to a particular decision maker, where n is the number of times we have observed

[18]To wit, this means that the Bonferroni correction would then be at most $\alpha/(k_1 + \cdots + k_m - m)$ instead of $\alpha/(k_1 + \cdots + k_m)$, alleviating the overcorrection entailed by that procedure. The same observation applies also to all step-down multiple-hypotheses testing procedures.

this person make a choice from S. In what follows, we denote the actual probability that x is chosen from S by π, that is, $\pi := \mathbb{P}(x, S)$. Moreover, relabelling if necessary, we assume that y is chosen from S more frequently than x in the data, that is, $L_n(x, S) \leq L_n(y, S)$.

In this setup, (3) takes a very simple form:

$$H_x : \pi \geq \lambda M_{\mathbb{P}}(S) \quad \text{and} \quad H_y : 1 - \pi \geq \lambda M_{\mathbb{P}}(S).$$

Besides, by Assumption 2, this reduces to a single hypothesis test with the null hypothesis $H_x : \pi \geq \lambda M_{\mathbb{P}}(S)$. On the other hand, a routine manipulation shows that $\pi \geq \lambda M_{\mathbb{P}}(S)$ iff $\pi \geq \frac{\lambda}{1+\lambda}$. Consequently, our null hypothesis in this instance can be stated simply as:

$$H : \pi \geq \tfrac{\lambda}{1+\lambda}.$$

Now, under Assumption 1, $L_n(x, S)$ is binomially distributed with parameters n and π. Using this fact, we wish to obtain a threshold test of the form: Reject H if the realization of $L_n(x, S)$ is smaller than some suitably chosen nonnegative integer. To this end, we recall the following elementary, and well-known, property of the binomial distribution.

Lemma 3. Let n be a positive integer, $\theta \in \{0, ..., n\}$, and $\gamma_1, \gamma_2 \in [0, 1]$. For any two binomially distributed random variables u and v with parameters n and γ_1, and n and γ_2, respectively, we have $\text{Prob}(v \leq \theta) < \text{Prob}(u \leq \theta)$ whenever $\gamma_1 < \gamma_2$.

For the null hypothesis $\pi = a$, where a is some number larger than $\frac{\lambda}{1+\lambda}$, the p-value is $\text{Prob}(u \leq L_n(x, S))$ where $u \sim \text{Binom}(n, a)$. Lemma 3 says that the largest of these p-values over all $a \geq \frac{\lambda}{1+\lambda}$ is obtained precisely when $a = \frac{\lambda}{1+\lambda}$. In other words, all of these p-values are less than α iff the p-value associated with the null hypothesis

$$H : \pi = \tfrac{\lambda}{1+\lambda}$$

is less than α. Put precisely, given the realization of $L_n(x, S)$, this p-value is given as

$$p := \text{Prob}(u \leq L_n(x, S)) \quad \text{where} \quad u \sim \text{Binom}(n, \tfrac{\lambda}{1+\lambda}). \tag{4}$$

Then, where $c_{\lambda,\alpha}(S)$ denotes the set of inferred "choices" from S, our procedure simply says:

Do not include x in $c_{\lambda,\alpha}(S)$ if $p \leq \alpha$ for the realization of $L_n(x, S)$.

(We do not make the dependence of $c_{\lambda,\alpha}(S)$ on n here only not to clutter the notation.) That is, given the observed value of $L_n(x, S)$, our inferred choice set at S is

$$c_{\lambda,\alpha}(S) = \begin{cases} \{x, y\}, & \text{if } \mathrm{Prob}(u \leq L_n(x, S)) > \alpha \\ \{y\}, & \text{otherwise,} \end{cases} \tag{5}$$

where u is the binomially distributed random variable with parameters n and $\frac{\lambda}{1+\lambda}$.

Alternatively, given λ and α, let $w_{\lambda,n}$ be the maximum number of times we need to observe x being chosen so as to reject the hypothesis that x should be included in $c_{\lambda,\alpha}(S)$. That is,

$$w_{\lambda,n} := \max\left\{\theta \in \{0, ..., n\} : \mathrm{Prob}(u \leq \theta) \leq \alpha\right\} \tag{6}$$

where $u \sim \mathrm{Binom}(n, \frac{\lambda}{1+\lambda})$. Then,

$$c_{\lambda,\alpha}(S) = \begin{cases} \{x, y\}, & \text{if } L_n(x, S) > w_{\lambda,n} \\ \{y\}, & \text{otherwise.} \end{cases} \tag{7}$$

The p-value in (4), as well as the threshold $w_{\lambda,n}$, and hence $c_{\lambda,\alpha}(S)$, are very easy to compute in practice.

Remark 3. With probability at least $1 - \alpha$, the procedure above infers a choice set that contains the "true" choices of the subject. That is,

$$\mathrm{Prob}(C_{\lambda,\mathbb{P}}(S) \subseteq c_{\lambda,\alpha}(S)) > 1 - \alpha. \tag{8}$$

Indeed, by (5), the probability of the event "$C_{\lambda,\mathbb{P}}(S) \nsubseteq c_{\lambda,\alpha}(S)$" is $\mathrm{Prob}(L_n(x, S) \in \{0, ..., w_{\lambda,n}\})$ (i.e., the probability that x is not included in $c_{\lambda,\alpha}(S)$) *when* $1-\pi = \mathbb{P}(y, S) < \lambda$ (which means x is in $C_{\lambda,\mathbb{P}}(S)$). But $L_n(x, S) \sim \mathrm{Binom}(n, \pi)$ while $1 - \pi < \lambda$ implies $\pi > \frac{1}{1+\lambda}$. By Lemma 3, therefore, we have

$$\mathrm{Prob}(L_n(x, S) \leq w_{\lambda,n}) < \mathrm{Prob}(u \leq w_{\lambda,n}) \leq \alpha$$

where $u \sim \mathrm{Binom}(n, \frac{1}{1+\lambda})$ and $1 - \pi < \lambda$, which proves (8). □

3.2.2 Discussion

Some other features of the test statistic $w_{\lambda,n}$ are worth emphasizing. It is easy to compute that, when $n = 4$, the p-value in (4) exceeds $\alpha := 0.05$ for all λs considered, which means that each member of the doubleton S is declared as "chosen." Put another way, if we have 4 repetitions or less, and the level of significance is set at .05, the choice of the level of selectivity λ is irrelevant and we include each item, even if one of these items is never chosen. Similarly, when $n \leq 7$, then $\text{Prob}(u \leq 1) > 0.05$ for all λs considered. Thus, any time we have at most 7 repetitions, and the level of significance is set at .05, the choice of λ is irrelevant and we include any item chosen at least once out 7 times. These observations illustrate how our criterion tends to be inclusive in the face of (relatively) small sample. This is to be expected, because we include an item as a choice unless we can reject with sufficient confidence that it should be excluded. If we observe only 7 repetitions, we cannot reject that an option chosen only once out of seven times; this option may in fact have a probability of choice as high as 50 percent (so this item is included in the choice set even for $\lambda = 1$.)

The construction becomes more selective when the sample size is larger. To illustrate, the following table calculates $w_{\lambda,20}$ and $w_{\lambda,50}$ for various choices for λ when $\alpha := 0.05$:

λ	0.1	0.3	0.5	0.7	0.9	1
$w_{\lambda,20}$	0	2	3	5	6	6
$w_{\lambda,50}$	1	7	11	15	18	19

In particular, an item chosen 6 times out of 20 (so that its relative choice frequency is 30%) is always included in the choice set for any λ. But an item chosen 3 times out of 20 (so that its relative choice frequency is 15%) is included for $\lambda \leq 0.5$, but excluded for $\lambda \geq 0.7$. On the other hand, an item chosen 19 out of 50 times is included in the choice set (from a doubleton menu), even for $\lambda = 1$. But if an item is chosen 18 times, then it is excluded from the choice set when $\lambda = 1$. When $n = 50$ and $\lambda \leq 0.1$, every item chosen at least once is included in the choice set.

These examples suggest a few properties of the threshold $w_{\lambda,n}$ as a function of λ and n. First, and this is a consequence of the fact that the the cdf of the binomial distribution with parameters n and p is decreasing in n, $w_{\lambda,n}$ is increasing in n. By contrast, $\frac{w_{\lambda,n}}{n}$ is decreasing in n only on average. Moreover, $\frac{w_{\lambda,n}}{n}$ converges (from below) to $\frac{\lambda}{1+\lambda}$ as n tends to infinity, which is the target threshold. When n is small, it is harder to confidently reject the hypothesis that the true underlying frequency is above a threshold, making our criterion more inclusive. As n grows, however, the small-sample nature of the data becomes less problematic, and the sample threshold converges to the one we would have with ideal data.

On the other hand, it is plain from Lemma 3 and the definition of $w_{\lambda,n}$ that this function is increasing in the level of selectivity λ, that is, we have $w_{\lambda,n} \leq w_{\lambda',n}$ whenever $0 \leq \lambda \leq \lambda' \leq 1$ (for any n). In particular, $w_{\lambda,n} \leq w_{1,n}$ for all $\lambda \in [0,1]$ and $n \in \mathbb{N}$. This is, of course, in the nature of things. As λ increases, it becomes harder to be admitted as a choice, and the estimated choice correspondence gets "thinner." The hardest test obtains at $\lambda = 1$, which yields the most exclusive (empirical) choice correspondence.

Remark 4. Describing the exact nature of the function $w_{\lambda,n}$ is not a trivial task, but it can be shown that this function is increasing in n, and it satisfies $\frac{w_{\lambda,n}}{n} \leq \frac{\lambda}{1+\lambda}$ for all $\lambda \in [0,1]$ and $\alpha \in [0, \frac{1}{4}]$ for sufficiently large n. (This is not an asymptotic result; it holds as soon as $n > \frac{1+\lambda}{\lambda}$.) Finally, and perhaps most important, we have $\lim \frac{w_{\lambda,n}}{n} = \frac{\lambda}{1+\lambda}$, that is, our test statistic $\frac{w_{\lambda,n}}{n}$ is a consistent estimator of the true threshold $\frac{\lambda}{1+\lambda}$ regardless of the choice of λ and α. (Proofs are given in the Appendix.) □

Remark 5. Experimental datasets are often quite small, so one cannot do better than applying the above small-sample procedure as stated. Yet, in other cases, researchers may have access to larger datasets—for example, marketing firms in the digital age may gain access to detailed observations of purchasing behaviors in frequent repetitions. If n is large and λ is not small, one can approximately determine the threshold $w_{\lambda,n}$ using the normal distribution. A long-standing convention in statistics is that one can safely approximate a binomial distribution with parameters n and q with the normal distribution provided that $nq > 5$. Thus, in the present setup, so long as $\frac{n\lambda}{1+\lambda} > 5$ holds, we have

$$\alpha = \text{Prob}(L_n(x,S) \leq w_{\lambda,n}) \approx \text{Prob}\left(\ell_n(x,S) \leq \frac{w_{\lambda,n}+0.5}{n}\right) \approx \Phi\left(\frac{\frac{w_{\lambda,n}+0.5}{n} - \frac{\lambda}{1+\lambda}}{\sqrt{\frac{\lambda}{n(1+\lambda)^2}}}\right)$$

by continuity correction. (Here Φ stands for the standard normal (cumulative) distribution function.) From this approximation, we readily get a normal-based approximation for the critical value we are after as

$$w_{\lambda,n} \approx n\frac{\lambda}{1+\lambda}\left(1 + \frac{\Phi^{-1}(\alpha)}{\sqrt{n\lambda}}\right) - 0.5.$$

□

3.2.3 Inferring a Choice Correspondence over Doubleton Menus

We now turn to the problem of inferring one's choice correspondence on any given collection of doubleton menus. Let us then fix any $m \in \mathbb{N}$ and any $S_1, ..., S_m \in \mathfrak{X}_2$, and suppose λ and α are chosen. Let us enumerate S_i as $\{x_i, y_i\}$ for each $i = 1, ..., m$, and denote the number of times we have observed the subject choose from the menu S_i by n_i. In what follows, we set $\pi_i := \mathbb{P}(x_i, S_i)$, and relabelling if necessary, assume $L_{n_i}(x_i, S_i) \leq L_{n_i}(y_i, S_i)$ for each $i = 1, ..., m$.

In this case, our $2m$-hypotheses testing problem (3) is:

$$H_{x_i} : \pi_i \geq \lambda M_{\mathbb{P}}(S_i),\ i = 1, ..., m \quad \text{and} \quad H_{y_i} : 1 - \pi \geq \lambda M_{\mathbb{P}}(S_i),\ i = 1, ..., m.$$

By Assumption 2, however, this reduces to the simpler m-hypotheses problem

$$H_1 : \pi_1 \geq \lambda M_{\mathbb{P}}(S_1) \quad \cdots \quad H_m : \pi_m \geq \lambda M_{\mathbb{P}}(S_m). \tag{9}$$

In addition, as in Section 3.2.1, we may invoke Lemma 3 to further simplify our problem as:

$$H_1 : \pi_1 = \tfrac{\lambda}{1+\lambda} \quad \cdots \quad H_m : \pi_m = \tfrac{\lambda}{1+\lambda}.$$

We have already computed the p-values for each of these hypotheses in Section 3.2.1: For each $i = 1, ..., m$, given the realization of $L_{n_i}(x_i, S_i)$, the p-value of the hypothesis H_i is given as

$$p_i := \text{Prob}(u \leq L_{n_i}(x_i, S_i)) \quad \text{where} \quad u \sim \text{Binom}(n_i, \tfrac{\lambda}{1+\lambda}).$$

Let us now order these p-values from smallest to the largest as

$$p_{(1)} \leq \cdots \leq p_{(m)}.$$

Next, we define

$$i^* := \max\left\{i \in \{1, ..., m\} : p_{(i)} \leq \tfrac{i}{m}\alpha\right\},$$

provided that there is at least one $i \in \{1, ..., m\}$ with $p_{(i)} \leq \frac{i}{m}\alpha$, and set $i^* := 0$ otherwise. The Benjamini-Hochberg procedure (with the FDR control level α) maintains that we accept all hyoptheses in (9) if $i^* = 0$, and

$$\text{reject } H_{(1)}, ..., H_{(i^*)} \quad \text{and} \quad \text{accept } H_{(i^*+1)}, ..., H_{(m)}$$

if $1 \leq i^* < m$, and reject all hypotheses in (9) if $i^* = m$. As a result, we obtain

$$c_{\lambda,\alpha}(S_{(i)}) = \begin{cases} \{x_i, y_i\}, & \text{if } i > i^* \\ \{y_i\}, & \text{otherwise,} \end{cases}$$

which completely describes the inferred choice correspondence on the domain $\{S_1, ..., S_m\}$. The results reported in the application of Section 4 are obtained by means of this method.

3.3 The Case of Arbitrary Choice Situations

3.3.1 Testing if a Particular Alternative is a Choice

Let us now take an arbitrary menu $S \in \mathfrak{X}_k$ where k is any integer with $k \geq 3$, and suppose again that λ and α are chosen. Let us enumerate S as $\{x_1, ..., x_k\}$, and simplify the notation by setting

$$\pi_i := \mathbb{P}(x_i, S) \quad \text{and} \quad \mathcal{L}_i := L_n(x_i, S), \quad i = 1, ..., k,$$

where n is the number of times we have observed this person make a choice from S. Thus, π_i is the (unknown) probability of the agent choosing x_i from S, while $\mathcal{L}_i$ is the number of times the individual has been observed to choose the item x_i from S in a choice experiment that is repeated n times. Obviously, $(\pi_1, ..., \pi_k)$ belongs to the $(k-1)$-dimensional unit simplex, while $\mathcal{L}_1 + \cdots + \mathcal{L}_k = n$. Moreover, by Assumption 1, we have

$$\text{Prob}(\mathcal{L}_1 = a_1, ..., \mathcal{L}_k = a_k) = \frac{n!}{a_1! \cdots a_k!} \pi_1^{a_1} \cdots \pi_k^{a_k}$$

where $(a_1, ..., a_k) \in \{0, ..., n\}^k$ and $a_1 + \cdots + a_k = n$. Thus, the distribution of $(\mathcal{L}_1, ..., \mathcal{L}_k)$ is multinomial with parameters n and $\pi_1, ..., \pi_k$ (but note that this distribution is singular in the sense that the covariance matrix of $(\mathcal{L}_1, ..., \mathcal{L}_k)$ is of rank $k-1$, due to the restriction $\mathcal{L}_1 + \cdots + \mathcal{L}_k = n$). Consequently, as proved in Chapter 35 of Johnson, Kotz and Balakrishnan (1997), we have:

> ***Lemma 4.*** (***a***) $\mathcal{L}_1$ is binomially distributed with parameters n and π_1;
>
> (***b***) If $\pi_1 = 1$, then $\mathcal{L}_1 = n$ and $\mathcal{L}_2 = \cdots = \mathcal{L}_k = 0$ almost surely. If $\pi_1 < 1$, then for any $(a_1, ..., a_k) \in \{0, ..., n\}^k$ with $a_1 + \cdots + a_k = n$,
>
> $$\text{Prob}(\mathcal{L}_2 = a_2, ..., \mathcal{L}_k = a_k \mid \mathcal{L}_1 = a_1) = \frac{(n-a_1)!}{a_2! \cdots a_k!} \prod_{i=2}^{k} \left(\frac{\pi_i}{1-\pi_1} \right)^{a_i};$$

that is, conditional on $\mathcal{L}_1 = a_1$, $(\mathcal{L}_2, ..., \mathcal{L}_k)$ is multinomially distributed with parameters $n - a_1$ and $\frac{\pi_2}{1-\pi_1}, ..., \frac{\pi_k}{1-\pi_1}$.

Our goal now is to find a p-value for the hypothesis that a particular member of S is a choice for the subject individual from that menu. For concreteness, but without loss of generality, we study this hypothesis for the alternative x_1. Formally, then, our null hypothesis is

$$H : \pi_1 \geq \lambda \max\{\pi_1, ..., \pi_k\}.$$

If $\lambda = 0$, then the problem is a simple hypothesis problem, so we do not have to elaborate on it here. In what follows, then, we assume $\lambda > 0$. As in the case of pairwise choice situations, it is enough to test for equality here, so we write this hypothesis as

$$H : n\pi_1 = \lambda \max\{n\pi_2, ..., n\pi_k\}. \tag{10}$$

To obtain the p-value we are after, and the associated threshold test, we consider the random variable

$$V_{a_1} := \mathcal{L}_2 \vee \cdots \vee \mathcal{L}_k$$

conditional on $\mathcal{L}_1 = a_1$, where $a_1 \in \{0, ..., n\}$. In words, V_{a_1} is the maximum value of the random variables $\mathcal{L}_2, ..., \mathcal{L}_k$ given that x_1 is chosen from S exactly a_1 many times. By Lemma 4,

$$\begin{aligned} \text{Prob}(V_{a_1} < v) &= \text{Prob}(\mathcal{L}_2 < v, ..., \mathcal{L}_k < v \mid \mathcal{L}_1 = a_1) \\ &= \sum_{(a_2,...,a_k) \in S(v;a_1)} \frac{(n-a_1)!}{a_2! \cdots a_k!} \prod_{i=2}^{k} \left(\frac{\pi_i}{1-\pi_1} \right)^{a_i} \end{aligned}$$

where $S(v; a_1)$ is the set of all $(k-1)$-vectors $(a_2, ..., a_k)$ of nonnegative integers such that $a_i < v$ for each $i = 2, ..., k$, and $a_2 + \cdots + a_k = n - a_1$.

We now define our test statistic (that depends on the chosen λ) as

$$W_\lambda := \mathcal{L}_1 - \lambda \left(\mathcal{L}_2 \vee \cdots \vee \mathcal{L}_k \right), \tag{11}$$

and note that this is a simple random variable whose (finite) range we denote by $\text{rng}(W_\lambda)$.[19]

[19]To be precise, let $J_0 := \{0\}$, and for any positive integer N, set $J_N := \{N, N-1, ..., \lfloor \frac{N}{k} \rfloor\} \backslash \{0\}$. Then, $\text{rng}(W_\lambda)$ is contained within the finite set $\bigcup_{i=0}^{n} (i - \lambda J_{n-i})$.

The p-value that we are after is thus given as

$$p = \text{ Prob}\left(W_\lambda \leq L_n(x_1, S) - \lambda \max_{i=2,\ldots,k} L_n(x_i, S)\right). \tag{12}$$

Of course, to compute this value, we need to find the distribution of W_λ. To this end, note that

$$\begin{aligned}
\text{Prob}(W_\lambda \leq w) &= 1 - \text{Prob}(W_\lambda > w) \\
&= 1 - \text{Prob}(\mathcal{L}_2 \vee \cdots \vee \mathcal{L}_k < \tfrac{1}{\lambda}(\mathcal{L}_1 - w)) \\
&= 1 - \sum_{a_1=0}^{n} \text{Prob}(V_{a_1} < \tfrac{1}{\lambda}(\mathcal{L}_1 - w) \mid \mathcal{L}_1 = a_1)\text{Prob}(\mathcal{L}_1 = a_1) \\
&= 1 - \sum_{a_1=0}^{n} \left(\sum_{(a_2,\ldots,a_k) \in S\left(\frac{a_1 - w}{\lambda}; a_1\right)} \frac{(n-a_1)!}{a_2! \cdots a_k!} \prod_{i=2}^{k} \left(\frac{\pi_i}{1-\pi_1}\right)^{a_i} \right) \frac{n! \pi_1^{a_1}(1-\pi_1)^{n-a_1}}{a_1!(n-a_1)!} \\
&= 1 - \sum_{a_1=0}^{n} \sum_{(a_2,\ldots,a_k) \in S\left(\frac{a_1 - w}{\lambda}; a_1\right)} \frac{n!}{a_1! a_2! \cdots a_k!} \prod_{i=1}^{k} \pi_i^{a_i}
\end{aligned}$$

for any real number w. This completely characterizes the distribution of W_λ for any probability vector $(\pi_1, ..., \pi_k)$. However, given that $k \geq 3$, the equations $\pi_1 = \lambda \max\{\pi_2, ..., \pi_k\}$ and $\pi_1 + \cdots + \pi_k = 1$ does not determine the vector $(\pi_1, ..., \pi_k)$ uniquely. Thus, unlike the case of pairwise choice situations, assuming the validity of the hypothesis (10) does not identify the distribution of W_λ. We thus have to use some suitable proxies for π_is, $i = 2, ..., k$, to determine a specific distribution for W_λ to evaluate the p-value in (12) exactly.

The procedure we propose here uses the sample relative frequencies as proxies for $\pi_2, ..., \pi_k$, and to account for the null hypothesis being true, sets π_1 as λ times the largest of these frequencies. That is, we define

$$\pi_2' := \ell_n(x_2, S), \ \ldots, \ \pi_k' := \ell_n(x_k, S) \quad \text{ and } \quad \pi_1' := \lambda \max\{\pi_2', ..., \pi_k'\}.$$

If $\pi_i' = 0$ for each $i = 2, ..., k$ here, then we are in the exceptional case in which the agent is observed to choose x_1 from S in every repetition of the experiment. In that case, of course, we do not reject our null hypothesis, and thus include x_1 in $c_{\lambda,\alpha}(S)$ as stipulated by Assumption 2. If, on the other hand, $\pi_i' > 0$ for at least one $i = 2, ..., k$, we normalize these numbers as

$$\hat{\pi}_1 := \frac{\pi_1'}{\pi_1' + \cdots + \pi_k'}, \ \ldots, \ \hat{\pi}_k := \frac{\pi_k'}{\pi_1' + \cdots + \pi_k'}.$$

These probabilities are, in turn, used in lieu of the actual probabilities $\pi_1, \ldots, \pi_k$ to determine the distribution of W_λ exactly. In other words, provided that the agent has chosen an alternative other than x_1 at least once from S, the p-value for testing "$x_1 \in c_{\lambda,\alpha}(S)$" is obtained as

$$p = \text{ Prob}\left(W \leq L_n(x_1, S) - \lambda \max_{i=2,\ldots,k} L_n(x_i, S)\right)$$

with W being a simple (rng(W_λ)-valued) random variable whose cdf is given by

$$\text{Prob}(W \leq w) = 1 - \sum_{a_1=0}^{n} \sum_{(a_2,\ldots,a_k)\in S\left(\frac{a_1-w}{\lambda};a_1\right)} \frac{n!}{a_1!a_2!\cdots a_k!} \prod_{i=1}^{k} \hat{\pi}_i^{a_i}. \tag{13}$$

We thus accept x_1 as a choice from S (given λ, and at the significance level α) iff either x_1 is chosen in S in every repetition or $p \geq \alpha$.

Remark 6. We can also reformulate this test as a threshold test. Define

$$w_{\lambda,n} := \max\left\{w \in \text{rng}(W_\lambda) : \text{Prob}(W \leq w) \leq \alpha\right\}$$

where W is the random variable whose cumulative distribution is given by (13). Then, our test maintains that we reject the hypothesis H if $\ell_n(x_1, S) - \lambda \max\{\hat{p}_2, \ldots, \hat{p}_k\} \leq \frac{1}{n} w_{\lambda,n}$, or equivalently, do not include x_1 in $c_{\lambda,\alpha}(S)$ if the realization of $\mathcal{L}_n(x_1, S) - \lambda\mathcal{L}_n(x_i, S)$ is less than or equal to $w_{\lambda,n}$ for each $i = 2, \ldots, k$. □

3.3.2 The General Testing Problem, Revisited

Let us now return to the general hypothesis testing problem at hand, namely, (3). Again, suppose λ and α are chosen. If all the menus $S_1, \ldots, S_m$ are doubletons, we have described how to test this problem at the control level α (by means of the Benjamini-Hochberg procedure) in Section 3.2.3. The general case can now be handled by the analogous approach. First, we delete from the system all hypotheses that correspond to alternatives that are most frequently chosen in the menus that belong, and include these as "choices" from those menus. For each of the remaining hypotheses, we determine the p-values as described in the previous subsection. Finally, we order these p-values, and apply the Benjamini-Hochberg procedure at the control level α to determine exactly which of the hypotheses of (3) are rejected. This furnishes the inferred choice correspondence of the individual subject to the selected values of λ and α.

4 Application: On the Transitivity of Preferences

To show how our method can be applied easily to choice data, we use it in the case of an eminent experiment: Experiment 1 of Tversky (1969). This is of the best known experiments in this area, with more than 3300 citations in Google Scholar at the time of this writing. The primary goal of Tversky's study was to show that preferences may violate stochastic transitivity in multi-attribute choice problems in which differences in one attribute are less prominent. In the modern literature on boundedly rational choice, Tversky (1969) is routinely cited as evidence of cyclical individual choice behavior, and hence the violation of WARP (cf. Manzini and Mariotti, 2007, Masatlioglu, Nakajima, and Ozbay, 2012, and Tserenjigmid, 2015). But Tversky's experiment uses stochastic choice data, and to see if it indeed provides evidence of widespread violation of WARP, one needs to infer subjects' deterministic choice correspondences. We apply our method to do so, and come to more nuanced conclusions.

Before proceeding, we note that the original experiment of Tversky had only 8 subjects, but was later replicated by Regenwetter et al. (2011) with 18 participants.[20] We pool the data to obtain the behavior of 26 individuals.

The Original Experiment. Tversky's experiment studied transitivity from doubleton choice problems and recorded the choice from all pairs of 5 gambles, named $a, b, ..., e$, each repeated 20 times (with additional 'decoy' choices in between). Gambles were such that those with names adjacent in the alphabet had similar probabilities of winning and more pronounced differences in payoffs; probabilities of winning differed substantially only across gambles with names further apart in the alphabet.[21] Tversky conjectured that this would lead to violations of transitivity of choices. To investigate this, he tested whether the empirical stochastic choice function satisfied Weak Stochastic Transitivity (WST) and found that this property failed significantly for 5 out of the 8 subjects (62%).[22]

[20]The only differences are the use of computers, updated payoffs (since decades have passed), and the implementation of the experiment in only one session (as opposed to Tversky's five sessions).

[21]If a gamble that pays x with probability p is identified by (p, x), the gambles were: $a = (\frac{7}{24}, 5)$, $b = (\frac{8}{24}, 4.75)$, $c = (\frac{9}{24}, 4.50)$, $d = (\frac{10}{24}, 4.25)$, $e = (\frac{11}{24}, 4)$. When comparing, for example, a and b, the difference between payoffs (5 vs. 4.75) seems more relevant than the difference between probabilities ($\frac{7}{24}$ vs. $\frac{8}{24}$).

[22]Recall that a stochastic choice function satisfies *Weak Stochastic Transitivity* if $P(x, \{x, y\}) \geq .5$ and $P(y, \{y, z\}) \geq .5$ imply $P(x, \{x, z\}) \geq .5$. Subsequent literature questioned the statistical tests used in Tver-

Applying our method. We use our procedure to infer a choice correspondence for each subject and test if this correspondence satisfies WARP. Maintaining $\alpha = .05$, we consider values of $\lambda \in \{0, .3, .5, .7, 1\}$. The first column of Table 1 displays the fraction of subjects whose computed choice correspondence abides by WARP for each λ.

As clear from the table, the majority of subjects are found rational unless λ is chosen very high: the fraction of WARP-abiding subjects ranges from 100% when $\lambda = 0$, to 62% when $\lambda = 0.5$, to 42% when $\lambda = 1$.[23] Even with $\lambda = .7$, the choices of about half of the subjects are consistent with rationality. One needs rather extreme values of the parameters, like λ close to one, to obtain a relevant majority violating rationality. Overall, contrary to what is often argued, this data does *not* appear to constitute evidence of widespread non-transitive or non-rational behavior in deterministic choice.

These results also allow us to ease another possible concern with our method. As discussed at length in Section 3, our inference problem takes the form of multiple-hypotheses testing, which one may worry may result in accepting all null hypotheses unless we have very large samples. In our context, this would translate into accepting any element as a choice (i.e., declaring $c_{\lambda,\alpha}(S) = S$ for each S), giving WARP by default.

The present data shows this is not the case. While the inferred choice correspondence is bound to be very inclusive with $\lambda = 0$, our procedure becomes discerning even with low values of λ, such as .3. Indeed, 19% of the subjects are found to violate WARP with $\lambda = .3$, while 48% does so with $\lambda = .5$.

WST vs. WARP. To further investigate the issue of rationality in this data, it may be useful to look at how testing WARP using our approach relate to testing violations of WST. There is a formal, albeit superficial, relation between the two tests: if $\lambda = 1$ and if we treat the observed data as if it were $\mathbb{P}$—that is, we disregard sampling errors—then an alternative belongs to the computed choice correspondence if, and only if, it is chosen at least half of the times. Thus, in such a special treatment, testing transitivity of the choice correspondence is identical to testing WST. However, the two tests differ once we depart from making either of these extreme assumptions. (The choice of $\lambda = 1$ is so demanding that it rules out options

sky (1969) and argued that more apt tests fail to find significant violations of Weak Stochastic Transitivity; cf. Iverson and Falmagne (1985) and Regenwetter et al. (2011). As mentioned explicitly in the latter, these concerns are rarely discussed, which is particularly striking given the prominence of Tversky's original paper.

[23]The fact that the fraction of rational subjects decreases with λ is a feature of this data but not a prediction of the model. As λ increases, fewer items are included in the choice correspondences, which may eliminate previous violations of WARP but may well add new ones. For example, in the data, subject #17 satisfies WARP with $\lambda = .3$ but violates it with $\lambda = .7$, while subject #5 violates WARP with $\lambda = .3$ but satisfies it with $\lambda = .7$.

λ	Satisfy WARP	Satisfy WST	Satisfy WARP and WST	Sat. WARP Violates WST	Sat. WST Violates WARP	Violates WARP and WST
0	100	46	46	54	0	0
.3	81	46	31	50	15	4
.5	62	46	31	31	15	23
.7	50	46	35	15	12	38
1	42	46	31	12	15	42

Table 1: Percentage of subject that satisfy WARP and WST in the data of Tversky (1969) and replication

chosen, say, 8 or 9 times out of 20 repetitions.)

This is easy to illustrate using Tversky's data. WST is satisfied by only 46% of the subjects, which implies that testing WARP using our procedure returns a higher fraction of rational subjects unless λ is very high. Importantly, violations of WARP and of WST are neither nested nor perfectly correlated. This is evident from Table 1, which reports the fraction of subjects that satisfy each property and their combination. For example, when $\lambda = .5$, an apparently reasonable choice for the level of selectivity, 62% of subjects satisfy WARP and 46% satisfy WST, but these are not nested groups: 31% satisfy WARP but not WST, another 15% satisfy WST and not WARP. Only 23% violates both, and only 31% satisfies both. Table 1 plainly shows how, across different values of λ, sizable fractions of subjects satisfy one property but not the other. Independently of the merit of each procedure, it seems to us that analyzing choice data by employing both approaches—that is, inferring a deterministic choice correspondence as well as deducing a stochastic choice function—provides a better insight into understanding one's choice behavior.

5 Conclusion

Despite being perhaps the most fundamental primitive of microeconomics, choice correspondences are not observable. Barring some specially designed experiments, all we can observe in general is a single choice made by an individual at a given time, and not the *set* of all her potential choices. In this paper we propose a method to "compute" a choice correspondence using data that come in the form of repeated observations of choices made by a decision maker.

Our approach constructs one's choice correspondence in two stages. First, the analyst needs to decide how to impute the choice correspondence if she had access to an 'ideal' dataset that provides the true choice probability of each option. There is no unexceptionable method of doing this, but we have underscored here a one-parameter family of choice imputations. These have the advantage of being mathematically simple, and as we have shown in Section 2, are erected on an axiomatic foothold (especially in the context of pairwise choice problems). Any one member of this family either includes everything that is chosen with positive probability in the choice set from a menu S, or keeps in that set only the options whose probability of choice is higher than λ times the choice probability of any other alternative in S. The parameter λ determines how inclusive the criterion is, and would be chosen by the analyst according to the problem at hand.

The second stage pertains to applying such a rule in the real world, where the analyst does not have access to an ideal data set, but is instead confronted with a finite number of observations. This brings a set of issues concerning sampling errors to the fore. To address these, we develop statistical methods to estimate a choice correspondence by means of hypothesis testing.

When combined, these two stages require the analyst to select two parameters—λ, indicating the inclusivity with ideal data, and α, indicating the level of significance for hypothesis testing—and provides practical formulae to infer choice correspondences. To illustrate the use of our overall method of elicitation, we considered here the repeated (within-subject) choice experiment of Tversky (1969) and estimated the deterministic choice correspondences of the subjects of this experiment. We show how our approach may yield novel insights into nature of the violations of rationality in this data.

APPENDIX: PROOFS

We begin with two lemmata.

> ***Lemma A.1.***[24] Let $\Psi : \mathsf{scf}(X) \to \mathsf{cc}(X)$ be a choice imputation that satisfies the properties A and B. Then, for every $\mathbb{P}, \mathbb{Q} \in \mathsf{scf}(X)$ and $S \in \mathfrak{X}_2$,
>
> $$x \in \Psi(\mathbb{P})(S) \text{ and } \mathbb{Q}(z,S) \geq \mathbb{P}(x,S) \quad \text{imply} \quad z \in \Psi(\mathbb{Q})(S).$$

Proof. As usual, we put $C_{\mathbb{P}} := \Psi(\mathbb{P})$ for any $\mathbb{P} \in \mathsf{scf}(X)$. Now take any $\mathbb{P}, \mathbb{Q} \in \mathsf{scf}(X)$ and $S \in \mathfrak{X}_2$ and pick any $x, z \in S$ such that $x \in C_{\mathbb{P}}(S)$ and $\mathbb{Q}(z,S) \geq \mathbb{P}(x,S)$. We wish to show that $z \in C_{\mathbb{Q}}(S)$. To this end, we take any $y \in X \backslash \{x,z\}$, put $A := \{x,y\}$, and pick a $\mathbb{P}_1 \in \mathsf{scf}(X)$ such that

$$\mathbb{P}_1(x,S) = \mathbb{P}(x,S) \quad \text{and} \quad \mathbb{P}_1(y,A) = \mathbb{P}(x,S).$$

As $\mathbb{P}_1(\cdot,S) = \mathbb{P}(\cdot,S)$, we have $C_{\mathbb{P}_1}(S) = C_{\mathbb{P}}(S)$ by property A, so $x \in C_{\mathbb{P}_1}(S)$. But then, since $\mathbb{P}_1(y,A) = \mathbb{P}(x,S) = \mathbb{P}_1(x,S)$, property B entails that $y \in C_{\mathbb{P}_1}(A)$.

Let us now pick any $\mathbb{P}_2 \in \mathsf{scf}(X)$ such that

$$\mathbb{P}_2(y,A) = \mathbb{P}_1(y,A) \quad \text{and} \quad \mathbb{P}_2(z,S) = \mathbb{Q}(z,S).$$

As $\mathbb{P}_2(\cdot,A) = \mathbb{P}_1(\cdot,A)$, we have $C_{\mathbb{P}_2}(A) = C_{\mathbb{P}_1}(A)$ by property A, so by what we have just found, $y \in C_{\mathbb{P}_2}(A)$. Moreover,

$$\mathbb{P}_2(z,S) = \mathbb{Q}(z,S) \geq \mathbb{P}(x,S) = \mathbb{P}_1(y,A) = \mathbb{P}_2(y,A),$$

so by property B, $z \in C_{\mathbb{P}_2}(S)$. But as $\mathbb{P}_2(\cdot,S) = \mathbb{Q}(\cdot,S)$, property A says that $C_{\mathbb{P}_2}(S) = C_{\mathbb{Q}}(S)$, so we conclude that $z \in C_{\mathbb{Q}}(S)$. ■

> ***Lemma A.2.*** Let $\Psi : \mathsf{scf}(X) \to \mathsf{cc}(X)$ be a choice imputation that satisfies the properties A and B. Then, for every $\mathbb{P}, \mathbb{Q} \in \mathsf{scf}(X)$ and $S, T \in \mathfrak{X}_2$,
>
> $$x \in \Psi(\mathbb{P})(S) \text{ and } \mathbb{Q}(z,T) \geq \mathbb{P}(x,S) \quad \text{imply} \quad z \in \Psi(\mathbb{Q})(T).$$

Proof. As usual, we put $C_{\mathbb{P}} := \Psi(\mathbb{P})$ for any $\mathbb{P} \in \mathsf{scf}(X)$. Now take any $\mathbb{P}, \mathbb{Q} \in \mathsf{scf}(X)$ and $S, T \in \mathfrak{X}_2$ and pick any $(x,z) \in T \times S$ such that $x \in C_{\mathbb{P}}(S)$ and $\mathbb{Q}(z,T) \geq \mathbb{P}(x,S)$. If $S = T$, we are done by Lemma A.1. Suppose, then, S and T are distinct. Let $\mathbb{P}_0$ be any element of $\mathsf{scf}(X)$ with $\mathbb{P}_0(\cdot,S) = \mathbb{P}(\cdot,S)$ and $\mathbb{P}_0(\cdot,T) = \mathbb{Q}(\cdot,T)$. By the property A, we then have $C_{\mathbb{P}_0}(S) = C_{\mathbb{P}}(S)$ and $C_{\mathbb{P}_0}(T) = C_{\mathbb{Q}}(T)$.

[24] Gerelt Tserenjigmid has suggested this lemma to us, which simplifies the subsequent argument.

It follows that $x \in C_{\mathbb{P}_0}(S)$ while $\mathbb{P}_0(z,T) \geq \mathbb{P}_0(x,S)$, whence property B entails $z \in C_{\mathbb{P}_0}(T)$, so we again find $z \in C_{\mathbb{Q}}(T)$, as desired. ■

Proof of Theorem 1

For any $\mathbb{P} \in \mathsf{scf}(X)$, we put $C_{\mathbb{P}} := \Psi(\mathbb{P})$, and define

$$\lambda_{\mathbb{P}} := \min\left\{\frac{m_{\mathbb{P}}(S)}{M_{\mathbb{P}}(S)} : C_{\mathbb{P}}(S) = S \in \mathfrak{X}_2\right\}.$$

Clearly, $0 < \lambda_{\mathbb{P}} \leq 1$. (If $\lambda_{\mathbb{P}} = 0$ were the case, then there would be an $S \in \mathfrak{X}_2$ with $S = C_{\mathbb{P}}(S)$ and $m_{\mathbb{P}}(S) = 0$, but this would contradict (1).) Now take any $\mathbb{P} \in \mathsf{scf}(X)$ and $S \in \mathfrak{X}_2$. Denote the elements of S by x and y so that $\mathbb{P}(x,S) \leq \mathbb{P}(y,S)$. If $x \in C_{\mathbb{P}}(S)$, then $\frac{\mathbb{P}(x,S)}{\mathbb{P}(y,S)} \geq \lambda_{\mathbb{P}}$ by definition of $\lambda_{\mathbb{P}}$. Conversely, suppose $\mathbb{P}(x,S) \geq \lambda_{\mathbb{P}}\mathbb{P}(y,S)$. Since $\mathfrak{X}_2$ is finite, there is an $T \in \mathfrak{X}_2$ with $C_{\mathbb{P}}(T) = T$ and $\frac{m_{\mathbb{P}}(T)}{M_{\mathbb{P}}(T)} = \lambda_{\mathbb{P}}$. Then, $\frac{\mathbb{P}(x,S)}{1-\mathbb{P}(x,S)} \geq \frac{m_{\mathbb{P}}(T)}{1-m_{\mathbb{P}}(T)}$, and it follows that $\mathbb{P}(x,S) \geq m_{\mathbb{P}}(T)$. But then, by property B, we obtain $x \in C_{\mathbb{P}}(S)$. Thus, $x \in C_{\mathbb{P}}(S)$ iff $\mathbb{P}(x,S) \geq \lambda_{\mathbb{P}}\mathbb{P}(y,S)$. As this property (as well as Lemma A.1) implies that $y \in C_{\mathbb{P}}(S)$, and $\mathbb{P}(y,S) \geq \lambda_{\mathbb{P}}M_{\mathbb{P}}(S)$ holds trivially, and because S was arbitrarily chosen above, we conclude that

$$C_{\mathbb{P}}(S) = \{x \in S : \mathbb{P}(x,S) \geq \lambda_{\mathbb{P}}M_{\mathbb{P}}(S)\}$$

for every $\mathbb{P} \in \mathsf{scf}(X)$ and $S \in \mathfrak{X}_2$.

To complete our proof, we define

$$\lambda := \inf_{\mathbb{P}\in\mathsf{scf}(X)} \lambda_{\mathbb{P}}.$$

Let us again fix arbitrary $\mathbb{P} \in \mathsf{scf}(X)$ and $S \in \mathfrak{X}_2$, and denote the elements of S by x and y so that $\mathbb{P}(x,S) \leq \mathbb{P}(y,S)$. Besides, let us pick a sequence $(\mathbb{Q}_k)$ in $\mathsf{scf}(X)$ such that $\lambda_{\mathbb{Q}_k} \downarrow \lambda$. As $\mathfrak{X}_2$ is finite, for every positive integer k, there is a $T_k \in \mathfrak{X}_2$ with $C_{\mathbb{Q}_k}(T_k) = T_k$ and $\lambda_{\mathbb{Q}_k} = \frac{m_{\mathbb{Q}_k}(T_k)}{M_{\mathbb{Q}_k}(T_k)}$. Since $\mathfrak{X}_2$ is finite, there must exist a constant subsequence of (T_k), so it is without loss of generality to assume that $T_1 = T_2 = \cdots = T$ for some $T \in \mathfrak{X}_2$. Then, $\lambda_{\mathbb{Q}_k} = \frac{m_{\mathbb{Q}_k}(T)}{M_{\mathbb{Q}_k}(T)}$ and $C_{\mathbb{Q}_k}(T) = T$ for each k. Next, for each $k \in \mathbb{N}$ we take any $\mathbb{P}_k \in \mathsf{scf}(X)$ with $\mathbb{P}_k(x,S) = m_{\mathbb{Q}_k}(T)$. Then, by Lemma A.2, we have $C_{\mathbb{P}_k}(S) = S$ for each k. Thus, $\lambda_{\mathbb{Q}_k} = \frac{m_{\mathbb{P}_k}(S)}{M_{\mathbb{P}_k}(S)}$ and $C_{\mathbb{P}_k}(S) = S$ for each k.

Consider first the case in which $\lambda = 0$. In this case, $\frac{m_{\mathbb{P}_k}(S)}{M_{\mathbb{P}_k}(S)} \downarrow 0$, whence $m_{\mathbb{P}_k}(S) \downarrow 0$. If $\mathbb{P}(x,S) > 0$, therefore, we have $\mathbb{P}(x,S) > m_{\mathbb{P}_k}(S)$ for large enough k. But then, Lemma A.2 entails that $x \in C_{\mathbb{P}}(S)$. Thus, $\mathrm{supp}(\mathbb{P}(\cdot,S)) \subseteq C_{\mathbb{P}}(S)$. As the converse inequality is ensured by (1), we conclude that $C_{\mathbb{P}}(S) = C_{\mathbb{P},0}(S)$, as desired.

We now assume that $\lambda > 0$. If $x \in C_{\mathbb{P}}(S)$, then $\mathbb{P}(x,S) \geq \lambda_{\mathbb{P}}M_{\mathbb{P}}(S) \geq \lambda M_{\mathbb{P}}(S)$, so $C_{\mathbb{P}}(S) \subseteq C_{\mathbb{P},\lambda}(S)$. Conversely, assume $\mathbb{P}(x,S) \geq \lambda M_{\mathbb{P}}(S)$, and note that this implies $\mathbb{P}(x,S) > 0$. If $\mathbb{P}(x,S) > \lambda M_{\mathbb{P}}(S)$, then, by definition of λ, there is a k large enough that $\frac{\mathbb{P}(x,S)}{M_{\mathbb{P}}(S)} > \lambda_{\mathbb{Q}_k} \geq \lambda$, whence $\frac{\mathbb{P}(x,S)}{1-\mathbb{P}(x,S)} > \frac{m_{\mathbb{P}_k}(S)}{1-m_{\mathbb{P}_k}(S)}$. Thus, $\mathbb{P}(x,S) > m_{\mathbb{P}_k}(T)$, whence, by Lemma A.2, we find $x \in C_{\mathbb{P}}(S)$, as desired. Finally, suppose that $\mathbb{P}(x,S) = \lambda M_{\mathbb{P}}(S)$. In this case, we have $\frac{m_{\mathbb{P}_k}(S)}{1-m_{\mathbb{P}_k}(S)} \downarrow \frac{\mathbb{P}(x,S)}{1-\mathbb{P}(x,S)}$, whence $m_{\mathbb{P}_k}(S) \downarrow \mathbb{P}(x,S)$. By property C, we thus again find $x \in C_{\mathbb{P}}(S)$. As Lemma A.1 implies $y \in C_{\mathbb{P}}(S)$, and $\mathbb{P}(y,S) \geq \lambda M_{\mathbb{P}}(S)$ holds trivially, we conclude that $C_{\mathbb{P}}(S) = \{x \in S : \mathbb{P}(x,S) \geq \lambda M_{\mathbb{P}}(S)\} = C_{\mathbb{P},\lambda}(S)$, as we sought.

Proof of Theorem 2

The proof of the "if" part of the claim is straightforward, so we focus on its "only if" part. Assume that $\Phi : \mathsf{scf}(X) \to \mathsf{cc}(X)$ satisfies the properties A, B, C and D. By Theorem 1, there exists a (unique) $\lambda \in [0,1]$ such that $C_{\mathbb{P}}(S) = C_{\mathbb{P},\lambda}(S)$ for every $S \in \mathfrak{X}_2$ and $\mathbb{P} \in \mathsf{scf}(X)$, where, as usual, we write $C_{\mathbb{P}}$ for $\Phi(\mathbb{P})$. Now take any $\mathbb{P} \in \mathsf{scf}(X)$ and $S \in \mathfrak{X}$ with $|S| \geq 3$. Let y be an element of S with $\mathbb{P}(y,S) = M_{\mathbb{P}}(S)$, and note that $y \in C_{\mathbb{P}}(S)$ by the property D (applied for $x = y$). Next, take any x in S, put $A := \{x,y\}$, and define $\mathbb{Q} \in \mathsf{scf}(X)$ as $\mathbb{Q}(\cdot,T) := \mathbb{P}(\cdot,T)$ for every $T \in \mathfrak{X}\backslash\{A\}$, and $\mathbb{Q}(x,A) := \frac{\mathbb{P}(x,S)}{\mathbb{P}(x,S)+\mathbb{P}(y,S)}$ and $\mathbb{Q}(y,A) := 1 - \mathbb{Q}(x,A)$.

Now suppose $x \in C_{\mathbb{P}}(S)$. As $A \neq S$ (because S does not belong to $\mathfrak{X}_2$), we have $\mathbb{P}(\cdot,S) = \mathbb{Q}(\cdot,S)$, and the property A thus implies that $x \in C_{\mathbb{Q}}(S)$. Besides,

$$\frac{\mathbb{Q}(x,A)}{\mathbb{Q}(y,A)} = \frac{\mathbb{P}(x,S)}{\mathbb{P}(y,S)} = \frac{\mathbb{Q}(x,S)}{\mathbb{Q}(y,S)},$$

so property D entails $x \in C_{\mathbb{Q}}(A)$. Since $A \in \mathfrak{X}_2$, this means that $\mathbb{Q}(x,A) \geq \lambda\mathbb{Q}(y,A)$, whence $\mathbb{P}(x,S) \geq \lambda\mathbb{P}(y,S) = \lambda M_{\mathbb{P}}(S)$. Conversely, suppose $\mathbb{P}(x,S) \geq \lambda M_{\mathbb{P}}(S)$ holds. Then,

$$\mathbb{Q}(x,A) = \left(\frac{\mathbb{Q}(x,A)}{\mathbb{Q}(y,A)}\right)\mathbb{Q}(y,A) = \left(\frac{\mathbb{P}(x,S)}{\mathbb{P}(y,S)}\right)M_{\mathbb{Q}}(A) \geq \lambda M_{\mathbb{Q}}(A)$$

so $x \in C_{\mathbb{Q}}(A)$. It then follows from the property D that $x \in C_{\mathbb{Q}}(S)$. Since $\mathbb{P}(\cdot,S) = \mathbb{Q}(\cdot,S)$, the property A thus implies $x \in C_{\mathbb{P}}(S)$. In view of the arbitrary choice of x in S, we thus conclude that

$$C_{\mathbb{P}}(S) = \{x \in S : \mathbb{P}(x,S) \geq \lambda M_{\mathbb{P}}(S)\},$$

that is, $C_{\mathbb{P}}(S) = C_{\mathbb{P},\lambda}(S)$. In view of the arbitrary choice of S and $\mathbb{P}$, we are done.

Proof of Lemma 3

For any $n \in \mathbb{N}$ and $c \in \{0,...,n\}$, define the self-map $\varphi_{n,c}$ on $[0,1]$ by

$$\varphi_{n,c}(\pi) := \sum_{i=0}^{c} \binom{n}{i} \pi^i (1-\pi)^{n-i}$$

and note that

$$\begin{aligned}
\varphi_{n,c}'(\pi) &= \sum_{i=0}^{c}\binom{n}{i}i\pi^{i-1}(1-\pi)^{n-i} - \sum_{i=0}^{c}\binom{n}{i}(n-i)\pi^{i}(1-\pi)^{n-i-1} \\
&= \sum_{i=1}^{c}\frac{n!}{(i-1)!(n-i)!}\pi^{i-1}(1-\pi)^{n-i} - \sum_{i=0}^{c}\frac{n!}{i!(n-i-1)!}\pi^{i}(1-\pi)^{n-i-1} \\
&= n\sum_{i=1}^{c}\binom{n-1}{i-1}\pi^{i-1}(1-\pi)^{n-i} - n\varphi_{n-1,c}(\pi) \\
&= n\left(\sum_{i=0}^{c-1}\binom{n-1}{i}\pi^{i}(1-\pi)^{n-i-1} - \varphi_{n-1,c}(\pi)\right) \\
&= n\left(\varphi_{n-1,c-1}(\pi) - \varphi_{n-1,c}(\pi)\right).
\end{aligned}$$

It follows that $\varphi_{n,c}' < 0$, and the lemma follows.[25]

Proofs of the Claims of Remark 4

For any $\pi \in [0,1]$, let (x_m) be a sequence of Bernoulli random variables on a given probability space that are i.i.d. with parameter π. For any positive integer n, put $u_n := x_1 + \cdots + x_n$; u_n is binomially distributed with parameters n and π. Let F_n stand for the cumulative distribution function of u_n. Notice that $F_n \geq F_{n+1}|_{(-\infty,n]}$ implies $w_{\lambda,n} \leq w_{\lambda,n+1}$, so it is enough to prove the former inequality to conclude that $w_{\lambda,n}$ is increasing in n. To this end, fix an $n \in \mathbb{N}$, take any $\theta \in \{0,...,n\}$, and note that

$$\begin{aligned}
F_{n+1}(\theta) &= \text{Prob}(u_{n+1} \leq \theta) \\
&= \text{Prob}(u_n < \theta) + \text{Prob}(u_{n+1} \leq \theta \mid u_n = \theta)\text{Prob}(u_n = \theta) \\
&= F_n(\theta - 1) + (1-\pi)\binom{n}{\theta}\pi^{\theta}(1-\pi)^{n-\theta} \\
&= \left(F_n(\theta-1) + \binom{n}{\theta}\pi^{\theta}(1-\pi)^{n-\theta}\right) - \pi\binom{n}{\theta}\pi^{\theta}(1-\pi)^{n-\theta} \\
&= F_n(\theta) - \pi\binom{n}{\theta}\pi^{\theta}(1-\pi)^{n-\theta} \\
&\leq F_n(\theta).
\end{aligned}$$

This proves our first assertion in Remark 4.

Next, fix any $n \in \mathbb{N}$ and $\pi \in [0,1]$ such that $\pi > \frac{1}{n}$. By Theorem 1 of Greenberg and Mohri (2014), we have $\text{Prob}(u \leq n\pi) \geq \frac{1}{4}$ for any binomially distributed random variable u with parame-

[25]This argument proves a bit more than what is needed for Lemma 3. It shows that $\frac{1}{n}\varphi_{n,c}'(\pi)$ equals $\text{Prob}(w \leq c-1)-$ $\text{Prob}(w \leq c)$ where w is a random variable with $w \sim \text{Binomial}(n-1,\pi)$.

ters n and π. Setting $\pi = \frac{\lambda}{1+\lambda}$ and picking any $\alpha \in [0, \frac{1}{4})$, it then follows from (6) that $w_{\lambda,n} \leq \frac{n\lambda}{1+\lambda}$ whenever $n > \frac{1+\lambda}{\lambda}$, as we claimed in Remark 4.

To prove the consistency assertion made in Remark 4, take any $\lambda \in (0, 1]$ and $n \in \mathbb{N}$, and let u_n be a binomially distributed random variable with parameters n and $\pi := \frac{\lambda}{1+\lambda}$. (Our claim in the case where $\lambda = 0$ is trivially true.) Note that $\text{Prob}(n\pi < u_n \leq \lceil n\pi \rceil)$ is either 0 (which happens when $n\pi$ is an integer) or it equals $\text{Prob}(u_n = \lceil n\pi \rceil)$. Since $\binom{n}{i}\pi^{i-1}(1-\pi)^{n-i} \to 0$ as $n \uparrow \infty$ for any positive integer i, therefore, we have

$$\text{Prob}(n\pi < u_n \leq \lceil n\pi \rceil) \to 0 \qquad \text{as } n \uparrow \infty.$$

Since every median of the binomial distribution with parameters n and π lies within $\lfloor n\pi \rfloor$ and $\lceil n\pi \rceil$ – see, for instance, Kaas and Buhrman (1980) – it follows that for every $\varepsilon > 0$ there is a positive integer N large enough that $\text{Prob}(u_n \leq n\pi) \geq \frac{1}{2} - \varepsilon$ for each $n \geq N$. In particular, for any α within, say, $(0, \frac{1}{4})$, we have

$$\text{Prob}(u_n \leq n\pi) > \alpha \qquad \text{for each } n \geq N,$$

which, by (6), means $w_{\lambda,n} \leq n\pi$ for each $n \geq N$. We conclude that $\limsup \frac{w_{\lambda,n}}{n} \leq \pi$.

On the other hand, by Hoeffding's Inequality,

$$\text{Prob}(u_n \leq (\pi - \varepsilon)n) \leq e^{-2\varepsilon^2 n} \qquad \text{for every } n \in \mathbb{N} \text{ and } \varepsilon > 0.$$

Consequently, for any $\varepsilon > 0$, $\text{Prob}(\frac{u_n}{n} \leq \pi - \varepsilon) \to 0$ as $n \uparrow \infty$. But by definition of $w_{\lambda,n}$,

$$\text{Prob}\left(\frac{u_n}{n} \leq \frac{w_{\lambda,n}}{n} - \frac{1}{n}\right) = \text{Prob}(u_n \leq w_{\lambda,n} - 1) \leq \alpha < \text{Prob}(u_n \leq w_{\lambda,n}) = \text{Prob}\left(\frac{u_n}{n} \leq \frac{w_{\lambda,n}}{n}\right)$$

for every $n \in \mathbb{N}$. It follows that, for any $\varepsilon > 0$, there is a positive integer N large enough that

$$\text{Prob}\left(\frac{u_n}{n} \leq \pi - \varepsilon - \frac{1}{n}\right) < \text{Prob}\left(\frac{u_n}{n} \leq \frac{w_{\lambda,n}}{n}\right) \qquad \text{for each } n \geq N,$$

whence

$$\pi - \varepsilon < \frac{w_{\lambda,n}}{n} + \frac{1}{n} \qquad \text{for each } n \geq N.$$

We conclude that $\pi - \varepsilon \leq \liminf \frac{w_{\lambda,n}}{n}$ for any $\varepsilon > 0$, which means $\pi \leq \liminf \frac{w_{\lambda,n}}{n}$. Combining this with what we have found in the previous paragraph yields $\lim \frac{w_{\lambda,n}}{n} = \pi = \frac{\lambda}{1+\lambda}$, as we sought.

References

Afriat, S., The construction of utility functions from expenditure data, *Int. Econ. Rev.* 8 (1967), 67-77.

Agranov, M., and P. Ortoleva, Stochastic Choice and Preferences for Randomization, *J. Pol. Econ.* 125 (2017), 40-68.

Agranov, M., and P. Ortoleva, Ranges of Randomization, *mimeo*, Princeton Univ., 2021.

Arrow, K.J., Rational choice functions and orderings, *Economica,* 102 (1959), 121-127.

Benjamimi, Y., and Y. Hochberg, Controlling the false discovery rate: a practical and forceful approach to multiple testing, *J. Royal Statist. Soc. Ser B.* 57 (1995), 289-300.

Bernheim, D., and A. Rangel, Toward choice-theoretic foundations for behavioral welfare economics, *Amer. Econ. Rev.* 97 (2007), 464-70.

Bernheim, D., and A. Rangel, Beyond revealed preference: Choice-theoretic foundations for behavioral welfare economics, *Quart. J. Econ.* 124 (2009), 51-104.

Bouacida, E., Eliciting choice correspondences: A general method and an experimental implementation, *mimeo*, Paris School of Economics, 2018.

Cerreia-Vioglio, S., D. Dillenberger, P. Ortoleva, and G. Riella, Deliberately stochastic, *Amer. Econ. Rev.* 109 (2019), 2425-45.

Cettolin, E. and A. Riedl, Revealed preferences under uncertainty: Incomplete preferences and preferences for randomization, *J. Econ. Theory* 181 (2019), 547–85.

Chambers, C., and F. Echenique, *Revealed Preference Theory*, Cambridge, Cambridge University Press, 2016.

Cohen, M., J-Y. Jaffray, and T. Said, Individual behavior under risk and uncertainty: An experimental study, *Theory Decis.* 18 (1985), 203-228.

Cohen, M., J-Y. Jaffray, and T. Said, Experimental comparison of individual behavior under risk and under uncertainty for gains and for losses, *Organ. Behav. Hum. Decis. Process.* 39 (1987), 1-22.

Costa-Gomes, M., C. Cueva, G. Gerasimou, and M. Tejiščák, Choice, deferral and consistency, *Quantitative Economics*, forthcoming, 2021.

Danan, E. and A. Ziegelmeyer, Are preferences complete?: An experimental measurement of indecisiveness under risk, *mimeo*, Max Planck Inst. of Economics, 2006.

Dwenger, N., D. Kübler, and G. Weizsäcker, Flipping a coin: Evidence from laboratory and field, *J. Pub. Econ.* 167 (2018), 240-250.

Eliaz, K. and E. A. Ok, "Indifference or Indecisiveness? Choice Theoretic Foundations of Incomplete Preferences," *Games Econ. Behav.* 56 (2006), 61-86.

Feldman, P. and J. Rehbeck, Revealing a preference for mixing: An experimental study of risk, *mimeo*, UC-San Diego, 2020.

Fishburn, P., Choice probabilities and choice functions, *J. Math. Psych.* 18 (1978), 201-219.

Gerasimou, Georgios, Model-Rich Approaches to Eliciting Possibly Weak or Incomplete Preferences: Evidence from a Multi-Valued Choice Experiment, *mimeo*, University of St. Andrews, 2021.

Greenberg, S. and M. Mohri, Tight lower bound on the probability of a binomial exceeding its expectation, *Stat. Prob. Lett.*, 86 (2014), 91-98.

Holm, S., A simple sequentially rejective multiple test procedure, *Scand. J. Statist.* 6 (1979), 65–70.

Houthakker, H.S., Revealed preference and the utility function, *Economica*, 17 (1950), 159-174.

Iverson, G., and J.-C. Falmagne, Statistical issues in measurement, *Math. Soc. Sci.* 10 (1985), 131-153.

Johnson, N., S. Kotz, and N. Balakrishnan, *Discrete Multivariate Distributions*, New York, Wiley, 1997.

Kaas, R. and J. Buhrman, Mean, median and mode in binomial distributions, *Stat. Neerlandica,* 34 (1980) 13-18.

Kircher, P., S. Ludwig, and A. Sandroni, On the difference between social and private goods, *BE J. Theoret. Econ.* 13 (2013), 1-27.

Lehmann, E., and J. Romano, *Testing Statistical Hypotheses*, 3rd ed., New York, Springer, 2005.

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

Manzini, P. and M. Mariotti, Sequentially rationalizable choice, *Amer. Econ. Rev.* 97 (2007), 1824-1830.

Manzini, P. and M. Mariotti, Dual random utility maximisation, *J. Econ. Theory* 177 (2018), 162-182.

Masatlioglu, Y., D. Nakajima, and E. Ozbay, Revealed attention, *Amer. Econ. Rev.* 102 (2012), 2183-2205.

Mas-Colell, A., M. Whinston, and J. Green, *Microeconomic Theory*, New York: Oxford University Press, 1995.

Miao, B. and S. Zhong, Probabilistic social preference: How Machina's Mom randomizes her choice, *Econ Theory* 65 (2018), 1-24.

Nishimura, H., and E. A. Ok, Non-Existence of Continuous Choice Functions, *J. Econ. Theory* 153 (2014), 376-391.

Ok, E. A., and G. Tserenjigmid, Comparative rationality of stochastic choice rules, *mimeo*, New York Univ., 2020.

Ok, E. A., and G. Tserenjigmid, Indifference, indecisiveness, experimentation and stochastic choice, *Theoretical Econ.* 17 (2022), 651-686.

Regenwetter, M., J. Dana, and C. Davis-Stober, Transitivity of preferences, *Psych. Rev.* 118 (2011), 42-56.

Rubinstein, A., Irrational diversification in multiple decision problems, *European Econ. Rev.* 46 (2002), 1369–78.

Samuelson, P.A., A note on the pure theory of consumer's behaviour, *Economica* 5 (1938), 61-71.

Samuelson, P.A., Consumption theory in terms of revealed preference, *Economica* 60 (1950), 243-253.

Sautua, S., Does uncertainty cause intertia in decision making? An experimental study of the role of regret aversion and indecisiveness, *J. Econ. Behav. Org.* 136 (2017), 1-14.

Tserenjigmid, G., Theory of decisions by intra-dimensional comparisons, *J. Econ. Theory* 159 (2015), 326-338.

Tversky, A., Intransitivity of preferences, *Psych. Rev.* 76 (1969), 31-48.