

# Weighing Sample Evidence

Szwagrzak, Karol

*Document Version*  
Final published version

*Publication date:*  
2021

*License*  
Unspecified

*Citation for published version (APA):*  
Szwagrzak, K. (2021). *Weighing Sample Evidence*. Copenhagen Business School [wp]. Working Paper / Department of Economics. Copenhagen Business School No. 03-2021

[Link to publication in CBS Research Portal](#)

## General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

## Take down policy

If you believe that this document breaches copyright please contact us (research.lib@cbs.dk) providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 13. May. 2025





**Copenhagen  
Business School**  
HANDELSHØJSKOLEN

# **Department of Economics**

**Copenhagen Business School**

**Working paper 03-2021**

**Weighing sample evidence**

**Karol Flores-Szwagrzak**

# Weighing sample evidence

Karol Flores-Szwagrzak\*

## Abstract

We model how sample evidence guides choice: An agent faces a number of alternative actions. For each action, she observes a sample of outcomes; she cannot see the distribution from where the sample was drawn. To make her choice, she evaluates the evidence for the hypothesis that an action is optimal. The *strength* of evidence in favor of the hypothesis is measured by the average decision utility of the outcomes in its sample; its *weight* gauges predictive validity and is approximated by the size of the sample. We identify necessary and sufficient conditions for her choice to be determined by the interaction of strength and weight, reflecting the determinants of confidence judgements documented in experiments (Griffin and Tversky, 1992). These conditions characterize a model consistent with non-trivial uncertainty attitudes and “frequentist” expected utility maximization.

**Keywords:** sample, sample size, uncertainty attitudes

**JEL codes:** D81, D83

---

\*Department of Economics, Copenhagen Business School, Porcelænshvaen 16A, Frederiksberg, Denmark. E-mail: ksz.eco@cbs.dk. Phone: (45)21178011. I thank Geir Asheim, Christopher Chambers, Tuomas Laiho, Bård Harstad, Ying He, David Jinkins, Andy Mackenzie, Isabelle Mairey, Thierry Marchant, Sergio Ocampo Diaz, Jurre Thiel, William Thomson, and Rafael Treibich for useful discussions. I am especially grateful to Asen Kochov for a number observations and challenging questions. Comments from seminar participants at the University of Oslo, Copenhagen Business School, and the University of Rochester are also gratefully appreciated. All errors are my own. My research is generously supported by the Danish Independent Research Council (DFR-Forskningsprojekt-1, 9038-00058A).

# 1 Introduction

After careful research and consultations, a person suffering from a rare illness has narrowed down two alternative treatments. Ideally, before making her choice, she would experiment with each treatment; if a degree of chance is involved in the effectiveness of a treatment, she would also experiment a large enough number of times to ensure predictive validity. In practice, the information available to her are the samples of outcomes experienced by subjects exposed to the two treatments. Indeed, the fundamental evidence across the social and medical sciences comes in the form of samples specifying the outcomes experienced by subjects exposed to an intervention or treatment.

Samples are useful when the processes generating outcomes are not directly observable. If they were observable, each action (e.g., a treatment choice) could be thought of as a draw from a known probability distribution. Several theories examine choice in this domain including expected utility (Von Neumann and Morgenstern, 1944), rank dependent expected utility (Quiggin, 1982; Yaari, 1987), and prospect theory (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992). In this paper, we study choice when each action can be thought of as a draw from an *unknown* distribution; the only information on the action is a sample of outcomes drawn from this distribution.

Of course, a sample can be used to make inferences about the underlying distribution. But, exact knowledge of this distribution cannot be derived from any finite number of observations; it is the limit approachable but not attainable by extended observation (Koopmans, 1949). Moreover, inference always requires assumptions about the distribution and the sampling process; the credibility of these assumptions is a subjective matter (Manski, 2003). Uncertainty in the sense of Knight (1921) is inescapable; even very good data will only mitigate the problem. The empirical frequency distributions or posterior distributions cannot simply be taken to stand for the actual distributions without a loss of relevant information, e.g., sample size, confidence intervals, prior or parameter assumptions, etc. Even if the use of a prior is fully justified, a posterior distribution is just a point estimate of the true distribution; its parameters have their own distributions.

We directly model how samples shape choice: An agent, the decision-maker, determines her confidence in the hypothesis that a sample signals her best draw. To do so, she accounts for the *strength* and *weight* of the evidence in favor of the hypothesis; these are the two determinants of confidence in judgements documented in experiments by [Griffin and Tversky \(1992\)](#). We quantify the strength of evidence in a sample as the average subjective utility of its outcomes; weight reflects predictive validity and we use sample size as its proxy. We provide necessary and sufficient axioms for choice to depend solely on strength and weight.

As a first approximation, decisions could be driven by strength alone. The attractiveness of outcomes is measured by a decision utility  $v$  assigning a numerical score to each outcome. The strength of a sample with  $n$  outcomes  $\alpha_1, \dots, \alpha_n$  is then the average decision utility over these outcomes,

$$\frac{1}{n} \sum_{i=1}^n v(\alpha_i). \quad (1)$$

If the agent chooses maximal strength alternatives, disregarding sample size, her behavior would be consistent with the “law of small numbers” ([Tversky and Kahneman, 1971](#)), the common but mistaken presumption that a small sample can be used to arrive at valid inferences about the distribution from where it was drawn.

In general, the agent factors both strength and weight as if equipped with an evaluation function  $V : \mathbb{R} \times \mathbb{N} \rightarrow \mathbb{R}$ . A sample consisting of  $n$  outcomes  $\alpha_1, \dots, \alpha_n$  is ranked according to the value

$$V\left(\frac{1}{n} \sum_{i=1}^n v(\alpha_i), n\right). \quad (2)$$

The value is always increasing in strength (the first argument). The interaction between strength and weight (the second argument) can be more intricate: If strength is above an endogenously determined threshold, increasing the sample size while keeping strength fixed increases the value; if strength is below a possibly different threshold, increasing sample size decreases the value. Depending on strength, the agent can thus prefer more certain or more uncertain information regarding the possible outcomes of her actions. However, as sample size increases, uncertainty

vanishes; the dependence on sample size diminishes. When the weight of evidence is large enough, actions are compared primarily on their strength.

Choice is thus consistent with the Law of Large Numbers and expected utility maximization: If the sample is sufficiently large, the relative frequency of each outcome approximates its probability; strength then approximates expected utility. The axiomatization of the model thus provides foundations for a subjective “frequentist” expected utility model accommodating non-trivial uncertainty attitudes.

The remainder of this paper is organized as follows: Section 2 explains where we depart from the closest references in the decision theory literature. Section 3 formally introduces the domain of decisions studied in this paper. Section 4 presents the model and examples. Section 5 presents the axioms and characterizations of the model. Section 6 concludes. The proofs are gathered in an Appendix.

## 2 Related literature

The classical account of decision under uncertainty (Savage, 1954) assumes an agent that understands all the possible resolutions of uncertainty conditional on the realized “state of nature.” For each action, if she knows the state, she knows the outcome. Sampling via randomized controlled trials is an especially powerful tool to identify and quantify causal relationships even without knowing the exact mechanisms driving these relationships, i.e., when the relevant states cannot be specified. Moreover, it is well understood that there are decision problems where the states of nature are not given naturally, nor can they be readily formulated (e.g., Gilboa and Schmeidler, 1995; Ahn, 2008). This has stimulated the development of alternative, complementary theories close to ours.

*Case-based decision theory* (Gilboa and Schmeidler, 1995, 2001; Gilboa et al., 2002) examines how decisions are guided by data on past “cases.” Each case specifies the decision problem, the action chosen for the problem, and the resulting outcome. Consider, for example, a medication (an action) treating different diseases (different problems); a case then records the health outcome experienced by a person afflicted by a given disease who used the medication. The key feature of case-based decision models is a similarity function describing how information is harnessed across

different related problems and actions (Gilboa and Schmeidler, 1996, 2001). Our framework is simpler: Similarities between actions can only be gauged on the basis of their outcomes, not on any other observable characteristics of the actions or problems.

Accordingly, we focus on the weighing of sample evidence, not on drawing analogies to past cases. In our framework, standard case-based models boil down to ranking actions on the basis of the average utility of the outcomes in their sample or on the basis of the summation of the utilities of these outcomes (pages 48, 74 in Gilboa and Schmeidler, 2001). The average is what we referred to as strength in (1) and the summation for a sample consisting of  $n$  outcomes  $\alpha_1, \dots, \alpha_n$  is

$$\sum_{i=1}^n v(\alpha_i).$$

There are basic drawbacks to both strength and summation as evaluation criteria: First, controlling for strength, a larger sample offers greater confidence and may carry more weight for an uncertainty averse agent. Second, summation puts an excessive premium on sample size. Large samples mechanically deliver high sums. The result is inertia, with choices driven by past popularity, i.e., large samples.<sup>1</sup> These observations suggest the agent will balance strength and sample size, the core issue examined in this paper.

Another strand of theories forgoing the states of nature concept models preferences over *menus*, i.e., sets of possible outcomes. The agent chooses a menu and then, in a later stage interpreted as that following the resolution of uncertainty, selects an outcome from her menu (Kreps, 1979); alternatively, the agent may have no control over the outcome in her chosen menu (Fishburn, 1972). Menus are fundamentally different from samples: A sample drawn from a distribution need not contain all foreseeable outcomes; the support of the distribution generally contains outcomes not present in the sample. Moreover, a sample may contain multiple observations of the same outcome, enabling learning about its likelihood; the sample's size may also influence its perceived predictive validity.

---

<sup>1</sup>Gilboa and Schmeidler (1996) addresses these issues by embedding case-based decision theory in a dynamic framework where an agent facing a multi-armed bandit adjusts her aspiration level.

The most relevant theories of menu choice concern situations of *complete uncertainty*, where the agent has no control over the outcome in her chosen menu. Decisions can then be modelled as maximizing expected utility under uniform beliefs over the outcomes in the menus (Fishburn, 1972; Gravel et al., 2012). The characterizations of this model rely on the following “betweenness” axiom:<sup>2</sup> If a menu is preferred over another, then the menu described by the union of the outcomes in the two original menus is between them in the preference ranking. Applying betweenness to samples would preclude size considerations, ruling out strict preferences over equal strength actions. These preferences may however reflect the varying confidence in outcome estimates across samples. Ruling them out would *require* decisions to be consistent with the spurious “law of small numbers” (Tversky and Kahneman, 1971).

Recent research on menu choice has sought sharper characterization results by enriching the choice set to include lotteries over outcomes (e.g., Dekel et al., 2001, 2007, 2009; Gul and Pesendorfer, 2001; Olszewski, 2007; Ahn, 2008; Dekel and Lipman, 2012). Under the assumption that the agent has no control over the outcome in her chosen menu, this framework thus enables modelling uncertainty about the distribution outcomes will be drawn from (Olszewski, 2007; Ahn, 2008). Preferences over menus reveal uncertainty attitudes: Singleton menus are certain since the agent knows the distribution she will draw from; larger sets are *objectively ambiguous/uncertain* since the agent does not know what distribution in the menu she will draw from. The agent’s problem is then to determine how much credence to put on the foreseeable distributions in each menu. Unlike in our paper, the question of where distributions come from is not addressed; distributions are directly observable. At a technical level, the domain specifies preferences over sets containing uncountably many distributions; this makes modelling attitudes toward sample size impossible. At a conceptual level, the key betweenness-type axioms characterizing the models proposed by Olszewski (2007) and Ahn (2008) are also incompatible with concerns for sample size, by essentially the same arguments discussed above.

Technically, our paper is close to that of Kothiyal, Spinu, and Wakker (2014) who arrived at a general characterization of an average utility model to evaluate sequence

---

<sup>2</sup>Gravel et al. (2012) refers to this axiom as “averaging.”



of outcomes of variable length. Their goal is to provide weak and transparent necessary and sufficient preference axioms for the average utility model, not to study the weighing of sample evidence and the associated uncertainty attitudes.

### 3 Definitions

A *sample* consists of a finite number of outcomes drawn from a set  $\mathbb{O}$  of possible outcomes. The collection of all samples is denoted by  $\mathbb{S}$ . For each sample  $S \in \mathbb{S}$  and each outcome  $\alpha \in \mathbb{O}$ ,  $S(\alpha)$  denotes the number of times the outcome appears in the sample.

An action is interpreted as a draw from an unknown distribution over  $\mathbb{O}$ . The agent can only observe a sample drawn from this distribution. Preferences over actions can thus be recast as preferences over the corresponding samples. Let  $\succsim$  denote a complete and transitive binary relation over  $\mathbb{S}$ . The symmetric and asymmetric parts of  $\succsim$  are denoted by  $\sim$  and  $\succ$ , respectively.

For each  $S \in \mathbb{S}$ , the *size* of  $S$  is number of outcomes in  $S$  and is denoted by  $|S|$ . Thus,  $|S| = \sum_{\alpha \in \mathbb{O}} S(\alpha)$ . For each pair  $S, T \in \mathbb{S}$ ,  $S \oplus T$  denotes the sample obtained by pooling the outcomes in  $S$  and  $T$ . Thus, for each  $\alpha \in \mathbb{O}$ ,  $(S \oplus T)(\alpha) = S(\alpha) + T(\alpha)$ . For each positive integer  $n$  and each  $S \in \mathbb{S}$ ,  $n \otimes S$  denotes the sample obtained pooling  $n$  copies of  $S$ . Thus, for each  $\alpha \in \mathbb{O}$ ,  $(n \otimes S)(\alpha) = n \times S(\alpha)$ . Abusing notation, the sample consisting of a single outcome  $\alpha \in \mathbb{O}$  is denoted by  $\alpha$ .

### 4 Model

Griffin and Tversky (1992) proposed an intuitive explanation of confidence formation in judgements: People assess the degree to which the available evidence supports a hypothesis, i.e., the *strength* of evidence in favor of the hypothesis, and then adjust for predictive validity, i.e., the *weight* of evidence. For example, in deciding if a coin is biased in favor of heads, the strength is the proportion of heads in a sample of coin tosses and the weight is the size of the sample. People value both strength and weight but the exact tradeoff between the two is an empirical question. In many settings

however people are systematically overly influenced by strength to the detriment of weight relative to a Bayesian benchmark.

We propose a model where strength and weight considerations explain choice behavior; special cases allow for the bias documented by [Griffin and Tversky \(1992\)](#) as well as natural Bayesian benchmarks. To arrive at her choice between two actions, an agent determines the extent to which a sample signals attractive outcomes, i.e., the strength of the evidence in the sample. The agent then adjusts for weight or credence.

We formalize **strength** as the average decision utility of the outcomes in a sample. Given a decision utility function  $v : \mathbb{O} \rightarrow \mathbb{R}$  and a sample  $S$ , the average value taken by  $v$  over the outcomes in the sample is denoted by  $A(v, S)$ , i.e.,

$$A(v, S) = \sum_{\alpha \in \mathbb{O}} \frac{S(\alpha)}{|S|} v(\alpha). \quad (3)$$

The range of values taken by this average over all samples is denoted by  $A_v$ , i.e.,

$$A_v = \{x \in \mathbb{R} : \exists S \in \mathbb{S}, A(v, S) = x\}.$$

A **strength & weight representation** of  $\succsim$  is a pair of functions  $v, V$  such that  $v : \mathbb{O} \rightarrow \mathbb{R}$  is not constant,  $V : A_v \times \mathbb{N} \rightarrow \mathbb{R}$  is strictly increasing in its first argument, and, for each pair  $S, T \in \mathbb{S}$ ,

$$S \succsim T \quad \Leftrightarrow \quad V(A(v, S), |S|) \geq V(A(v, T), |T|). \quad (4)$$

We will introduce two additional properties that a strength & weight representation may satisfy. Before doing this, we provide three examples of representations satisfying these properties. The first example illustrates that our model is consistent with the textbook example of Bayesian updating.

**Example 1.** *There are two possible outcomes,  $\mathbb{O} = \{\alpha, \beta\}$ . Outcome  $\alpha$  is preferred so  $v(\alpha) = 1$  and  $v(\beta) = 0$ . The agent has a prior over the probability of an  $\alpha$  realization, a Beta distribution with parameters  $a, b > 0$ ; the mean of this prior is  $p = \frac{a}{a+b}$ . She evaluates a sample according to the posterior probability of the outcome being  $\alpha$  conditional on observing the sample. That is, for each  $S \in \mathbb{S}$ ,*

$$V(A(v, S), |S|) = \text{Prob}(\alpha|S) = \frac{|S|}{a + b + |S|} A(v, S) + \frac{a + b}{a + b + |S|} p.$$

A similar strength & weight representation applies to situations with more than two outcomes.

**Example 2.** *The agent starts out expecting an outcome  $\alpha_0 \in \mathbb{O}$ . As evidence accumulates in the form of a larger sample, she turns more sensitive to sample information. Then, for each  $S \in \mathbb{S}$ ,*

$$V(\mathbf{A}(v, S), |S|) = w_{|S|} \mathbf{A}(v, S) + (1 - w_{|S|}) v(\alpha_0)$$

where  $\{w_n\}$  is an increasing sequence of real numbers in  $[0, 1]$  with  $w_n \rightarrow 1$ . Large samples are thus given more weight.

**Example 3.** *The outcomes are possible changes in wealth and  $\mathbb{O}$  can be taken to be the real line. Outcome 0 represents no change in wealth and we let  $v(0) = 0$ . The agent evaluates strength values with a function  $\mathcal{V} : \mathbb{R} \rightarrow \mathbb{R}$  that is continuous, strictly increasing, and such that  $\mathcal{V}(0) = 0$ . She accounts for weight by discounting geometrically according to a parameter  $r \in (0, 1)$ . For each  $S \in \mathbb{S}$ ,*

$$V(\mathbf{A}(v, S), |S|) = (1 - r^{|S|}) \mathcal{V}(\mathbf{A}(v, S)).$$

*This strength & weight representation is inspired by the representation of time preferences proposed by [Fishburn and Rubinstein \(1982\)](#) where a monetary gain or loss in the future is discounted geometrically, allowing for impatience in the domain of gains and procrastination in the domain of losses.*

The first additional property of a strength & weight representation, “convergence,” specifies that, when two samples become large enough, they are primarily compared on the basis of their strength. The represented preferences are then consistent with expected utility maximization relative to the empirical distribution functions inferred from the samples. A strength & weight representation  $v, V$  is **convergent** if there is a strictly increasing function  $\mathcal{V} : \mathbf{A}_v \rightarrow \mathbb{R}$  such that,

$$\text{for each } x \in \mathbf{A}_v, V(x, n) \rightarrow \mathcal{V}(x) \text{ as } n \rightarrow \infty.$$

The second additional property of a strength & weight representation “orders” uncertainty comparisons across strength-weight pairs. Keeping strength constant,

more certainty in the form of a larger sample is either always good news or it is always bad news. Moreover, if more certainty is good (bad) news at a given strength level, it is also good (bad) news at a higher (lower) strength level. A strength & weight representation  $v, V$  is **ordered** if its range of strength values  $\mathbf{A}_v$  can be partitioned into up to three sets  $B, N, G$  such that,

$$\begin{aligned} &\text{for each } (x, y, z) \in B \times N \times G, x < y < z, \text{ and, for each positive integer } n, \\ &V(x, n) > V(x, n + 1), \quad V(y, n) = V(y, n + 1), \quad V(z, n) < V(z, n + 1). \end{aligned} \quad (5)$$

Informally,  $B$  consists of the strength levels of samples that deliver “bad” outcomes on average,  $N$  consists of the strength levels of samples delivering “neutral” outcomes on average, and  $G$  consists of the strength levels of samples delivering “good” outcomes on average. Some of these sets may be empty: The case where  $B$  and  $N$  are empty can be thought of as describing a “gains domain” where a higher weight is always desirable. The case where  $N$  and  $G$  are empty describes a “losses domain” where the opposite is true. When  $B$  and  $G$  are empty, sample size is irrelevant and evaluation is driven by strength alone.

In Example 1, the partition depends on the value of the prior mean  $p$ :  $B$  consists of all strength levels strictly smaller than  $p$ ,  $N$  consists of  $p$ , and  $G$  consists of all strength levels strictly greater than  $p$ . Analogously, in Example 2, the partition depends on the value  $v(\alpha_0)$ . In Example 3,  $B$  consists of all negative real numbers,  $N$  consists of 0, and  $G$  consists of all positive real numbers.

Figure 1 illustrates a convergent and ordered strength & weight representation  $v, V$  when the second argument in  $V$  takes values  $n$  and  $k$  with  $n > k$ . The representation also becomes weight-independent for sufficiently large samples: The 45° line describes an evaluation function that does not depend on its second argument, i.e., that depends only on strength. As the second argument increases from  $k$  to  $n$  the evaluation function comes closer to the 45° line. Evaluation is primarily based on strength for large samples. For smaller samples, e.g., of sizes  $n$  and  $k$ , uncertainty attitudes can display a form of reference dependence: Actions can be evaluated differently depending on whether their strength is below  $c$  (in  $B$ ) or above  $d$  (in  $G$ ). Holding strength constant beyond  $d$ , increasing sample size yields a preferable action; holding strength constant below  $c$ , increasing sample size yields a worse one.

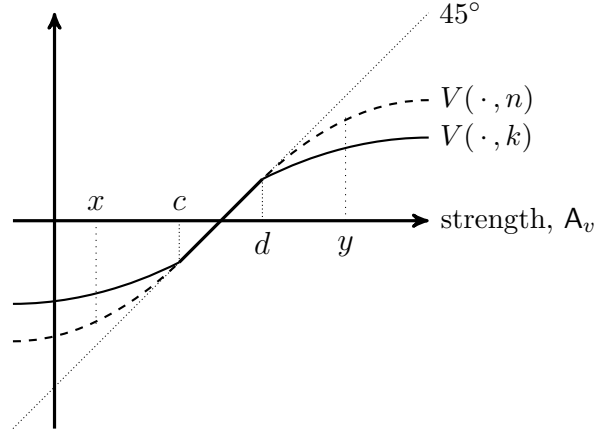


Figure 1: An illustration of a convergent and ordered strength & weight representation  $v, V$ . Here  $B$  is the range of strength values below  $c$ ,  $N$  is the range of strength values between  $c$  and  $d$ , and  $G$  is the range of strength values above  $d$ . Observe that, for each pair of positive integers  $k, n$  with  $k < n$ ,  $V(x, n) < V(x, k)$  if  $x < c$ ,  $V(y, n) > V(y, k)$  if  $y > d$ , and  $V(z, n) = V(z, k)$  if  $z$  is between  $c$  and  $d$ .

## 5 Axioms

The first axiom specifies that sample data matters to the agent. She ought to be *sensitive* towards outcome variations across samples. However, with very few observations, these variations may be due to chance. We therefore only require that the agent can express strict preferences between samples of a possibly very large size.

**A1 Sensitivity:** There are  $S, T \in \mathbb{S}$  of equal size satisfying  $S \succ T$ .

A weaker “non-triviality” axiom would specify strict preferences between two arbitrary samples. Conforming with non-triviality but not with sensitivity would imply indifference between any two samples of the same size. Implausibly, preferences would then be driven solely by sample size.

The second axiom is an independence condition. It specifies under what circumstances further outcome observations do not affect the evaluation of two samples. Suppose that the agent is comparing distributions  $F$  and  $G$  on the basis of their samples and determines that  $F$  offers the better draw. Now suppose that, just before taking her draw from  $F$ , she observes one further outcome drawn from  $F$  and one

further outcome drawn from  $G$ . Our axiom specifies that she won't change her mind *if* the new outcomes are identical and the samples are of the same size. The marginal effect of one further equal observation is not sufficient to overturn her choice. The extra outcomes *cancel* out in the comparison.

**A2 Cancellation:** For each pair  $S, T \in \mathbb{S}$  of equal size and each  $\alpha \in \mathbb{O}$ ,  $S \succ T$  implies  $S \oplus \alpha \succ T \oplus \alpha$  and  $S \sim T$  implies  $S \oplus \alpha \sim T \oplus \alpha$ .

The restriction to samples of equal size is important: Adding an observation to a sample with ten outcomes enriches it more than adding it to a sample with ten thousand. This reflects a fundamental psychophysical principle, the Weber-Fechner law: Stimuli perception is decreasing in the level of the stimulus, i.e., change is perceived in a log scale. In our context, the relevance of an outcome observation is marginally decreasing in the size of the sample.

In situations where there are two possible outcomes (a success and failure, a win and a loss, survival and death, etc.) sensitivity and cancellation are necessary and sufficient conditions for preferences to have a strength & weight representation. Additionally, the average decision utility and the fraction of preferred outcomes are identical measures of strength, as noted in the equivalence of the second and third statements below.

**Theorem 1.** *If the set of outcomes  $\mathbb{O}$  has a cardinality of two, then the following statements are equivalent.*

- i.  $\succsim$  satisfies sensitivity and cancellation.*
- ii.  $\succsim$  has a strength  $\mathcal{E}$  weight representation  $v, V$ .*
- iii. There are  $\alpha, \beta \in \mathbb{O}$  satisfying  $\alpha \succ \beta$  and a function  $W : [0, 1] \times \mathbb{N} \rightarrow \mathbb{R}$  that is strictly increasing in its first argument such that, for each pair  $S, T \in \mathbb{S}$ ,*

$$S \succsim T \iff W(S(\alpha)/|S|, |S|) \geq W(T(\alpha)/|T|, |T|).$$

Independence axioms akin to cancellation but lacking the equal size restriction have been used extensively in theories of menu choice (Barberà et al., 2004). They are central and referred to as “monotonicity” and sometimes as “cancellation” axioms in the theory of measurement (Krantz et al., 1971). The restriction to equally sized menus of foreseeable outcomes has been used in the complete uncertainty framework;

it is one of the axioms characterizing the uniform expected utility criterion (Gravel et al., 2012). An analogous restriction has also been used to characterize the average utility criterion in the comparison of sequences of outcomes of different lengths (Kothiyal et al., 2014).<sup>3</sup>

## 5.1 Identifying uncertainty attitudes

Choice can be driven by the attractiveness of outcomes and by their perceived uncertainty, as inferred from sample data. We can identify the role of uncertainty by varying its degree while controlling for outcomes. Comparing two samples with identical relative outcome frequencies, the larger sample yields a more precise signal of the underlying distribution. This can be interpreted as a decrease in uncertainty since the larger sample enables a more precise approximation of the likelihood of each outcome, yet the likelihood estimate based on the relative frequency is unchanged. All else equal, the agent may then prefer an action with a larger sample. However, she may also prefer a smaller sample: An action with a sample consisting solely of a single “bad” outcome may seem safer than an action with a sample of a hundred observations of the exact same bad outcome; the sample for the second action may suggest that the realization of the bad outcome is more certain. There are also situations where the agent may avoid more precise information: People at risk for health conditions often avoid costless medical tests that should help them make better decisions (Golman et al., 2017; Hertwig and Engel, 2016).

Given samples  $S$  and  $T$ , we will say that  $S$  *reduces uncertainty in*  $T$  if the relative frequencies of outcomes in  $S$  and  $T$  are identical and  $S$  is larger than  $T$ .<sup>4</sup> If sample  $S$  reduces uncertainty over sample  $T$  and  $S \succ T$ , we will say that *reducing uncertainty is beneficial*. Having identified uncertainty attitudes, we can specify consistency properties these attitudes may satisfy.

The third axiom specifies that preferences are *monotone* with respect to uncertainty reductions: After a beneficial reduction in uncertainty, each successive

---

<sup>3</sup>The analogue to cancellation in Gravel et al. (2012) is referred to as “restricted independence;” in Kothiyal et al. (2014) it is referred to as “expansion independence.”

<sup>4</sup>For each  $\alpha \in \mathbb{O}$ ,  $S(\alpha)/|S| = T(\alpha)/|T|$ , and  $|S| > |T|$ .

reduction is beneficial as well.

**A3 Monotonicity:** For each  $S \in \mathbb{S}$  and each pair  $k, l$  of strictly positive integers,  $(k + 1) \otimes S \succsim k \otimes S$  implies  $(l + 1) \otimes S \succsim l \otimes S$ .

Knowing whether an agent finds reducing uncertainty beneficial for a sample  $S$  can also enable us to infer if she finds reducing uncertainty beneficial for a sample  $T$  not related to  $S$  by an uncertainty reduction. Suppose that reducing uncertainty is beneficial for a sample  $T$ . Consider a sample  $S$  of the same size as  $T$  that is not related to  $T$  by an increase in uncertainty. A preference for  $S$  is then due to the attractiveness of its outcomes vis-à-vis those in  $T$ . Then, reducing uncertainty in  $S$ , making its better outcomes more certain, ought to be beneficial as well. This criterion enables *sorting* the samples of a given size into those that benefit from uncertainty reductions and those that do not. The fourth axiom summarizes these observations.

**A4 Sorting:** For each pair  $S, T \in \mathbb{S}$  of equal size with  $S \succsim T$ ,  $2 \otimes T \succ T$  implies  $2 \otimes S \succ S$  and  $2 \otimes S \prec S$  implies  $2 \otimes T \prec T$ .

The fifth axiom specifies when two samples are “large enough” to mitigate uncertainty in their comparison. Given a sample, an agent can reasonably become confident in her knowledge of the underlying distribution if the relative frequency of outcomes stabilises after successive sample enlargements. By the Law of Large Numbers, the relative frequency of each outcome then approaches its probability. If we compared two samples and found that after successively scaling both up a large enough number of times, the agent prefers one of the two scaled up samples. Sample size then ceases to be a determinant factor in the comparison.

**A5 Approachability:** For each pair  $S, T \in \mathbb{S}$ , there is a strictly positive integer  $n$  such that, if  $n \otimes S \succ n \otimes T$ , then  $k \otimes S \succ l \otimes T$  for all integers  $k, l$  greater than  $n$ .

Approachability reflects the idea that the agent is ultimately interested in the outcomes of her choice. Sample size is just a proxy for outcome uncertainty; it only has instrumental value in learning about the outcome distribution.

**Theorem 2.** *If the set of outcomes  $\mathbb{O}$  has a cardinality of two, then the following statements are equivalent.*



- i.  $\succsim$  satisfies sensitivity, cancellation, monotonicity, sorting, and approachability.
- ii.  $\succsim$  has a convergent and ordered strength & weight representation  $v, V$ .

## 5.2 Continuous outcomes

Outcomes like commodity bundles, changes in wealth, and waiting times are often modelled as continuous variables. Sample comparisons can then be required to be “continuous,” i.e., robust to small outcome perturbations. For example, if samples  $S$  and  $T$  satisfy  $S \succ T$ , we should be able to slightly perturb the outcomes in  $S$  and thereby obtain a new sample  $R$  such that  $R \succ T$ . Formally, assuming that the space of possible outcomes is a convex subset of Euclidean space, there is a sufficiently small number  $\varepsilon > 0$  such that  $R \succ T$  if the greatest Euclidean distance between the outcomes in  $S$  and those in  $R$  is less than  $\varepsilon$ . The fact that  $R$  is obtained by perturbing the outcomes in  $S$  requires  $R$  and  $S$  to be of equal size;  $T$  can be of a different size.

For each pair of equally sized samples  $R$  and  $S$ , the *distance between  $R$  and  $S$*  is the greatest Euclidean distance between the outcomes in  $R$  and  $S$ . We denote this distance by  $d(R, S)$ .<sup>5</sup> A subset of samples  $\mathcal{S} \subseteq \mathbb{S}$  is *open* if each of its elements is in its interior.<sup>6</sup>

**A6 Continuity:** For each  $S \in \mathbb{S}$ ,  $\{T \in \mathbb{S} : T \succ S\}$  and  $\{T \in \mathbb{S} : S \succ T\}$  are open.

Under continuity, our characterizations can be extended to the case where the set of outcomes is a convex subset of Euclidean space such as a commodity space. The corresponding strength & weight representation  $v, V$  is then such that  $v$  is continuous and  $V$  is continuous in its first argument.

**Theorem 3.** *If the set of outcomes  $\mathbb{O}$  is a convex subset of Euclidean space, then the following statements are equivalent.*

- i.  $\succsim$  satisfies sensitivity, cancellation, and continuity.
- ii.  $\succsim$  has a strength & weight representation  $v, V$ . Moreover,  $v$  is continuous and  $V$  is continuous in its first argument.

---

<sup>5</sup> $d(R, S)$  is the maximum value attained by  $\|\alpha - \beta\|$  over  $\alpha, \beta \in \mathbb{O}$  with  $R(\alpha) \geq 1$  and  $S(\beta) \geq 1$ .

<sup>6</sup>For each  $S \in \mathcal{S}$ , there is  $\varepsilon > 0$  such that  $\{R \in \mathbb{S} : d(S, R) < \varepsilon\}$  is contained in  $\mathcal{S}$ .

**Theorem 4.** *If the set of outcomes  $\mathbb{O}$  is a convex subset of Euclidean space, then the following statements are equivalent.*

- i.  $\succsim$  satisfies sensitivity, cancellation, continuity, monotonicity, sorting, and approachability.*
- ii.  $\succsim$  has a convergent and ordered strength  $\mathcal{E}$  weight representation  $v, V$ . Moreover,  $v$  is continuous and  $V$  is continuous in its first argument.*

### 5.3 Discrete outcomes

We now turn to situations where  $\mathbb{O}$  is either finite or countably infinite. One further axiom is required: A preference between two equally sized samples can be made sufficiently strong (by increasing sample size while keeping the relative empirical distributions in each sample intact) so as not to be reversed by a finite number of further outcome observations in both samples.

**A7 Archimedean:** For each pair  $Q, R \in \mathbb{S}$  of equal size with  $Q \succ R$  and each pair  $S, T \in \mathbb{S}$  of equal size, there is a positive integer  $n$  such that  $(n \otimes Q) \oplus S \succsim (n \otimes R) \oplus T$ .

This is a variant of the Archimedean axioms used to derive additive representations of weak orders (Roberts and Luce, 1968; Krantz et al., 1971). The axiom's formulation is similar to that used by Kothiyal et al. (2014) to characterize the average utility criterion.

**Theorem 5.** *If the set of outcomes  $\mathbb{O}$  is either finite or countably infinite, then the following statements are equivalent.*

- i.  $\succsim$  satisfies sensitivity, cancellation, and Archimedean.*
- ii.  $\succsim$  has a strength  $\mathcal{E}$  weight representation  $v, V$ .*

**Theorem 6.** *If the set of outcomes  $\mathbb{O}$  is either finite or countably infinite, then the first statement below implies the second one.*

- i.  $\succsim$  satisfies sensitivity, cancellation, Archimedean, monotonicity, sorting, and approachability.*
- ii.  $\succsim$  has a convergent strength  $\mathcal{E}$  weight representation  $v, V$ .*

Theorem 6 establishes that sensitivity, cancellation, Archimedean, monotonicity, sorting, and approachability are sufficient conditions for preferences to have a convergent strength & weight representation. Theorem 7 in the Appendix describes a weaker necessary ordering property for a strength & weight representation to satisfy these axioms.

## 6 Conclusion

We modelled how past outcome data informs decisions. Framing these decisions in terms of precise probabilities without reference to their derivation ignores uncertainty when the outcome generating processes are not observable. For example, consider the choice between two experimental treatments for a life-threatening disease: Each of the 20 subjects receiving the first treatment survived compared to only 960 out of the 1000 subjects receiving the second treatment. The 100% survival rate under the first treatment may be a fluke. The second treatment can even be credited with saving more lives than the first one. Knowing the actual survival probabilities would make the treatment choice trivial; it is not.

There are at least two further related reasons to avoid framing decisions in terms of probability distributions:

- i. Cognitive psychology research documents that people are hardwired to process and compare ensembles like samples ([Gallistel, 1990](#); [Gallistel and Gelman, 1992](#); [Butterworth, 2001, 2007](#); [Feigenson et al., 2004](#); [Dehaene, 2011](#)). Presenting information in a sample format, “natural frequencies,” improves judgement ([Gigerenzer and Hoffrage, 1995](#); [Hoffrage et al., 2000](#)); this contrasts with the systemic biases documented in probabilistic judgements (see [Tversky and Kahneman, 1971](#); [Kahneman and Tversky, 1972, 1973](#); [Tversky and Kahneman, 1974](#), for the seminal contributions). This suggests a model of choice that relies on samples as primitives.
- ii. When a data generating process is not observable, the distribution describing it is typically derived from observations gathered in sample form. Modelling decision-making as a choice over distributions then involves artificial primitives

and inconsistencies specific to these. For example, the “reduction of compound lotteries” is a typical assumption in modelling preferences over lotteries.<sup>7</sup> Expected utility theory yields inconsistent predictions when it is not met (Kahneman and Tversky, 1979). It may be possible to avoid inconsistencies specific to the probabilistic domain by presenting the available outcome information in sample form.

In our model, the agent judges the plausibility of the hypothesis that a given action will deliver her optimal outcome on the basis of the available sample data. She can distinguish between the strength of evidence supporting an action (the average utility of the outcomes in its sample) and the weight given to this evidence (sample size). She then combines the strength and weight of the evidence supporting the hypothesis, as suggested by the intuitive explanation of confidence formation provided by Griffin and Tversky (1992). There are strength & weight representations that are descriptive in the sense that they are consistent with the overconfidence in strength pattern documented by Griffin and Tversky (1992); there are others that are normative in that they combine the strength and weight of evidence consistently with a Bayesian benchmark (e.g., Example 1).

Future work may use other measures of strength: A pessimistic agent may assess strength by the minimum utility of the outcomes in a sample; an optimistic agent may use the maximum utility instead. More generally, the agent may be equipped with a rank-dependent measure of strength that can capture the average, the minimum, or the maximum utilities as special cases of one functional.<sup>8</sup> However, the average is a natural benchmark: From a descriptive viewpoint, psychologists docu-

---

<sup>7</sup>A compound lottery is a probabilistic mixture over probability distributions. The reduction of compound lotteries requires the agent is able to transform a compound lottery into a probability distribution over outcomes with the same overall outcome probabilities as the compound lottery; moreover, the agent ought to be indifferent between the resulting probability distribution and the compound lottery.

<sup>8</sup>The characterization of a model with these features requires technical topological or algebraic conditions on the domain of preferences and space of outcomes (Wakker, 1991, 1993). It also would rely on a weaker version of the cancellation axiom whereby the preferences between two samples of the same size are invariant to adding a common observation that is either better or worse than any outcome in the original samples.

ment that people have a natural ability to spontaneously use quite precise average approximations to summarize ensembles of objects of the same general kind and the same number of elements such as samples of circles with varying areas (Ariely, 2001; Chong and Treisman, 2003; Kahneman, 2003). From a normative viewpoint, the use of the average as measure of strength builds a bridge between sample evaluation as we proposed and expected utility theory under risk. Under reasonable independence assumptions, as samples grow, the empirical distribution function approaches the true distribution generating the outcomes in the sample. Our axiomatic analysis justifies that, as one approaches the true distributions, choice will be consistent with expected utility maximization.

**Extensions.** We modelled the determinants of choice when the only information describing an action is its sample of outcomes. Of course, decisions may be driven by more information, especially when the available samples are small. For example, in the case of medications, preferences for a drug with a particular chemical composition may reflect information or beliefs held by the agent.

To express this formally, consider a set of observable characteristics  $\mathbb{X}$  reflecting features of an action other than outcomes. For example,  $x \in \mathbb{X}$  could summarize the observable chemical composition of a drug, whether it is generic, and its brand; it could also specify a known prior distribution over the possible outcomes. Suppose that the agent accounts for these characteristics as well as outcomes in forming her preferences; this can be modelled by equipping the agent with a preference relation  $\succsim^*$  on  $\mathbb{X} \times \mathbb{S}$ . An action is identified by a pair  $(x, S) \in \mathbb{X} \times \mathbb{S}$  where  $x$  describes the action’s observable attributes and  $S$  its sample of past outcomes. In this paper, we focused on the case where  $x \in \mathbb{X}$  is fixed across actions. Of course, the observable characteristics may be important in the decision. However, as evidence accumulates in the form of large samples, the decision may again be driven by outcomes; the agent endorsing this view would ascribe to *asymptotic consequentialism*: For each pair  $x, y \in \mathbb{X}$  there is a positive integer  $n$  such that, for each  $S \in \mathbb{S}$  larger than  $n$ ,  $(x, S) \sim^* (y, S)$ .

Even more generally, a medication may have been used to treat different health problems, not just the problem faced by the agent, revealing its broader therapeutic

and side-effects. This information can then be harnessed across the different problems. Following the case-based approach pioneered by [Gilboa and Schmeidler \(1995, 2001\)](#), we can introduce a set of decision problems  $\mathbb{P}$ . The evidence supporting a decision now comes in the form of triples  $(p, x, S) \in \mathbb{P} \times \mathbb{X} \times \mathbb{S}$ . Here,  $(p, x, S)$  is the past case where problem  $p$  was tackled  $|S|$  times using an action with observable characteristics  $x$ , resulting in the outcomes in sample  $S$ . A memory  $\mathbb{M} \subseteq \mathbb{P} \times \mathbb{X} \times \mathbb{S}$  consists of all the cases the agent is aware of. An action with observable characteristics  $x$  for problem  $p$  is then identified with the case  $(p, x, S)$  and the agent forms preferences over  $\mathbb{P} \times \mathbb{X} \times \mathbb{S}$  given her memory of past cases.<sup>9</sup> To assess the relevance of these cases to problem  $p$  and an action with observable characteristics  $x$ , she relies on a similarity function  $\mathfrak{s} : [\mathbb{P} \times \mathbb{X}]^2 \rightarrow [0, 1]$  with values ranging from  $\mathfrak{s}((p, x), (q, y)) = 1$  if  $(q, y)$  is completely similar to  $(p, x)$  and  $\mathfrak{s}((p, x), (q, y)) = 0$  if it is completely different. A conceivable case-based decision model for this domain would represent the ranking of an action specified as  $(p, x, S)$  by

$$U(p, x, S) = \sum_{(q, y, T) \in \mathbb{M}} \mathfrak{s}((p, x), (q, y)) V(\mathbf{A}(v, T), |T|)$$

where  $v, V$  is a strength & weight representation of the form introduced in Section 4. An analysis of the above model is beyond the scope of this paper.

**Social welfare functions.** A long-standing problem in welfare economics is to determine the appropriateness of a utilitarian, an average utilitarian, or some other social welfare function when the population of interest varies in size; one of the most thorny issues here is the “repugnant conclusion” ([Parfit, 1976, 1982](#)). A strength & weight representation is formally similar to a social welfare function comparing welfare across varying populations ([Blackorby and Donaldson, 1984; Blackorby et al., 2005](#)).<sup>10</sup> This social welfare function has two arguments: Population size and the average utility across people in a social state. Individual utilities vary continuously and are observable. The analysis is normative, concentrating on ethical issues in

---

<sup>9</sup>[Gilboa and Schmeidler \(1995\)](#) use the term “act” referring to what we interpreted as observable characteristics.

<sup>10</sup>I thank Geir Asheim for pointing this out.

welfare evaluation. In our context, a sample can be viewed as the collection of outcomes experienced by the members of a population under a given policy, i.e., an action. The decision utility function used to evaluate strength in our model can describe how a social planner evaluates each of these outcomes; individual preferences are not observable. Our analysis however focuses on individual decision-making.

## References

- Ahn, D. S. (2008). Ambiguity without a state space. *The Review of Economic Studies* 75(1), 3–28.
- Ariely, D. (2001). Seeing sets: Representation by statistical properties. *Psychological Science* 12(2), 157–162.
- Barberà, S., W. Bossert, and P. K. Pattanaik (2004). Ranking sets of objects. In *Handbook of Utility Theory*, pp. 893–977. Springer.
- Blackorby, C., W. Bossert, and D. J. Donaldson (2005). *Population issues in social choice theory, welfare economics, and ethics*. Number 39. Cambridge University Press.
- Blackorby, C. and D. Donaldson (1984). Social criteria for evaluating population change. *Journal of Public Economics* 25(1-2), 13–33.
- Butterworth, B. (2001). Statistics: What seems natural? *Science* 292(5518), 853–855.
- Butterworth, B. (2007). Why frequencies are natural. *Behavioral and Brain Sciences* 30(3), 259–260.
- Chong, S. C. and A. Treisman (2003). Representation of statistical properties. *Vision research* 43(4), 393–404.
- Debreu, G. (1959). Topological methods in cardinal utility theory. Technical report, Cowles Foundation for Research in Economics, Yale University.
- Dehaene, S. (2011). *The number sense: How the mind creates mathematics*. Oxford University Press.
- Dekel, E. and B. L. Lipman (2012). Costly self-control and random self-indulgence. *Econometrica* 80(3), 1271–1302.

- Dekel, E., B. L. Lipman, and A. Rustichini (2001). Representing preferences with a unique subjective state space. *Econometrica* 69(4), 891–934.
- Dekel, E., B. L. Lipman, and A. Rustichini (2009). Temptation-driven preferences. *The Review of Economic Studies* 76(3), 937–971.
- Dekel, E., B. L. Lipman, A. Rustichini, and T. Sarver (2007). Representing preferences with a unique subjective state space: A corrigendum. *Econometrica* 75(2), 591–600.
- Feigenson, L., S. Dehaene, and E. Spelke (2004). Core systems of number. *Trends in Cognitive Sciences* 8(7), 307–314.
- Fishburn, P. C. (1972). Even-chance lotteries in social choice theory. *Theory and Decision* 3(1), 18–40.
- Fishburn, P. C. and A. Rubinstein (1982). Time preference. *International Economic Review*, 677–694.
- Gallistel, C. R. (1990). *The Organization of Learning*. MIT Press.
- Gallistel, C. R. and R. Gelman (1992). Preverbal and verbal counting and computation. *Cognition* 44(1-2), 43–74.
- Gigerenzer, G. and U. Hoffrage (1995). How to improve Bayesian reasoning without instruction: frequency formats. *Psychological Review* 102(4), 684.
- Gilboa, I. and D. Schmeidler (1995). Case-based decision theory. *The Quarterly Journal of Economics* 110(3), 605–639.
- Gilboa, I. and D. Schmeidler (1996). Case-based optimization. *Games and Economic Behavior* 15(1), 1–26.
- Gilboa, I. and D. Schmeidler (2001). *A theory of case-based decisions*. Cambridge University Press.
- Gilboa, I., D. Schmeidler, and P. P. Wakker (2002). Utility in case-based decision theory. *Journal of Economic Theory* 105(2), 483–502.
- Golman, R., D. Hagmann, and G. Loewenstein (2017). Information avoidance. *Journal of Economic Literature* 55(1), 96–135.
- Gravel, N., T. Marchand, and A. Sen (2012). Uniform expected utility criteria for decision making under ignorance or objective ambiguity. *Journal of Mathematical Psychology* 56(5), 297–315.
- Griffin, D. and A. Tversky (1992). The weighing of evidence and the determinants



- of confidence. *Cognitive psychology* 24(3), 411–435.
- Gul, F. and W. Pesendorfer (2001). Temptation and self-control. *Econometrica* 69(6), 1403–1435.
- Hertwig, R. and C. Engel (2016). Homo ignorans: Deliberately choosing not to know. *Perspectives on Psychological Science* 11(3), 359–372.
- Hoffrage, U., S. Lindsey, R. Hertwig, and G. Gigerenzer (2000). Communicating statistical information. *Science* 290(5500), 2261–2262.
- Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *American Economic Review* 93(5), 1449–1475.
- Kahneman, D. and A. Tversky (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology* 3(3), 430–454.
- Kahneman, D. and A. Tversky (1973). On the psychology of prediction. *Psychological Review* 80(4), 237.
- Kahneman, D. and A. Tversky (1979). Prospect theory: An analysis of decision under risk. *Econometrica* 47(2), 363–391.
- Knight, F. H. (1921). *Risk, uncertainty and profit*. Boston: Houghton Mifflin Company.
- Koopmans, T. C. (1949). Identification problems in economic model construction. *Econometrica*, 125–144.
- Kothiyal, A., V. Spinu, and P. P. Wakker (2014). Average utility maximization: a preference foundation. *Operations Research* 62(1), 207–218.
- Krantz, D., R. Luce, P. Suppes, and A. Tversky (1971). *Foundations of Measurement: Additive and Polynomial Representations*. Academic Press.
- Kreps, D. (1988). *Notes on the theory of choice*. Westview Press.
- Kreps, D. M. (1979). A representation theorem for “preference for flexibility”. *Econometrica*, 565–577.
- Manski, C. F. (2003). Partial identification of probability distributions: Springer series in statistics.
- Olszewski, W. (2007). Preferences over sets of lotteries. *The Review of Economic Studies* 74(2), 567–595.
- Parfit, D. (1976). On doing the best for our children. In M. D. Bayles (Ed.), *Ethics and Population*, Volume 100-102, pp. 239–253. Schenkman.

- Parfit, D. (1982). Future generations: Further problems. *Philosophy & Public Affairs*, 113–172.
- Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior and Organization* 3(4), 323–343.
- Roberts, F. S. and R. D. Luce (1968). Axiomatic thermodynamics and extensive measurement. *Synthese*, 311–326.
- Savage, L. (1954). *The Foundations of Statistics*. Dover Books on Mathematics. Dover Publications.
- Tversky, A. and D. Kahneman (1971). Belief in the law of small numbers. *Psychological Bulletin* 76(2), 105.
- Tversky, A. and D. Kahneman (1974). Judgment under uncertainty: Heuristics and biases. *Science* 185(4157), 1124–1131.
- Tversky, A. and D. Kahneman (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty* 5(4), 297–323.
- Von Neumann, J. and O. Morgenstern (1944). *Theory of Games and Economic Behavior*. Princeton University Press.
- Wakker, P. (1988). The algebraic versus the topological approach to additive representations. *Journal of Mathematical Psychology* 32(4), 421–435.
- Wakker, P. (1991). Additive representations on rank-ordered sets. i. the algebraic approach. *Journal of Mathematical Psychology* 35(4), 501–531.
- Wakker, P. (1993). Additive representations on rank-ordered sets: Ii. the topological approach. *Journal of Mathematical Economics* 22(1), 1–26.
- Yaari, M. E. (1987). The dual theory of choice under risk. *Econometrica*, 95–115.

## Online Appendix

### Preliminary observations

**Proposition 1.** *Suppose that  $\succsim$  satisfies cancellation. Then, for each pair  $S, T \in \mathbb{S}$  of equal size and each  $n \in \mathbb{N}$ ,  $S \succsim T$  if and only if  $n \otimes S \succsim n \otimes T$ .*

*Proof.* Let  $S, T \in \mathbb{S}$  be such that  $|S| = |T|$ . Suppose that  $S \succsim T$ . By repeatedly

applying cancellation,  $S \oplus S \succsim T \oplus S$  and  $S \oplus T \succsim T \oplus T$ . Since,  $T \oplus S = S \oplus T$ , by transitivity,  $S \oplus S \succsim T \oplus T$ . Equivalently,  $2 \otimes S \succsim 2 \otimes T$ . Continuing in this way,  $n \otimes S \succsim n \otimes T$ . Thus,  $S \succsim T$  implies  $n \otimes S \succsim n \otimes T$ . To prove the converse, suppose that  $n \otimes S \succsim n \otimes T$  but  $S \prec T$ . Repeating the above argument, with  $\prec$  instead of  $\succsim$ , we obtain  $n \otimes S \prec n \otimes T$ , a contradiction. Thus,  $S \succsim T$ , as desired.  $\square$

**Proposition 2.** *Sensitivity and cancellation together imply there are  $\alpha, \beta \in \mathbb{O}$  such that  $\alpha \succ \beta$ .*

*Proof.* By sensitivity, there are  $S, T \in \mathbb{S}$  such that  $|S| = |T|$  and  $S \succ T$ . Let  $n = |S|$ . Let the  $n$  outcomes in  $S$  be denoted by  $\alpha_1, \dots, \alpha_n$  so that  $\alpha_1 \succ \dots \succ \alpha_n$ . Let the  $n$  outcomes in  $T$  be denoted by  $\beta_1, \dots, \beta_n$  so that  $\beta_1 \succ \dots \succ \beta_n$ .

We first show that

$$n \otimes \alpha_1 \succsim S \quad \text{and} \quad T \succsim n \otimes \beta_n. \quad (6)$$

The proof of (6) is by induction: If  $n = 1$ , then  $S = \alpha_1 = n \otimes \alpha_1$ . Let  $h \in \mathbb{N}$  and assume that  $n \otimes \alpha_1 \succsim S$  is true whenever  $n < h$ . Let  $n = h$  and let  $S' \in \mathbb{S}$  consist of all of the outcomes in  $S$  except  $\alpha_h$ . By the inductive hypothesis,  $(h-1) \otimes \alpha_1 \succsim S'$ . By cancellation,  $[(h-1) \otimes \alpha_1] \oplus \alpha_1 \succsim S' \oplus \alpha_1$ . Equivalently,  $h \otimes \alpha_1 \succsim S' \oplus \alpha_1$ . Since  $\alpha_1 \succ \alpha_h$ , cancellation yields  $S' \oplus \alpha_1 \succ S' \oplus \alpha_h = S$ . Thus,  $h \otimes \alpha_1 \succsim S$ . By induction, for each  $n \in \mathbb{N}$ ,  $n \otimes \alpha_1 \succsim S$ . The fact that  $T \succsim n \otimes \beta_n$  is proven analogously.

By (6),  $n \otimes \alpha_1 \succsim S \succ T \succsim n \otimes \beta_n$ . Thus,  $n \otimes \alpha_1 \succ n \otimes \beta_n$ . By Proposition 1,  $\alpha_1 \succ \beta_n$ .  $\square$

## Proof of Theorem 1

**Proposition 3.** *Assume that  $\mathbb{O} = \{\alpha, \beta\}$  and that  $\succsim$  satisfies  $\alpha \succ \beta$  and cancellation. Then, for each pair  $S, T \in \mathbb{S}$  such that  $S$  and  $T$  are of equal size,  $S \succsim T$  if and only if  $S(\alpha) \geq T(\alpha)$ .*

*Proof.* Let  $S, T \in \mathbb{S}$  be of size  $n$ . Define  $s = S(\alpha)$  and  $t = T(\alpha)$ . If  $s = t$ , then  $S = T$  and there is nothing to prove. Assume instead that  $s \neq t$ .

Assume that  $s > t$ . By Proposition 1,  $(s - t) \otimes \alpha \succ (s - t) \otimes \beta$ . By cancellation,  $[(s - t) \otimes \alpha] \oplus [t \otimes \alpha] \succ [(s - t) \otimes \beta] \oplus [t \otimes \alpha]$ . Since

$$\begin{aligned} S &= [(s - t) \otimes \alpha] \oplus [t \otimes \alpha] \oplus [(n - s) \otimes \beta] \text{ and} \\ T &= [(s - t) \otimes \beta] \oplus [t \otimes \alpha] \oplus [(n - s) \otimes \beta], \end{aligned}$$

by cancellation,  $S \succ T$ .

Assume that  $S \succsim T$ . If  $s < t$  the argument in the preceding paragraph implies that  $S \prec T$ , a contradiction. Thus,  $s > t$ .  $\square$

*Proof of Theorem 1.* Let  $\mathbb{O} = \{\alpha, \beta\}$  and assume that  $\succsim$  satisfies sensitivity and cancellation. By Proposition 2, without loss of generality,  $\alpha \succ \beta$ . Let  $S, T$  denote two samples of size  $k$  and let  $n > k$ . Let  $S' = (n \otimes \alpha) \oplus S$  and  $T' = (n \otimes \alpha) \oplus T$ . Then,  $S'(\alpha) \geq n > k = T'(\alpha)$ . By Proposition 3,  $S' \succ T'$ . Thus,  $\succsim$  satisfies Archimedean. By Theorem 5,  $\succsim$  has a strength & weight representation  $v, V$ . Thus, the first statement in Theorem 1 implies the second one. We omit the straightforward proof of the converse. The equivalence of the second and third statements follows by letting  $v(\alpha) = 1$  and  $v(\beta) = 0$ .  $\square$

### Proof of Theorem 3

**Notation:** For each  $n \in \mathbb{N}$  and each  $x \in \mathbb{O}^n$ , let  $[x]$  denote the sample consisting of observations  $x_1, \dots, x_n$ . We can extend  $\succsim$  from  $\mathbb{S}$  to  $\cup_{n \in \mathbb{N}} \mathbb{O}^n$  defining an extended preference relation  $\succsim_e$  as follows: For each pair  $x, y \in \cup_{n \in \mathbb{N}} \mathbb{O}^n$ ,

$$x \succsim_e y \quad \text{if and only if} \quad [x] \succsim [y]. \quad (7)$$

For each  $n \in \mathbb{N}$ , let  $\mathbb{S}_n$  consist of all  $S \in \mathbb{S}$  such that  $|S| = n$ .

For each  $S \in \mathbb{S}$  and each  $\xi \in \mathbb{O}$ , the expression  $\xi \in S$  is taken to mean that  $S(\xi) \geq 1$ . For each  $S \in \mathbb{S}$  and each  $u : \mathbb{O} \rightarrow \mathbb{R}$ ,  $\sum_{\xi \in S} u(\xi)$  is shorthand for  $\sum_{\xi \in \mathbb{O}} S(\xi)u(\xi)$ .

For each  $\varepsilon \in \mathbb{R}_{++}$  and each  $\gamma \in \mathbb{O}$ , let  $B_\varepsilon(\gamma)$  denote the open ball of radius  $\varepsilon$  around  $\gamma$  in  $\mathbb{O}$ . Thus,  $B_\varepsilon(\gamma) = \{\eta \in \mathbb{O} : \|\gamma - \eta\| < \varepsilon\}$  where  $\|\cdot\|$  is the Euclidean norm.

**Proposition 4.** For each  $n \in \mathbb{N}$  and each pair  $A, B \in \mathbb{S}$  with  $|A| = |B| = n$ ,

- i.  $d(A, B) \geq 0$  and  $d(A, B) = 0$  if and only if  $A = B$ ,
- ii.  $d(A, B) = d(B, A)$ ,
- iii.  $d(A, B) \leq d(A, C) + d(C, B)$  for each  $C \in \mathbb{S}$  of size  $n$ .

*Proof.* Let  $A, B, C \in \mathbb{S}$  each contain  $n$  outcomes. From the definition of  $d$ , it is obvious that  $d(A, B) \geq 0$ ,  $d(A, B) = 0$  if and only if  $A = B$ , and that  $d(A, B) = d(B, A)$ . Then, for each  $\alpha \in A$ , each  $\beta \in B$ , and each  $\gamma \in C$ ,

$$\begin{aligned} d(C, A) + d(C, B) &\geq \|\gamma - \alpha\| + \|\gamma - \beta\| \\ &= \|\alpha - \gamma\| + \|\gamma - \beta\| \\ &\geq \|\alpha - \gamma + \gamma - \beta\| = \|\alpha - \beta\|. \end{aligned}$$

Taking the maximum over all  $\alpha \in A$  and  $\beta \in B$ ,  $d(C, A) + d(C, B) \geq d(A, B)$ .  $\square$

**Proposition 5.** Suppose that  $\succsim$  is continuous and that  $n \in \mathbb{N}$ ,  $S \in \mathbb{S}$ , and  $\alpha, \beta \in \mathbb{O}$  are such that  $n \otimes \alpha \succsim S \succsim n \otimes \beta$ . Then, there is  $\gamma \in \mathbb{O}$  such that  $n \otimes \gamma \sim S$ .

*Proof.* Suppose, by way of contradiction, that there is no  $\gamma \in \mathbb{O}$  such that  $n \otimes \gamma \sim S$ . Then, the sets  $A = \{\gamma \in \mathbb{O} : n \otimes \gamma \succ S\}$  and  $B = \{\gamma \in \mathbb{O} : n \otimes \gamma \prec S\}$  partition  $\mathbb{O}$  and are non-empty since  $\alpha \in A$  and  $\beta \in B$ .

Moreover,  $A$  and  $B$  are open: Let  $\gamma \in A$ . By continuity,  $C = \{T \in \mathbb{S} : T \succ S\}$  is open. Thus, since it contains  $n \otimes \gamma$ , there is  $\varepsilon > 0$  such that

$$D = \{T \in \mathbb{S} : d(n \otimes \eta, T) < \varepsilon\} \subseteq C.$$

Then, for each  $\eta \in \mathbb{O}$  such that  $\|\gamma - \eta\| < \varepsilon$ ,  $d(n \otimes \gamma, n \otimes \eta) = \|\gamma - \eta\| < \varepsilon$ . Thus,  $n \otimes \eta \in D \subseteq C$ . Thus, for each  $\eta \in \mathbb{O}$  such that  $\|\gamma - \eta\| < \varepsilon$ ,  $\eta \in A$ . Thus,  $\gamma$  is in the interior of  $A$ . Thus, every element of  $A$  is in its interior. Thus,  $A$  is open. Similarly,  $B$  is open.

Thus,  $A$  and  $B$  are open, non-empty, and partition a connected set  $\mathbb{O}$ . This contradiction establishes that, in fact, there is  $\gamma \in \mathbb{O}$  such that  $n \otimes \gamma \sim S$ .  $\square$

**Proposition 6.** Suppose that  $\succsim$  satisfies cancellation and continuity. Then, for each  $S \in \mathbb{S}$ , there is  $\alpha \in \mathbb{O}$  such that  $|S| \otimes \alpha \sim S$ .

*Proof.* Suppose that  $\succsim$  satisfies continuity and cancellation. Let  $S \in \mathbb{S}$ ,  $n = |S|$ , and label the  $n$  outcomes in  $S$  by  $\alpha_1, \alpha_2, \dots, \alpha_n$  so that  $\alpha_1 \succsim \alpha_2 \succsim \dots \succsim \alpha_n$ . By the same induction argument used to establish (6),  $n \otimes \alpha_1 \succsim S \succsim n \otimes \alpha_n$ . By Proposition 5, this implies that there is  $\alpha \in \mathbb{O}$  such that  $n \otimes \alpha \sim S$ .  $\square$

**Proposition 7.** *Suppose that  $\succsim$  is continuous. Then, for each  $n \in \mathbb{N}$ , the preference relation  $\succsim_e$  on  $\mathbb{O}^n$  is continuous: For each  $x \in \mathbb{O}^n$ ,*

$$\{y \in \mathbb{O}^n : y \succ_e x\} \text{ and } \{y \in \mathbb{O}^n : x \succ_e y\}$$

*are open with respect to the product topology on  $\mathbb{O}^n$  where  $\mathbb{O}$  is endowed with the Euclidean topology.*

*Proof.* Let  $n \in \mathbb{N}$  and  $x \in \mathbb{O}^n$ . We will show that  $\{y \in \mathbb{O}^n : y \succ_e x\}$  is open with respect to the product topology on  $\mathbb{O}^n$  where  $\mathbb{O}$  is endowed with the Euclidean topology. Let  $z \in \{y \in \mathbb{O}^n : y \succ_e x\}$ . Thus,  $z \succ_e x$  and, by (7),  $[z] \succ [x]$ . By continuity,  $\{Y \in \mathbb{S} : Y \succ [x]\}$  is open so there is  $\varepsilon > 0$  such that

$$\{Y \in \mathbb{S} : d([z], Y) < \varepsilon\} \subseteq \{Y \in \mathbb{S} : Y \succ [x]\}.$$

Let  $y \in \mathbb{O}^n$  be such that  $\sum_{i=1}^n \|z_i - y_i\| < \varepsilon$ . Then,  $d([z], [y]) \leq \sum_{i=1}^n \|z_i - y_i\| < \varepsilon$ . Thus,  $[y] \succ [x]$  and, by (7),  $y \succ_e x$ . Thus,  $\{y \in \mathbb{O}^n : \sum_{i=1}^n \|z_i - y_i\| < \varepsilon\}$  is contained in  $\{y \in \mathbb{O}^n : y \succ_e x\}$ . Thus, if  $z \succ_e x$ , then  $z$  is in the interior of  $\{y \in \mathbb{O}^n : y \succ_e x\}$ . Thus,  $\{y \in \mathbb{O}^n : y \succ_e x\}$  is open. The proof that  $\{y \in \mathbb{O}^n : x \succ_e y\}$  is open is analogous.  $\square$

**Proposition 8.** *Suppose that  $\succsim$  satisfies sensitivity, continuity, and cancellation. Then, for each  $n \in \mathbb{N}$  such that  $n \geq 3$ , there is a continuous non-constant function  $u^n : \mathbb{O} \rightarrow \mathbb{R}$  such that, for each pair  $S, T \in \mathbb{S}_n$ ,  $S \succsim T$  if and only if  $\sum_{\xi \in S} u^n(\xi) \geq \sum_{\xi \in T} u^n(\xi)$ . Moreover,  $u^n$  is unique up to a positive affine transformation.*

*Proof.* It is straightforward to verify that  $\succsim$  satisfies cancellation, continuity, comparability, and sensitivity if  $\succsim$  can be represented as specified in the Proposition 8. The affine uniqueness of the representation is also straightforward to verify.

Conversely, suppose that  $\succsim$  satisfies sensitivity, continuity, and cancellation. Let  $n \in \mathbb{N}$  and  $N = \{1, \dots, n\}$ . By Proposition 2, there are  $\alpha, \beta \in \mathbb{O}$  such that  $\alpha \succ \beta$ .

Let  $k \in N$  and  $x, y \in \mathbb{O}^n$  be such that  $x_k = \alpha$ ,  $y_k = \beta$ , and, for each  $h \in N \setminus \{k\}$ ,  $x_h = y_h$ . By cancellation,  $[x] \succ [y]$ . Thus, the extended preference relation  $\succsim_e$  satisfies the following property:

$$\forall x \in \mathbb{O}^n, \forall k = 1, \dots, n, \exists x'_k \in \mathbb{O} \quad \text{s.t.} \quad (x'_k, x_{-k}) \succ_e x \quad \text{or} \quad x \succ_e (x'_k, x_{-k}). \quad (8)$$

Let  $x, y, z, w \in \mathbb{O}^n$  and  $I \subseteq N$ . By cancellation,  $[x_I] \oplus [z_{N \setminus I}] \succsim [y_I] \oplus [z_{N \setminus I}]$  if and only if  $[x_I] \succsim [y_I]$ . Similarly,  $[x_I] \oplus [w_{N \setminus I}] \succsim [y_I] \oplus [w_{N \setminus I}]$  if and only if  $[x_I] \succsim [y_I]$ . Thus,

$$\begin{aligned} (x_I, z_{-I}) \succsim_e (y_I, z_{-I}) &\Leftrightarrow [x_I] \oplus [z_{N \setminus I}] \succsim [y_I] \oplus [z_{N \setminus I}] \\ &\Leftrightarrow [x_I] \succsim [y_I] \\ &\Leftrightarrow [x_I] \oplus [w_{N \setminus I}] \succsim [y_I] \oplus [w_{N \setminus I}] \\ &\Leftrightarrow (x_I, w_{-I}) \succsim_e (y_I, w_{-I}). \end{aligned}$$

Thus,

$$\forall x, y, z, w \in \mathbb{O}^n, \forall I \subseteq N, \quad (x_I, z_{-I}) \succsim_e (y_I, z_{-I}) \Leftrightarrow (x_I, w_{-I}) \succsim_e (y_I, w_{-I}). \quad (9)$$

Then, since  $\succsim_e$  is continuous on  $\mathbb{O}^n$  (Proposition 7) and  $n \geq 3$ , by Debreu's characterization of separable utility functions (Debreu, 1959; Wakker, 1988), conditions (8) and (9), ensure that,

$$\forall x, y \in \mathbb{O}^n, \quad x \succsim_e y \Leftrightarrow \sum_{i=1}^n u_i^n(x_i) \geq \sum_{i=1}^n u_i^n(y_i) \quad (10)$$

where  $u_i^n : \mathbb{O} \rightarrow \mathbb{R}$  is continuous and unique up to an affine transformation so that there are scalars  $a > 0$  and  $b_i$  such that any alternative additive representation would hold if and only if  $v_i^n = au_i^n + b_i$ .

It remains to show that we can drop the subscript in  $u_i^n$  to define  $u^n$ . By way of contradiction, suppose that there are  $i, j \in N$  and  $\alpha, \beta \in \mathbb{O}$  such that  $u_i^n(\alpha) + u_j^n(\beta) \neq u_i^n(\beta) + u_j^n(\alpha)$ . Without loss of generality, suppose that  $u_i^n(\alpha) + u_j^n(\beta) > u_i^n(\beta) + u_j^n(\alpha)$ . Let  $x, y \in \mathbb{O}^2$  be such that  $x_i = \alpha$ ,  $x_j = \beta$  and  $y_i = \beta$ ,  $y_j = \alpha$ , respectively. By (10),  $x \succ_e y$ . However,  $[x] = [y]$  implies  $[x] \sim [y]$ . This contradicts the definition of  $\succsim_e$  as an extension of  $\succsim$ . Thus,

$$\forall i, j \in N, \forall \alpha, \beta \in \mathbb{O}, \quad u_i^n(\alpha) + u_j^n(\beta) = u_i^n(\beta) + u_j^n(\alpha).$$

Thus, for each pair  $i, j \in N$  and each pair  $\alpha, \beta \in \mathbb{O}$ ,  $u_i^n(\alpha) - u_j^n(\alpha) = u_i^n(\beta) - u_j^n(\beta)$ . Thus,  $u_i - u_j$  is constant. Since each of the  $u_i$  functions in (10) is unique up to an affine transformations,  $u_i^n - u_j^n$  can be taken to be zero. Thus, given  $i \in N$ , we can define  $u^n = u_i^n$ .

Let  $S, T \in \mathbb{S}_n$  and let  $x, y \in \mathbb{O}^n$  be such that  $[x] = S$  and  $[y] = T$ . By (7),  $S \succsim T$  if and only if  $x \succsim_e y$ . By (10) and the observations in the preceding paragraphs,  $x \succsim_e y$  if and only if  $\sum_{i=1}^n u^n(x_i) \geq \sum_{i=1}^n u^n(y_i)$  where  $u^n$  is continuous, represents  $R$ , and is unique up to an affine transformation. Finally,  $\sum_{i=1}^n u^n(x_i) \geq \sum_{i=1}^n u^n(y_i)$  is equivalent to  $\sum_{\xi \in S} u^n(\xi) \geq \sum_{\xi \in T} u^n(\xi)$ . Thus,  $S \succsim T$  if and only if  $\sum_{\xi \in S} u^n(\xi) \geq \sum_{\xi \in T} u^n(\xi)$ .  $\square$

**Proposition 9.** *Suppose that  $\succsim$  satisfies sensitivity, continuity, and cancellation. Then, there is a continuous function  $v : \mathbb{O} \rightarrow \mathbb{R}$  such that, for each pair  $S, T \in \mathbb{S}$  of equal size,  $S \succsim T$  if and only if  $\sum_{\xi \in S} v(\xi) \geq \sum_{\xi \in T} v(\xi)$ . Moreover,  $v$  is unique up to a positive affine transformation.*

*Proof.* It is straightforward to verify that  $\succsim$  satisfies sensitivity, continuity, and cancellation if  $\succsim$  can be represented as specified in Proposition 9.

Conversely, suppose that  $\succsim$  satisfies sensitivity, cancellation, and continuity. Let  $n \in \mathbb{N}$  be such that  $n \geq 3$ . By Proposition 8, there is a continuous  $u^n : \mathbb{O} \rightarrow \mathbb{R}$  such that,

$$\text{for each pair } S, T \in \mathbb{S}_n, S \succsim T \Leftrightarrow \sum_{\xi \in S} u^n(\xi) \geq \sum_{\xi \in T} u^n(\xi). \quad (11)$$

Let  $k \in \mathbb{N}$  be such that  $k \neq n$ . By Proposition 8, there is a continuous  $u^{kn} : \mathbb{O} \rightarrow \mathbb{R}$  such that,

$$\text{for each pair } S, T \in \mathbb{S}_{kn}, S \succsim T \Leftrightarrow \sum_{\xi \in S} u^{kn}(\xi) \geq \sum_{\xi \in T} u^{kn}(\xi). \quad (12)$$

By Proposition 1,

$$\text{for each pair } S, T \in \mathbb{S}_n, S \succsim T \Leftrightarrow k \otimes S \succsim k \otimes T. \quad (13)$$

Combining (11), (12), and (13), for each pair  $S, T \in \mathbb{S}_n$ ,

$$\sum_{\xi \in S} u^n(\xi) \geq \sum_{\xi \in T} u^n(\xi) \Leftrightarrow \sum_{\xi \in k \otimes S} u^{kn}(\xi) \geq \sum_{\xi \in k \otimes T} u^{kn}(\xi) \Leftrightarrow k \sum_{\xi \in S} u^{kn}(\xi) \geq k \sum_{\xi \in T} u^{kn}(\xi).$$



Since  $u^n$  is unique up to an affine transformation, it follows that  $u^{kn}$  is an affine transformation of  $u^n$ . Thus,

$$\text{for each pair } S, T \in \mathbb{S}_{kn}, S \succsim T \Leftrightarrow \sum_{\xi \in S} u^n(\xi) \geq \sum_{\xi \in T} u^n(\xi). \quad (14)$$

By Proposition 8, there is a continuous  $u^k : \mathbb{O} \rightarrow \mathbb{R}$  such that,

$$\text{for each pair } S, T \in \mathbb{S}_k, S \succsim T \Leftrightarrow \sum_{\xi \in S} u^k(\xi) \geq \sum_{\xi \in T} u^k(\xi). \quad (15)$$

By Proposition 1,

$$\text{for each pair } S, T \in \mathbb{S}_k, S \succsim T \Leftrightarrow n \otimes S \succsim n \otimes T. \quad (16)$$

Combining (14), (15), and (16), for each pair  $S, T \in \mathbb{S}_k$ ,

$$\sum_{\xi \in S} u^k(\xi) \geq \sum_{\xi \in T} u^k(\xi) \Leftrightarrow \sum_{\xi \in n \otimes S} u^n(\xi) \geq \sum_{\xi \in n \otimes T} u^n(\xi) \Leftrightarrow n \sum_{\xi \in S} u^n(\xi) \geq n \sum_{\xi \in T} u^n(\xi).$$

Since  $u^k$  is unique up to an affine transformation, it follows that  $u^n$  is an affine transformation of  $u^k$ . Thus,

$$\text{for each } k \in \mathbb{N} \text{ and each pair } S, T \in \mathbb{S}_k, S \succsim T \Leftrightarrow \sum_{\xi \in S} u^n(\xi) \geq \sum_{\xi \in T} u^n(\xi).$$

To complete the proof, let  $v = u^n$ . By Proposition 8,  $v$  is unique up to a positive affine transformation.  $\square$

**Proposition 10.** *Let  $I$  denote an interval in the real line, let  $N$  denote a finite set, and let  $\succsim^N$  denote a complete and transitive binary relation on  $I \times N$  such that,*

$$\begin{aligned} &\text{if } x, y, z \in I \text{ and } k, n \in N \text{ satisfy } (x, n) \succsim^N (y, k) \succsim^N (z, n), \\ &\text{then there is } y' \in I \text{ such that } (y', n) \sim^N (y, k), \end{aligned} \quad (17)$$

and

$$\text{for each } k \in N \text{ and each pair } x, y \in I, (x, k) \succsim^N (y, k) \Leftrightarrow x \geq y. \quad (18)$$

Then, there is a function  $W : I \times N \rightarrow \mathbb{R}$  that represents  $\succsim^N$  and is strictly increasing and continuous in its first argument.

*Proof.* Let all of the notation and assumptions be as in the statement. The proof is by induction on the cardinality of  $N$ .

Induction basis: Suppose that  $N$  has a cardinality of one. Let  $W : I \times N \rightarrow \mathbb{R}$  be bounded, and strictly increasing and continuous in its first argument. For each pair  $(x, n), (y, k) \in I \times N$ ,  $n = k$  since  $N$  contains a single element, and thus, by (18),  $(x, n) \succsim^N (y, k)$  if and only if  $x \geq y$ ; moreover, since  $W$  is strictly increasing in its first argument,  $x \geq y$  if and only if  $W(x, n) \geq W(y, k)$ . Thus,  $W$  represents  $\succsim^N$ , is bounded, and is strictly increasing and continuous in its first argument.

Induction hypothesis: Let  $m \in \mathbb{N}$  and suppose that, for each  $N$  such that  $|N| < m$ , there is a function  $\tilde{W} : I \times N \rightarrow \mathbb{R}$  that represents  $\succsim^N$ , is bounded and strictly increasing and continuous in its first argument.

Induction: Suppose that  $N$  has cardinality  $m$ . Fix an  $n \in N$  such that,

$$\text{for each } k \in N \setminus \{n\} \text{ and each } x \in I, \text{ there is } y \in I \text{ with } (x, n) \succsim^N (y, k). \quad (19)$$

By the induction hypothesis, there is a function  $\tilde{W} : I \times [N \setminus \{n\}] \rightarrow \mathbb{R}$  that represents the restriction of  $\succsim^N$  to  $I \times [N \setminus \{n\}]$ , that is bounded, and that is strictly increasing and continuous in its first argument. By (19),  $I$  can be partitioned into

$$\begin{aligned} A &= \{x \in I : \exists (y, k) \in I \times [N \setminus \{n\}], (x, n) \sim^N (y, k)\}, \text{ and} \\ B &= \{x \in I : \forall (y, k) \in I \times [N \setminus \{n\}], (x, n) \succ^N (y, k)\}. \end{aligned}$$

Then,<sup>11</sup>

$$A \text{ and } B \text{ are intervals such that } \inf A = \inf I, \sup A = \inf B, \sup B = \sup I. \quad (20)$$

We consider the two possible cases that follow, showing that we can always define an appropriate representation of  $\succsim^N$  on  $I \times N$ .

---

<sup>11</sup> Let  $x \in B$  and  $z \in I$  be such that  $z > x$ . Then,  $x \in B$  implies  $(x, n) \succ^N (y, k)$  for each  $(y, k) \in I \times [N \setminus \{n\}]$  and  $z > x$  implies, by (18),  $(z, n) \succ^N (x, n)$ . Combining these,  $(z, n) \succ^N (y, k)$  for each  $(y, k) \in I \times [N \setminus \{n\}]$ . Thus,  $z \in B$ . Thus,  $B$  is an interval and there is no  $x' \in A$  such that  $x' > x$ . Thus,  $A$  is an interval as well and  $\sup A = \inf B$ .

Case 1:  $A$  is empty. Let  $f : I \rightarrow \mathbb{R}$  be bounded, strictly increasing, continuous, and such that, for each  $(y, k) \in I \times [N \setminus \{n\}]$ ,  $f(x) > \tilde{W}(y, k)$ . Define  $W : I \times N \rightarrow \mathbb{R}$  by

$$W(x, k) = \begin{cases} \tilde{W}(x, k) & \text{if } x \in I \text{ and } k \neq n \\ f(x) & \text{if } x \in I \text{ and } k = n \end{cases} \quad \text{for each } (x, k) \in I \times N.$$

Then, by construction,  $W$  represents  $\succsim^N$ , is bounded, and is strictly increasing and continuous in its first argument.

Case 2:  $A$  is not empty. Fix an  $i \in N \setminus \{n\}$  such that, for each  $x \in A$ , there is a  $y \in I$  such that  $(y, i) \succsim^N (x, n)$ .<sup>12</sup> By (19), for each  $x \in I$ , there is  $y' \in I$  such that  $(x, n) \succsim^N (y', i)$ . Altogether,

$$\text{for each } x \in A, \text{ there are } y, y' \in I \text{ such that } (y, i) \succsim^N (x, n) \succsim^N (y', i). \quad (21)$$

Thus, by (17), for each  $x \in A$ , there is an  $x' \in I$  such that  $(x', i) \sim^N (x, n)$ ; such an  $x'$  is unique by (18). Thus, define  $g : A \rightarrow I$  by [for each  $x \in A$ ,  $(g(x), i) \sim^N (x, n)$ ]. Define  $h : B \rightarrow \mathbb{R}$  to be bounded, strictly increasing, continuous, and such that

$$\inf\{h(x) : x \in B\} = \sup\{\tilde{W}(g(x), i) : x \in A\}. \quad (22)$$

Define  $W : I \times N \rightarrow \mathbb{R}$  by

$$W(x, k) = \begin{cases} \tilde{W}(x, k) & \text{if } x \in I \text{ and } k \neq n \\ \tilde{W}(g(x), i) & \text{if } x \in A \text{ and } k = n \\ h(x) & \text{if } x \in B \text{ and } k = n \end{cases} \quad \text{for each } (x, k) \in I \times N.$$

The demonstration that  $W$  represents  $\succsim^N$  and is strictly increasing and continuous in its first argument follows in four successive steps.

Step 1.  $g$  is strictly increasing: Let  $x, y \in A$  be such that  $x < y$ . Then, by (18),  $(y, n) \succ^N (x, n)$ . Thus,  $(g(x), i) \sim^N (x, n)$  and  $(g(y), i) \sim^N (y, n)$  implies  $(g(y), i) \succ^N (g(x), i)$ . By (18),  $g(x) < g(y)$ , confirming that  $g$  is strictly increasing.

Step 2.  $g$  is continuous. Let  $x, y \in A$  be such that  $x < y$ . Since  $g$  is strictly increasing,  $g(x) < g(y)$ . Let  $z' \in (g(x), g(y))$ . By (18),  $(g(y), i) \succ^N (z', i) \succ^N$

---

<sup>12</sup>Such an  $i$  exists for otherwise  $x \in B$ . However,  $B$  is disjoint from  $A$ .

$(g(x), i)$ . By the definition of  $g$ ,  $(y, n) \sim^N (g(y), i)$  and  $(x, n) \sim^N (g(x), i)$ . Thus,  $(y, n) \succ^N (z', i) \succ^N (x, n)$ . By (17), there is  $z \in I$  such that  $(z, n) \sim^N (z', i)$ . Thus,  $(y, n) \succ^N (z, n) \succ^N (x, n)$ . By (18),  $z \in (x, y)$ . By the definition of  $g$ ,  $g(z) = z'$ . Altogether,

$$\begin{aligned} & \text{for each pair } x, y \in A \text{ such that } x < y \text{ and each } z' \in (g(x), g(y)), \\ & \text{there is } z \in (x, y) \text{ such that } g(z) = z'. \end{aligned} \quad (23)$$

Since  $g$  is increasing, (23) implies that  $g$  is continuous.<sup>13</sup>

Step 3.  $W$  is strictly increasing and continuous in its first argument: For each  $k \in N \setminus \{n\}$ ,  $\tilde{W}(\cdot, k)$  is strictly increasing and continuous. Thus,  $W(\cdot, k)$  is strictly increasing and continuous. Since  $i \in N \setminus \{n\}$  and  $g$  is strictly increasing and continuous on  $A$ ,  $W(\cdot, n) = \tilde{W}(g(\cdot), i)$  is strictly increasing and continuous on  $A$ . Since  $h$  is strictly increasing and continuous on  $B$ ,  $W(\cdot, n)$  is strictly increasing and continuous on  $B$ . By (20) and (22),  $W(\cdot, n)$  is then continuous on  $I$ . In particular, if either  $A$  or  $B$  are empty,  $W(\cdot, n)$  is strictly increasing and continuous.

It remains to show  $W(\cdot, n)$  is strictly increasing when  $A$  and  $B$  are not empty. Let  $s = \sup I$  and  $a = \sup A$ . Recall that  $\inf B = a$  by (20).

- If  $s \in I$ , then  $a \in A$ .<sup>14</sup> Since  $A$  and  $B$  are disjoint,  $a \notin B$ . Then, for each pair  $x, y \in I$  such that  $x < a < y$ ,

$$W(x, n) < W(a, n) = \tilde{W}(g(a), i) < h(y) = W(y, n).$$

- If  $s \notin I$ , then  $a \in B$ .<sup>15</sup> Since  $A$  and  $B$  are disjoint,  $a \notin A$ . Then, for each pair  $x, y \in I$  such that  $x < a < y$ ,

$$W(x, n) = \tilde{W}(g(x), i) < h(a) = W(a, n) < W(y, n).$$

---

<sup>13</sup>By way of contradiction, suppose that  $g$  is discontinuous at a certain  $y \in A$ . Since  $g$  is increasing, either  $\lim_{w \uparrow y} g(w) < g(y)$  or  $g(y) < \lim_{w \downarrow y} g(w)$ . Assume, for example, the former and let  $x < y$ . Since  $g$  is increasing,  $g(x) \leq \lim_{w \uparrow y} g(w) < g(y)$ . If  $z' \in (\lim_{w \uparrow y} g(w), g(y)) \subseteq (g(x), g(y))$ , then there is no  $z \in (x, y)$  such that  $g(z) = z'$ . This contradicts (23). Thus,  $g$  is continuous.

<sup>14</sup>By (21), for each  $x \in A$ , there is  $y \in I$  such that  $(y, i) \succ^N (x, n)$ . By (18),  $s \geq y$  implies  $(s, i) \succ^N (y, i)$ . By definition, for each  $z \in B$ ,  $(z, n) \succ^N (s, i)$ . Thus, for each  $x \in A$  and each  $z \in B$ ,  $(z, n) \succ^N (s, i) \succ^N (x, n)$ . By (17), there is  $s' \in I$  such that  $(s', n) \sim^N (s, i)$ . Thus, for each  $x \in A$ ,  $(s', n) \succ^N (x, n)$ . By (18), for each  $x \in A$ ,  $s' \geq x$  so  $s'$  is an upper bound of  $A$ . Since  $i \neq n$ ,  $(s', n) \sim^N (s, i)$  implies  $s' \in A$ . Thus,  $s' = a$ .

<sup>15</sup>By way of contradiction, suppose that  $a \in A$ . Then, there is  $(x, k) \in I \times [N \setminus \{n\}]$  such that

Thus,  $W(\cdot, n)$  is strictly increasing.

Step 4.  $W$  represents  $\succsim^N$ : Since  $W$  coincides with  $\tilde{W}$  on  $I \times [N \setminus \{n\}]$ ,  $W$  represents the restriction of  $\succsim^N$  to  $I \times [N \setminus \{n\}]$ . It remains to show that, for each pair  $x, y \in I$  and each  $k \in N \setminus \{n\}$ ,

$$(x, k) \succsim^N (y, n) \Leftrightarrow W(x, k) \geq W(y, n), \quad (24)$$

$$(x, n) \succsim^N (y, k) \Leftrightarrow W(x, n) \geq W(y, k). \quad (25)$$

To establish (24), note that either  $(x, k) \succsim^N (y, n)$  or  $W(x, k) \geq W(y, n)$  imply  $y \notin B$ . Thus,  $y \in A$ . If  $y \in A$ , then (24) follows from

$$(x, k) \succsim^N (y, n) \Leftrightarrow (x, k) \succsim^N (g(y), i) \Leftrightarrow \tilde{W}(x, k) \geq \tilde{W}(g(y), i) \Leftrightarrow W(x, k) \geq W(y, n).$$

If  $x \in A$ , then (25) is established by

$$(x, n) \succsim^N (y, k) \Leftrightarrow (g(x), i) \succsim^N (y, k) \Leftrightarrow \tilde{W}(g(x), i) \geq \tilde{W}(y, k) \Leftrightarrow W(x, n) \geq W(y, k).$$

If  $x \in B$ , then (25) is established as follows: Note that there is a  $z \in A$  such that  $(z, n) \succsim^N (y, k)$ .<sup>16</sup> By (20),  $z \in A$  and  $x \in B$  imply  $x \geq z$ .

- Suppose that  $(x, n) \succsim^N (y, k)$ . Since  $z \in A$ ,  $(g(z), i) \sim^N (z, n)$ . Thus,  $(g(z), i) \sim^N (z, n) \succsim^N (y, k)$  implies  $\tilde{W}(g(z), i) \geq \tilde{W}(y, k)$ . By definition,  $W(z, n) = \tilde{W}(g(z), i)$  and  $W(y, k) = \tilde{W}(y, k)$ . Thus,  $W(z, n) \geq W(y, k)$ . Since  $W(\cdot, n)$  is increasing,  $W(x, n) \geq W(z, n)$ . Thus,  $W(x, n) \geq W(y, k)$ .
- Suppose that  $W(x, n) \geq W(y, k)$ . By (18),  $x \geq z$  implies  $(x, n) \succsim^N (z, n)$ . Since  $(z, n) \succsim^N (y, k)$ ,  $(x, n) \succsim^N (y, k)$ . □

---

$(x, k) \sim^N (a, n)$ . Since  $x < s$ , there is  $y \in I$  be such that  $x < y < s$ . Then, by (18),  $(y, k) \succ^N (x, k)$ . By the definition,  $z \in B$  implies  $(z, n) \succ^N (y, k)$ . Altogether,  $(z, n) \succ^N (y, k) \succ^N (a, n)$ . By (17), there is  $y' \in I$  such that  $(y', n) \sim^N (y, k)$ . Thus,  $(y', n) \succ^N (a, n)$ . By (18),  $y' > a$ . However,  $(y', n) \sim^N (y, k)$  implies  $y' \in A$ . This is a contradiction since  $a$  is the supremum of  $A$ .

<sup>16</sup>By way of contradiction, suppose that, for each  $z \in A$ ,  $(y, k) \succ^N (z, n)$ . Let  $b \in B$ . By definition,  $(b, n) \succ^N (y, k)$ . Thus, for each  $z \in A$ ,  $(b, n) \succ^N (y, k) \succ^N (z, n)$ . By (17), there is  $y' \in I$  such that  $(y', n) \sim^N (y, k)$ . Thus, for each  $z \in A$ ,  $(y', n) \succ^N (z, n)$ . Thus,  $y' \notin A$ . However,  $(y', n) \sim^N (y, k)$  implies that  $y' \notin B$ . This is a contradiction since  $A$  and  $B$  partition  $I$ .

**Proposition 11.** *Let  $I$  denote an interval in the real line and let  $\succsim^*$  denote a complete and transitive binary relation on  $I \times \mathbb{N}$  such that,*

$$\begin{aligned} & \text{if } x, y, z \in I \text{ and } k, n \in \mathbb{N} \text{ satisfy } (x, n) \succsim^* (y, k) \succsim^* (z, n), \\ & \text{then there is } y' \in I \text{ such that } (y', n) \sim^* (y, k), \end{aligned} \quad (26)$$

and

$$\text{for each } k \in \mathbb{N} \text{ and each pair } x, y \in I, (x, k) \succsim^* (y, k) \Leftrightarrow x \geq y. \quad (27)$$

Then, there is a function  $W : I \times \mathbb{N} \rightarrow \mathbb{R}$  that represents  $\succsim^*$  and is strictly increasing and continuous in its first argument.

*Proof.* Let all the notation and assumptions be as in the statement. By (26), (27), and Proposition 10, for each  $n \in \mathbb{N}$ , the restriction of  $\succsim^*$  to  $I \times \{1, \dots, n\}$  can be represented by a function  $W^n : I \times \{1, \dots, n\} \rightarrow \mathbb{R}$  that is strictly increasing and continuous in its first argument. Moreover, for each  $n \in \mathbb{N}$ , both  $W^n$  and  $W^{n+1}$  represent the restriction of  $\succsim^*$  to  $I \times \{1, \dots, n\}$ . Thus, we can further specify that,

$$\text{for each } n \in \mathbb{N} \text{ and each } (x, k) \in I \times \{1, \dots, n\}, \quad W^n(x, k) = W^{n+1}(x, k). \quad (28)$$

Taking condition (28) as given, define  $W : I \times \mathbb{N} \rightarrow \mathbb{R}$ , for each  $(x, n) \in I \times \mathbb{N}$ , by  $W(x, n) = W^n(x, n)$ . Since each  $W^n(\cdot, n)$  is strictly increasing and continuous,  $W$  is strictly increasing and continuous in its first argument.

It remains to show that  $W$  represents  $\succsim^*$ . Let  $(x, n), (y, k) \in I \times \mathbb{N}$ . Then,  $W(x, n) = W^n(x, n)$  and  $W(y, k) = W^k(y, k)$ . Without loss of generality,  $n \leq k$ . By (28),  $W^n(x, n) = W^{n+1}(x, n) = \dots = W^k(x, n)$ . Thus,  $W(x, n) \geq W(y, k)$  if and only if  $W^k(x, n) \geq W^k(y, k)$ . Since  $W^k$  represents  $\succsim^*$  on  $I \times \{1, \dots, k\}$ ,  $W^k(x, n) \geq W^k(y, k)$  if and only if  $(x, n) \succsim^* (y, k)$ . Thus,  $W(x, n) \geq W(y, k)$  if and only if  $(x, n) \succsim^* (y, k)$ . Thus,  $W$  represents  $\succsim^*$ .  $\square$

*Proof of Theorem 3.* We start showing that the first statement in Theorem 3 implies the second one. Let  $\succsim$  satisfy sensitivity, continuity, and cancellation. By Proposition 9, there is a continuous function  $v : \mathbb{O} \rightarrow \mathbb{R}$  such that,

$$\text{for each pair } S, T \in \mathbb{S} \text{ with } |S| = |T|, S \succsim T \Leftrightarrow \sum_{\xi \in S} v(\xi) \geq \sum_{\xi \in T} v(\xi). \quad (29)$$

Since  $\mathbb{O}$  is convex and  $v$  is continuous,  $v(\mathbb{O})$  is an interval in  $\mathbb{R}$ . Define a binary relation  $\succsim^*$  on  $v(\mathbb{O}) \times \mathbb{N}$  as follows:

$$\begin{aligned} & \text{for each pair } (x, n), (y, k) \in v(\mathbb{O}) \times \mathbb{N}, \\ & (x, n) \succsim^* (y, k) \text{ if there are } \alpha, \beta \in \mathbb{O} \text{ with } x = v(\alpha), y = v(\beta), n \otimes \alpha \succsim k \otimes \beta. \end{aligned} \quad (30)$$

By (30), for each pair  $(x, n), (y, k) \in v(\mathbb{O}) \times \mathbb{N}$ ,  $(x, n) \succsim^* (y, k)$  implies there are  $\alpha', \beta' \in \mathbb{O}$  with  $x = v(\alpha')$ ,  $y = v(\beta')$ , and  $n \otimes \alpha' \succsim k \otimes \beta'$ . By (29), for each pair  $\alpha, \beta \in \mathbb{O}$ ,  $v(\alpha') = v(\alpha)$  implies that  $n \otimes \alpha' \sim n \otimes \alpha$  and  $v(\beta') = v(\beta)$  implies that  $k \otimes \beta' \sim k \otimes \beta$ . Thus,  $n \otimes \alpha' \succsim k \otimes \beta'$  implies  $n \otimes \alpha \succsim k \otimes \beta$ . Thus,

$$\begin{aligned} & \text{for each pair } \alpha, \beta \in \mathbb{O} \text{ and each pair } k, n \in \mathbb{N}, \\ & n \otimes \alpha \succsim k \otimes \beta \text{ if } \alpha, \beta \in \mathbb{O} \text{ are such that } (v(\alpha), n) \succsim^* (v(\beta), k). \end{aligned} \quad (31)$$

The relation  $\succsim^*$  satisfies the following properties:

- **Completeness:** Let  $(x, n), (y, k) \in v(\mathbb{O}) \times \mathbb{N}$ . Then, there are  $\alpha, \beta \in \mathbb{O}$  such that  $v(\alpha) = x$  and  $v(\beta) = y$ . Since  $\succsim$  is complete, either  $n \otimes \alpha \succsim k \otimes \beta$  or  $n \otimes \alpha \precsim k \otimes \beta$ . By (30),  $n \otimes \alpha \succsim k \otimes \beta$  implies  $(x, n) \succsim^* (y, k)$  and  $n \otimes \alpha \precsim k \otimes \beta$  implies  $(y, k) \succsim^* (x, n)$ . Thus, either  $(x, n) \succsim^* (y, k)$  or  $(y, k) \succsim^* (x, n)$ .
- **Transitivity:** Let  $(x, n), (y, k), (z, l) \in v(\mathbb{O}) \times \mathbb{N}$  be such that  $(x, n) \succsim^* (y, k)$  and  $(y, k) \succsim^* (z, l)$ . Then, there are  $\alpha, \beta, \gamma \in \mathbb{O}$  such that  $v(\alpha) = x$ ,  $v(\beta) = y$ , and  $v(\gamma) = z$ . By (31),  $(x, n) \succsim^* (y, k)$  implies  $n \otimes \alpha \succsim k \otimes \beta$  and  $(y, k) \succsim^* (z, l)$  implies  $k \otimes \beta \succsim l \otimes \gamma$ . Since,  $\succsim$  is transitive,  $n \otimes \alpha \succsim l \otimes \gamma$ . By (30),  $(x, n) \succsim^* (z, l)$ .
- **Condition (26):** Let  $x, y, z \in v(\mathbb{O})$  and  $k, n \in \mathbb{N}$  be such that  $(x, n) \succsim^* (y, k) \succsim^* (z, n)$ . Then, there are  $\alpha, \beta, \gamma \in \mathbb{O}$  such that  $v(\alpha) = x$ ,  $v(\beta) = y$ , and  $v(\gamma) = z$ . By (31),  $n \otimes \alpha \succsim k \otimes \beta \succsim n \otimes \gamma$ . By Proposition 5, there is  $\eta \in \mathbb{O}$  such that  $n \otimes \eta \sim k \otimes \beta$ . By (30), letting  $y' = v(\eta)$ ,  $(y', n) \sim^* (y, k)$ .
- **Condition (27):** Let  $n \in \mathbb{N}$  and  $x, y \in v(\mathbb{O})$ . Then, there are  $\alpha, \beta \in \mathbb{O}$  such that  $v(\alpha) = x$  and  $v(\beta) = y$ . By (29),  $nv(\alpha) \geq nv(\beta)$  if and only if  $n \otimes \alpha \succsim n \otimes \beta$ . Thus,  $x \geq y$  if and only if  $n \otimes \alpha \succsim n \otimes \beta$ . By (30) and (31),  $n \otimes \alpha \succsim n \otimes \beta$  if and only if  $(x, n) \succsim^* (y, n)$ . Thus,  $x \geq y$  if and only if  $(x, n) \succsim^* (y, n)$ .

Since  $v(\mathbb{O})$  is an interval,  $v(\mathbb{O}) = \mathbf{A}_v$ . By Proposition 11, there is  $V : \mathbf{A}_v \times \mathbb{N} \rightarrow \mathbb{R}$  that represents  $\succsim^*$  and is strictly increasing and continuous in its first argument.

By Proposition 6, for each pair  $S, T \in \mathbb{S}$ , there are  $\alpha, \beta \in \mathbb{O}$  such that  $|S| \otimes \alpha \sim S$  and  $|T| \otimes \beta \sim T$ . By (30) and (31),  $V(v(\alpha), |S|) \geq V(v(\beta), |T|)$  if and only if  $|S| \otimes \alpha \succsim |T| \otimes \beta$ . By (29),  $|S| \otimes \alpha \sim S$  if and only if  $|S|v(\alpha) = |S|A(v, S)$  and  $|T| \otimes \beta \sim T$  if and only if  $|T|v(\beta) = |T|A(v, T)$ . Thus,

$$\text{for each pair } S, T \in \mathbb{S}, S \succsim T \Leftrightarrow V(A(v, S), |S|) \geq V(A(v, T), |T|).$$

We omit the straightforward proof that the second statement in Theorem 3 implies the first one.  $\square$

## Proof of Theorem 5

*Proof.* Assume that  $\mathbb{O}$  is countable and that  $\succsim$  satisfies sensitivity, cancellation, and Archimedean.

**Part 1 (construction of  $v$ ):** Fix  $\theta \in \mathbb{O}$  and define  $\succsim^\theta$  as follows: For each pair  $S, T \in \mathbb{S}$ ,  $S \succsim^\theta T$  if there is an integer  $n$  at least as large as  $|S|$  and  $|T|$  such that  $S \oplus [(n - |S|) \otimes \theta] \succsim T \oplus [(n - |T|) \otimes \theta]$ . We show that  $\langle \mathbb{S}, \succsim^\theta, \oplus \rangle$  is an “extensive system” (Roberts and Luce, 1968, page 320). This enables an additive representation of  $\succsim^\theta$  which in turn will be used to construct  $v$ . Claims 1, 2, 3, and 4 below show that  $\langle \mathbb{S}, \succsim^\theta, \oplus \rangle$  satisfies each of the conditions defining an extensive system.

*Claim 0:* For each pair  $S, T \in \mathbb{S}$ , if  $n \geq \max\{|S|, |T|\}$  is such that

$$S \oplus [(n - |S|) \otimes \theta] \succsim T \oplus [(n - |T|) \otimes \theta], \quad (32)$$

then, for each  $m \geq \max\{|S|, |T|\}$ ,  $S \oplus [(m - |S|) \otimes \theta] \succsim T \oplus [(m - |T|) \otimes \theta]$ .

Suppose that  $S, T \in \mathbb{S}$  and  $n$  are as specified in Claim 0. First, we show that we can replace  $n$  by  $m = n + 1$ . By cancellation, (32) implies that

$$S \oplus [(n - |S|) \otimes \theta] \oplus \theta \succsim T \oplus [(n - |T|) \otimes \theta] \oplus \theta.$$

In the above line, the sample in left hand side of  $\succsim$  is equal to  $S \oplus [(n + 1 - |S|) \otimes \theta]$  and the sample in its right hand side is equal to  $T \oplus [(n + 1 - |T|) \otimes \theta]$ . Claim 0 is thus true for  $m = n + 1$ . Repeating the argument, it is also true for each  $m > n$ .



Second, we show that if  $n > \max\{|S|, |T|\}$  we can replace  $n$  by  $m = n - 1$ . In (32), the sample in the left hand side of  $\succsim$  is equal to  $S \oplus [(n - 1 - |S|) \otimes \theta] \oplus \theta$  and the sample in its right hand side is equal to  $T \oplus [(n - 1 - |T|) \otimes \theta] \oplus \theta$ . By cancellation, (32) thus implies that

$$S \oplus [(n - 1 - |S|) \otimes \theta] \succsim T \oplus [(n - 1 - |T|) \otimes \theta].$$

Claim 0 is thus true for  $m = n - 1$ . Repeating the argument, it is also true for each  $m \geq \max\{|S|, |T|\}$  with  $m < n$ .

*Claim 1:  $\succsim^\theta$  is complete and transitive.*

Let  $S, T \in \mathbb{S}$  and suppose that it is not the case that  $S \succsim^\theta T$ . Then, for each  $n \geq \max\{|S|, |T|\}$ ,  $S \oplus [(n - |S|) \otimes \theta] \prec T \oplus [(n - |T|) \otimes \theta]$ . Thus, by definition,  $T \succsim^\theta S$ . Thus,  $\succsim^\theta$  is complete.

Let  $R, S, T \in \mathbb{S}$  and suppose that  $R \succsim^\theta S$  and  $S \succsim^\theta T$ . By the definition of  $\succsim^\theta$  and Claim 0, there is  $n \geq \max\{|R|, |S|, |T|\}$  satisfying

$$R \oplus [(n - |R|) \otimes \theta] \succsim S \oplus [(n - |S|) \otimes \theta] \succsim T \oplus [(n - |T|) \otimes \theta].$$

Since  $\succsim$  is transitive,  $R \succsim T$ . Thus,  $\succsim^\theta$  is transitive.

*Claim 2: For each triple  $R, S, T \in \mathbb{S}$ ,  $S \oplus R \sim^\theta R \oplus S$  and  $[S \oplus R] \oplus T \sim^\theta S \oplus [R \oplus T]$ .*

This follows immediately since, for each triple  $R, S, T \in \mathbb{S}$ ,  $S \oplus R = R \oplus S$  and  $[S \oplus R] \oplus T = S \oplus [R \oplus T]$ .

*Claim 3: For each triple  $R, S, T \in \mathbb{S}$ ,  $R \succsim^\theta S$  if and only if  $R \oplus T \succsim^\theta S \oplus T$ .*

Let  $R, S, T \in \mathbb{S}$  and  $n \geq \max\{|R| + |T|, |S| + |T|\}$ . Then,

$$\begin{aligned} R \succsim^\theta S &\Leftrightarrow R \oplus [(n - |R|) \otimes \theta] \succsim S \oplus [(n - |S|) \otimes \theta] \\ &\Leftrightarrow R \oplus [(n - |R|) \otimes \theta] \oplus T \succsim S \oplus [(n - |S|) \otimes \theta] \oplus T \\ &\Leftrightarrow [R \oplus T] \oplus [(n - |R \oplus T|) \otimes \theta] \oplus [|T| \otimes \theta] \succsim [S \oplus T] \oplus [(n - |S \oplus T|) \otimes \theta] \oplus [|T| \otimes \theta] \\ &\Leftrightarrow [R \oplus T] \oplus [(n - |R \oplus T|) \otimes \theta] \succsim [S \oplus T] \oplus [(n - |S \oplus T|) \otimes \theta] \\ &\Leftrightarrow R \oplus T \succsim^\theta S \oplus T \end{aligned}$$

where the first equivalence follows from the definition of  $\succsim^\theta$  and Claim 0, the second equivalence follows from cancellation, the third equivalence follows from rearranging the previous line, the fourth equivalence follows from cancellation, and the last equivalence follows from the definition of  $\succsim^\theta$ .

*Claim 4:* For each pair  $Q, R \in \mathbb{S}$  with  $Q \succ^\theta R$  and each pair  $S, T \in \mathbb{S}$ , there is a positive integer  $n$  such that  $[n \otimes Q] \oplus S \succsim^\theta [n \otimes R] \oplus T$ .

Let  $Q, R \in \mathbb{S}$  be such that  $Q \succ^\theta R$  and  $S, T \in \mathbb{S}$ . Let  $q, r, s, t$  denote the sizes of  $Q, R, S, T$  respectively. Let  $k \geq \max\{q, r\}$  and  $l \geq \max\{s, t\}$ . By Observation 2,  $Q \succ^\theta R$  implies that  $Q \oplus [(k - q) \otimes \theta] \succ R \oplus [(k - r) \otimes \theta]$ . Thus, by the Archimedean axiom, there is a positive integer  $n$  such that

$$n \otimes \{Q \oplus [(k - q) \otimes \theta]\} \oplus \{S \oplus [(l - s) \otimes \theta]\} \succ n \otimes \{R \oplus [(k - r) \otimes \theta]\} \oplus \{T \oplus [(l - t) \otimes \theta]\}.$$

Rearranging,

$$\{[n \otimes Q] \oplus S\} \oplus \{[nk + l - (nq + s)] \otimes \theta\} \succ \{[n \otimes R] \oplus S\} \oplus \{[nk + l - (nr + t)] \otimes \theta\}$$

Since  $[n \otimes Q] \oplus S$  has size  $nq + s$ ,  $[n \otimes R] \oplus S$  has size  $nr + t$ , and  $nk + l \geq \max\{nq + s, nr + t\}$ , by the definition of  $\succsim^\theta$ ,  $[n \otimes Q] \oplus S \succsim^\theta [n \otimes R] \oplus T$ .

By Claims 1, 2, 3, and 4,  $(\mathbb{S}, \succsim^\theta, \oplus)$  is an extensive system. By Theorem 3 in [Roberts and Luce \(1968\)](#) this is a necessary and sufficient condition for the existence of  $w : \mathbb{S} \rightarrow \mathbb{R}$  such that, for each pair  $S, T \in \mathbb{S}$ ,

- (i)  $w(S \oplus T) = w(S) + w(T)$ ;
- (ii)  $S \succsim^\theta T$  if and only if  $w(S) \geq w(T)$ ;
- (iii) if  $\tilde{w} : \mathbb{S} \rightarrow \mathbb{R}$  satisfies (i) and (ii), then there is a scalar  $a > 0$  such that  $\tilde{w} = aw$ .

By repeatedly applying (i), for each  $S \in \mathbb{S}$ ,  $w(S) = \sum_{\alpha \in \mathbb{O}} S(\alpha)w(\alpha)$ . By the definition of  $\succsim^\theta$ , for each pair  $S, T \in \mathbb{S}$  with  $|S| = |T|$ ,  $S \succsim^\theta T$  if and only if  $S \succsim T$ . Thus, by (ii) and (iii), if  $v$  is a positive affine transformation of  $w$ , then

$$\forall S, T \in \mathbb{S} \text{ with } |S| = |T|, \quad S \succsim T \Leftrightarrow \mathbf{A}(v, S) \geq \mathbf{A}(v, T). \quad (33)$$

By sensitivity,  $v$  is not-constant.

**Part 2 (construction of  $V$ ):** For each  $n \in \mathbb{N}$ , each sample of size  $n$  can be specified as a point in  $\mathbb{O}^n$ . Since  $\mathbb{O}$  is countable, the standard Cantor diagonalization argument shows that  $\mathbb{O}^n$  is itself countable; the set of all samples of size  $n$  is thus countable. Thus,  $\mathbb{S}$ , the union over  $n \in \mathbb{N}$  of all sets of samples of size  $n$  is itself countable. Since  $\succsim$  is a complete and transitive relation on  $\mathbb{S}$ , it has a utility representation  $U : \mathbb{S} \rightarrow \mathbb{R}$  (e.g., [Kreps, 1988](#)): For each pair  $S, T \in \mathbb{S}$ ,

$$U(S) \geq U(T) \text{ if and only if } S \succsim T.$$

Let  $v : \mathbb{O} \rightarrow \mathbb{R}$  denote the non-constant function identified in the first part of the proof. By (33), we can then define  $V : \mathbb{R} \times \mathbb{N} \rightarrow \mathbb{R}$  to be such that,

1. for each  $S \in \mathbb{S}$ ,  $V(\mathbf{A}(v, S), |S|) = U(S)$ ,
2. for each  $k \in \mathbb{N}$ ,  $V(\cdot, k)$  is strictly increasing.

Then, for each pair  $S, T \in \mathbb{S}$ , applying the definition of  $U$  followed by that of  $V$ ,

$$\begin{aligned} S \succsim T &\Leftrightarrow U(S) \geq U(T) \\ &\Leftrightarrow V(\mathbf{A}(v, S), |S|) \geq V(\mathbf{A}(v, T), |T|). \end{aligned}$$

Thus, the first statement in [Theorem 5](#) implies the second one. We omit the straightforward proof that the second statement in [Theorem 5](#) implies the first one.  $\square$

**Remark 1.** *Note that sensitivity is only invoked in the above proof to guarantee that  $v$  is non-constant. The proof can thus also be used to characterize a generalized strength  $\mathcal{E}$  weight representation  $v, V$  where the requirement that  $v$  is non-constant is dropped if the sensitivity axiom is dropped. Thus,  $\succsim$  satisfies cancellation and Archimedean if and only if  $\succsim$  has a strength  $\mathcal{E}$  weight representation  $v, V$  where  $v$  may be constant.*

## Proof of [Theorems 2, 4, and 6](#)

**Proposition 12.** *Suppose that  $\succsim$  satisfies monotonicity, sorting, and that it has a strength  $\mathcal{E}$  weight representation  $v, V$ . For each  $x \in \mathbf{A}_v$ , let  $S^x \in \mathbb{S}$  denote the smallest sample with  $\mathbf{A}(v, S^x) = x$  and let  $s^x = |S^x|$ . Let*

$$B = \{x \in \mathbf{A}_v : \exists k, l \in \mathbb{N}, l < k, V(x, ks^x) < V(x, ls^x)\},$$

$$N = \{x \in \mathbf{A}_v : \exists k, l \in \mathbb{N}, l < k, V(x, ks^x) = V(x, ls^x)\},$$

$$G = \{x \in \mathbf{A}_v : \exists k, l \in \mathbb{N}, l < k, V(x, ks^x) > V(x, ls^x)\}.$$

Then, the following conditions hold:

(a) For each  $n \in \mathbb{N}$  and each  $x \in \mathbf{A}_v$ ,

$$V(x, (n+1)s^x) < V(x, ns^x) \text{ if } x \in B,$$

$$V(x, (n+1)s^x) = V(x, ns^x) \text{ if } x \in N,$$

$$V(x, (n+1)s^x) > V(x, ns^x) \text{ if } x \in G.$$

(b) The sets  $B, N, G$  partition  $\mathbf{A}_v$ .

(c) For each  $(x, y, z) \in B \times N \times G$ ,  $x < y < z$ .

(d) For each  $n \in \mathbb{N}$ , each  $x \in \mathbf{A}_v$ , and each  $S \in \mathbb{S}$  such that  $x = \mathbf{A}(v, S)$ ,

$$V(x, n|S|) > V(x, (n+1)|S|) \text{ if } x \in B,$$

$$V(x, n|S|) = V(x, (n+1)|S|) \text{ if } x \in N,$$

$$V(x, n|S|) < V(x, (n+1)|S|) \text{ if } x \in G.$$

*Proof.* Let all the assumptions and notation be as specified in the above statement.

(a) Let  $n \in \mathbb{N}$  and  $x \in B$ . Since  $x \in B$ , there is a pair  $k, l$  of positive integers such that  $k > l$  and  $V(x, ks^x) < V(x, ls^x)$ . By way of contradiction, suppose that  $n$  is such that  $V(x, (n+1)s^x) \geq V(x, ns^x)$ . Since  $v, V$  is a strength & weight representation,  $k \otimes S^x \prec l \otimes S^x$  and  $(n+1) \otimes S^x \succeq n \otimes S^x$ . By monotonicity, for each  $m \in \mathbb{N}$ ,  $(m+1) \otimes S^x \succeq m \otimes S^x$ . Thus,  $k \otimes S^x \succeq (k-1) \otimes S^x \succeq \dots \succeq l \otimes S^x$ . Thus,  $k \otimes S^x \succeq l \otimes S^x$ , a contradiction. Thus,  $(n+1) \otimes S^x \prec n \otimes S^x$ . Since  $v, V$  is a strength & weight representation,  $V(x, (n+1)s^x) < V(x, ns^x)$ . The proofs that  $V(x, (n+1)s^x) = V(x, ns^x)$  if  $x \in N$  and  $V(x, (n+1)s^x) > V(x, ns^x)$  if  $x \in G$  are analogous.

(b) For each  $x \in \mathbf{A}_v$ , one of the following conditions holds:  $V(x, 2s^x) < V(x, s^x)$ ,  $V(x, 2s^x) = V(x, s^x)$ , or  $V(x, 2s^x) > V(x, s^x)$ . By (a),

$$B = \{x \in \mathbf{A}_v : V(x, 2s^x) < V(x, s^x)\},$$

$$N = \{x \in \mathbf{A}_v : V(x, 2s^x) = V(x, s^x)\},$$

$$G = \{x \in \mathbf{A}_v : V(x, 2s^x) > V(x, s^x)\}.$$

Thus,  $B, N$ , and  $G$  partition  $\mathbf{A}_v$ .

(c) Let  $(x, y, z) \in B \times N \times G$ . By way of contradiction, suppose that  $x \geq y$ . Let  $X = s^y \otimes S^x$  and  $Y = s^x \otimes S^y$  and note that  $|X| = |Y|$ . Since  $\mathbf{A}(v, X) = x$ ,

$A(v, Y) = y$ , and  $v, V$  is a strength & weight representation,  $X \succsim Y$ . By (a),  $V(x, |X|) > V(x, 2|X|)$  and  $V(y, |Y|) = V(y, 2|Y|)$ . Since  $v, V$  is a strength & weight representation,  $X \succ 2 \otimes X$  and  $Y \sim 2 \otimes Y$ . By sorting,  $X \succsim Y$  and  $X \succ 2 \otimes X$  imply that  $Y \succ 2 \otimes Y$ , a contradiction. Thus,  $x < y$ . It remains to show that  $y < z$ . Suppose instead that  $y \geq z$ . Let  $Y = s^z \otimes S^y$  and  $Z = s^y \otimes S^z$  and note that  $|Y| = |Z|$ . Since  $A(v, X) = x$ ,  $A(v, Y) = y$ , and  $v, V$  is a strength & weight representation,  $Y \succsim Z$ . By (a),  $V(y, |Y|) = V(y, 2|Y|)$  and  $V(z, |Z|) < V(z, 2|Z|)$ . Since  $v, V$  is a strength & weight representation,  $Y \sim 2 \otimes Y$  and  $2 \otimes Z \succ Z$ . By sorting,  $Y \succsim Z$  and  $2 \otimes Z \succ Z$  imply that  $2 \otimes Y \succ Y$ , a contradiction. Thus,  $y < z$ . Altogether,  $x < y < z$ .

(d) Let  $x \in B$ ,  $n \in \mathbb{N}$ ,  $S \in \mathbb{S}$  be such that  $A(v, S) = x$ . If  $S = S^x$ , then (d) follows immediately from (a). Suppose instead that  $S \neq S^x$ . By way of contradiction, suppose that  $V(x, (n+1)|S|) \geq V(x, n|S|)$ . Since,  $v, V$  is a strength & weight representation,  $(n+1) \otimes S \succsim n \otimes S$ . Let  $k = |S| + s^x$ . By monotonicity,

$$ks^x \otimes S \succsim (ks^x - 1) \otimes S \succsim \cdots \succsim s^x \otimes S.$$

Thus,  $ks^x \otimes S \succsim s^x \otimes S$ . By (a),  $x \in B$  implies  $V(A(v, S^x), k|S|s^x) < V(A(v, S^x), |S|s^x)$ . Since  $v, V$  is a strength & weight representation,  $k|S| \otimes S^x \prec |S| \otimes S^x$ . Since  $A(v, s^x \otimes S) = x = A(v, |S| \otimes S^x)$ , the sizes of  $s^x \otimes S$  and  $|S| \otimes S^x$  are equal, and  $v, V$  is strength & weight representation,  $s^x \otimes S \sim |S| \otimes S^x$ . Thus,  $ks^x \otimes S \succ k|S| \otimes S^x$ . However, since  $A(v, k|S| \otimes S^x) = x = A(v, ks^x \otimes S)$ , the sizes of  $k|S| \otimes S^x$  and  $ks^x \otimes S$  are the same, and  $v, V$  is strength & weight representation,  $k|S| \otimes S^x \sim ks^x \otimes S$ . This contradiction confirms that  $V(x, n|S|) > V(x, (n+1)|S|)$ . The proof of the conditions for  $N$  and  $G$  in (d) are analogous.  $\square$

**Proposition 13.** *Suppose that  $\succsim$  satisfies monotonicity, sorting, approachability, and that it has a strength & weight representation  $v, V$ . Then  $v, V$  is convergent.*

*Proof.* Let all the assumptions and notation be as specified in the above statement. Additionally, let  $B, N, G$  denote the partition of  $A_v$  described in Proposition 12. The first step is showing that,

$$\begin{aligned} &\text{for each } x \in A_v \text{ and each pair } S, T \in \mathbb{S} \text{ with } A(v, S) = x = A(v, T), \\ &\{V(x, n|S|)\}_n \text{ and } \{V(x, n|T|)\}_n \text{ converge to the same limit.} \end{aligned} \tag{34}$$

Let  $x \in \mathbf{A}_v$  and  $S, T \in \mathbb{S}$  be such that  $\mathbf{A}(v, S) = x = \mathbf{A}(v, T)$ . For each  $n \in \mathbb{N}$ , define  $a_n = V(x, n|S|)$  and  $b_n = V(x, n|T|)$ . By Proposition 12,

$$\begin{aligned} & \text{the sequences } \{a_n\} \text{ and } \{b_n\} \text{ are strictly decreasing if } x \in B, \\ & \text{constant if } x \in N, \text{ and} \\ & \text{strictly increasing if } x \in G. \end{aligned} \tag{35}$$

We now show that

$$\begin{aligned} & \{a_n\} \text{ and } \{b_n\} \text{ have a common subsequence,} \\ & \text{are bounded above if there is } y \in \mathbf{A}_v \text{ with } y > x, \text{ and} \\ & \text{are bounded below if there is } y \in \mathbf{A}_v \text{ with } y < x. \end{aligned} \tag{36}$$

For each  $k \in \mathbb{N}$ ,  $a_{k|S||T|} = V(x, k|T||S|) = b_{k|S||T|}$ . Thus,  $\{a_n\}$  and  $\{b_n\}$  have a common subsequence. Let  $y \in \mathbf{A}_v$  and  $Q \in \mathbb{S}$  be such that  $\mathbf{A}(v, Q) = y$ . Define  $S' = |Q| \otimes S$  and  $Q' = |S| \otimes Q$ . Since  $S'$  and  $Q'$  are of equal size,  $\mathbf{A}(v, S') = x$ ,  $\mathbf{A}(v, Q') = y$ , and  $\succsim$  has a strength & weight representation, the following is true:  $Q' \succ S'$  if  $y > x$  and  $Q' \prec S'$  if  $y < x$ . By Proposition 1, for each positive integer  $m$ ,  $m \otimes Q' \succ m \otimes S'$  if  $y > x$  and  $m \otimes Q' \prec m \otimes S'$  if  $y < x$ . Thus, by approachability, there is a positive integer  $m$  such that, for each integer  $k$  greater than  $m$ ,  $m \otimes Q' \succ k \otimes S'$  if  $y > x$  and  $k \otimes Q' \prec m \otimes S'$  if  $y < x$ . Thus, there is a positive integer  $m$  such that, for each integer  $k$  greater than  $m$ ,  $a_{k|S'|} = V(x, k|S'|) < V(y, m|Q'|)$  if  $y > x$  and  $a_{k|S'|} = V(x, k|S'|) > V(y, m|Q'|)$  if  $y < x$ . Thus,  $\{a_n\}$  has a subsequence that is bounded above if  $y < x$  and it has a subsequence that is bounded below if  $y < x$ . By (35),  $\{a_n\}$  is monotone. Thus,  $\{a_n\}$  is itself is bounded above if  $y > x$  and bounded below if  $y < x$ . The same argument can be applied to reach the same conclusions for  $\{b_n\}$ .

The sequences  $\{a_n\}$  and  $\{b_n\}$  converge to the same limit:

Case 1.  $x \in B$ . By (35), if  $x \in B$ , then  $\{a_n\}$  and  $\{b_n\}$  are decreasing. By (36), if there is  $y \in \mathbf{A}_v$  such that  $y < x$ , both sequences are bounded below. Thus, both sequences are convergent. If there is no  $y \in \mathbf{A}_v$  such that  $y < x$ , then  $\{a_n\}$  and  $\{b_n\}$  are also convergent since they are decreasing, this allowing for the possibility that the limit of one or both of the sequences is  $-\infty$ . By (36),  $\{a_n\}$  and  $\{b_n\}$  have a common subsequence and thus converge to same limit.

Case 2.  $x \in N$ . By (35),  $\{a_n\}$  and  $\{b_n\}$  are constant. By (36), both sequences have a common subsequence and are thus identical. Trivially, both sequences converge to the same limit.

Case 3.  $x \in G$ . By (35), if  $x \in G$ , then  $\{a_n\}$  and  $\{b_n\}$  are increasing. By (36), if there is  $y \in \mathbf{A}_v$  with  $y > x$ , both sequences are bounded above. Thus, both sequences are convergent. If there is no  $y \in \mathbf{A}_v$  with  $y > x$ , then  $\{a_n\}$  and  $\{b_n\}$  are also convergent since they are increasing, this time allowing for the possibility that the limit of one or both of the sequences is  $\infty$ . By (36),  $\{a_n\}$  and  $\{b_n\}$  have a common subsequence and thus converge to same limit.

The convergence of  $\{a_n\}$  and  $\{b_n\}$  to a common limit proves (34).

Let  $\mathcal{V} : \mathbf{A}_v \rightarrow \mathbb{R} \cup \{-\infty, \infty\}$  be such that, for each  $x \in \mathbf{A}_v$ ,  $\mathcal{V}(x)$  is the limit of the sequence  $\{V(x, n|S|)\}_n$  where  $S \in \mathbb{S}$  is such that  $\mathbf{A}(v, S) = x$ . By (34), the function  $\mathcal{V}$  is well defined.

To show that  $\mathcal{V}$  is increasing, let  $x, y \in \mathbf{A}_v$  be such that  $x > y$ . Let  $S, T \in \mathbb{S}$  be such that  $\mathbf{A}(v, S) = x$  and  $\mathbf{A}(v, T) = y$ . Let  $S' = |T| \otimes S$  and  $T' = |S| \otimes T$ . Since  $\mathbf{A}(v, S') = x$ ,  $\mathbf{A}(v, T') = y$ ,  $|S'| = |T'|$ , and  $\succsim$  has a strength & weight representation,  $T' \succ S'$ . By Proposition 1, for each positive integer  $m$ ,  $m \otimes T' \succ m \otimes S'$ . Thus, by approachability, there is a positive integer  $m$  such that, for each integer  $k$  greater than  $m$ ,  $k \otimes T' \succ m \otimes S'$  and  $m \otimes T' \succ k \otimes S'$ . Thus, there is a positive integer  $m$  such that,

$$\begin{aligned} \text{for each integer } k \text{ greater than } m, V(y, k|T'|) &> V(x, m|S'|) \text{ and} \\ V(y, m|T'|) &> V(x, k|S'|). \end{aligned} \tag{37}$$

Case 1.  $y \in B$ . By (c) in Proposition 12,  $y \in B$  and  $y > x$  imply that  $x \in B$ . By (d) in Proposition 12,  $\{V(y, k|T'|)\}_k$  and  $\{V(x, k|S'|)\}_k$  are strictly decreasing. By the definition of  $\mathcal{V}$ , the first of these sequences converges to  $\mathcal{V}(y)$  and the second one converges to  $\mathcal{V}(x)$ . By the first inequality in (37),  $V(x, m|S'|)$  is a lower bound for  $\{V(y, k|T'|)\}_k$ . Thus,  $\mathcal{V}(y) \geq V(x, m|S'|) > \mathcal{V}(x)$ .

Case 2.  $y \in N$ . By (c) in Proposition 12,  $y \in N$  and  $y > x$  imply that either  $x \in B$  or  $x \in N$ . By (d) in Proposition 12,  $\{V(y, k|T'|)\}_k$  is a constant sequence and  $\{V(x, k|S'|)\}_k$  is strictly decreasing if  $x \in B$  and constant if  $x \in N$ . By

the definition of  $\mathcal{V}$ , the first of these sequences converges to  $\mathcal{V}(y)$  and the second one converges to  $\mathcal{V}(x)$ . Since  $S'$  and  $T'$  are of the same size and  $y > x$ ,  $V(y, m|T'|) > V(x, m|S'|)$ . Thus,  $\mathcal{V}(y) = V(y, m|T'|) > V(x, m|S'|) \geq \mathcal{V}(x)$ .

Case 3.  $y \in G$ . By (d) in Proposition 12,  $\{V(y, k|T')\}_k$  is strictly increasing. By the second inequality in (37),  $V(y, m|T')$  is an upper bound for  $\{V(x, k|S')\}_k$ . By the definition of  $\mathcal{V}$ , the first of these sequences converges to  $\mathcal{V}(y)$  and the second one converges to  $\mathcal{V}(x)$ . Thus,  $\mathcal{V}(y) > V(y, m|T') \geq \mathcal{V}(x)$ .

Thus, in each of the possible cases,  $\mathcal{V}(y) > \mathcal{V}(x)$  if  $y > x$ . The strength & weight representation  $v, V$  is thus convergent.  $\square$

*Proof of Theorem 2.* Suppose that the set of outcomes has a cardinality of two and let  $\mathbb{O} = \{\alpha, \beta\}$ . Suppose that  $\succsim$  satisfies sensitivity, cancellation, monotonicity, sorting, and approachability. By Theorem 1,  $\succsim$  has a strength and weight representation  $v, V$ . By Proposition 13,  $v, V$  is convergent.

It remains to show that  $v, V$  is ordered. For each  $x \in \mathbf{A}_v$ , let  $S^x \in \mathbb{S}$  denote the smallest sample with  $\mathbf{A}(v, S^x) = x$  and let  $s^x = |S^x|$ . We first show that,

$$\text{for each } x \in \mathbf{A}_v \text{ and each } S \in \mathbb{S} \text{ with } \mathbf{A}(v, S) = x, |S| \text{ is a multiple of } s^x. \quad (38)$$

Let  $x \in \mathbf{A}_v$  and  $S \in \mathbb{S}$  be such that  $\mathbf{A}(v, S) = x$ . If  $s^x = 1$ , then (38) is true since  $|S|$  is a multiple of one. Suppose instead that  $s^x > 1$ . Since  $v$  is not constant,  $\mathbf{A}(v, S) = x$  implies that  $S(\alpha)/|S| = S^x(\alpha)/s^x$ . Letting  $n$  denote the largest integer not exceeding  $S(\alpha)/S^x(\alpha)$ ,

$$|S| = \frac{S(\alpha)}{S^x(\alpha)} s^x = n s^x + \frac{S(\alpha) - n S^x(\alpha)}{S^x(\alpha)} s^x.$$

If the second term in the above summation is zero, then (38) is proven. By way of contradiction suppose that this is not the case. Since  $|S|$  and  $n s^x$  are integers, is an integer and  $\frac{S(\alpha) - n S^x(\alpha)}{S^x(\alpha)} s^x \equiv k$  is also an integer. Then,  $\frac{S^x(\alpha)}{s^x} = \frac{S(\alpha) - n S^x(\alpha)}{k}$ . By the definition of  $n$ ,  $S(\alpha) - n S^x(\alpha) < S^x(\alpha)$ . Then,  $k < s_x$ . Let  $T \in \mathbb{S}$  be such that  $|T| = k$  and  $T(\alpha) = S(\alpha) - n S^x(\alpha)$ . Then,  $\mathbf{A}(v, T) = x$  and  $|T| < s_x$ . This is a contradiction since, by definition,  $S^x$  as the smallest sample such that  $\mathbf{A}(v, S) = x$ . Thus, there is  $n \in \mathbb{N}$  such that  $|S| = n s^x$ , confirming (38).



By Proposition 12,  $\mathbf{A}_v$  can be partitioned into up to three sets  $B, N, G$  such that, for each  $(x, y, z) \in B \times N \times G$ ,  $x < y < z$  and, for each  $n \in \mathbb{N}$ ,

$$\begin{aligned} V(x, (n+1)s^x) &< V(x, ns^x) \text{ if } x \in B, \\ V(y, (n+1)s^y) &= V(y, ns^y) \text{ if } y \in N, \\ V(z, (n+1)s^z) &> V(z, ns^z) \text{ if } z \in G. \end{aligned}$$

By (38), for each  $x \in B$  and each  $n \in \mathbb{N}$ , there are no  $S \in \mathbb{S}$  and  $k \in \mathbb{N}$  such that  $\mathbf{A}(v, S) = x$  and  $|S| = ns^x + k \leq (n+1)s^x - 1$ . Thus, the values

$$V(x, ns^x + 1), V(x, ns^x + 2), \dots, V(x, (n+1)s^x - 1)$$

can be chosen so that

$$V(x, ns^x) > V(x, ns^x + 1) > \dots > V(x, (n+1)s^x - 1) > V(x, (n+1)s^x).$$

Then, for each  $x \in B$  and each  $n \in \mathbb{N}$ ,  $V(x, n+1) < V(x, n)$ . Similarly, we can specify the values of  $V$  so that, for each  $x \in N$  and each  $n \in \mathbb{N}$ ,  $V(x, n+1) = V(x, n)$ , and, for each  $x \in G$  and each  $n \in \mathbb{N}$ ,  $V(x, n+1) > V(x, n)$ . Thus,  $v, V$  satisfies condition (5) and is hence ordered. The first statement in Theorem 2 thus implies the second one. We omit the straightforward proof that the second statement in Theorem 2 implies the first one.  $\square$

*Proof of Theorem 4.* Suppose that  $\mathbb{O}$  is a convex subset of Euclidean space.

Suppose that  $\succsim$  satisfies sensitivity, cancellation, monotonicity, sorting, approachability, and continuity. By Theorem 3,  $\succsim$  has a strength and weight representation  $v, V$ . By Proposition 13,  $v, V$  is convergent.

It remains to show that  $v, V$  is ordered. For each  $x \in \mathbf{A}_v$ , let  $S^x \in \mathbb{S}$  denote the smallest sample with  $\mathbf{A}(v, S^x) = x$  and let  $s^x = |S^x|$ . By Theorem 3,  $v$  is continuous. Thus,  $v(\mathbb{O})$  is an interval in the real line and  $v(\mathbb{O}) = \mathbf{A}_v$ . Thus, by Proposition 12,  $v(\mathbb{O})$  can be partitioned into up to three sets  $B, N, G$  so that, letting  $(x, y, z) \in B \times N \times G$ ,  $x < y < z$ . Thus, there are  $\alpha, \beta, \gamma \in \mathbb{O}$  such that  $v(\alpha) = x$ ,  $v(\beta) = y$ , and  $v(\gamma) = z$ . Thus,  $s^\alpha = s^\beta = s^\gamma = 1$ . By (a) in Proposition 12,

$$V(x, n) > V(x, n+1), V(y, n) = V(y, n+1), V(z, n) < V(z, n+1).$$

Thus,  $v, V$  satisfies condition (5) and is hence ordered. The first statement in Theorem 4 thus implies the second one. We omit the straightforward proof that the second statement in Theorem 4 implies the first one.  $\square$

*Proof of Theorem 6.* Theorem 6 is an immediate corollary of Theorem 7 below.  $\square$

**Theorem 7.** *If the set of outcomes  $\mathbb{O}$  is either finite or countably infinite, then the following statements are equivalent.*

- i.  $\succsim$  satisfies sensitivity, cancellation, Archimedean, monotonicity, sorting, and approachability.
- ii.  $\succsim$  has a convergent strength  $\mathcal{E}$  weight representation  $v, V$  with the following property:  $A_v$  can be partitioned in up to three sets  $B, N, G$  such that, for each  $(x, y, z) \in B \times N \times G$ ,  $x < y < z$  and, for each  $n \in \mathbb{N}$ ,
  - $V(x, n|S|) > V(x, (n+1)|S|)$  if  $x \in B$  and  $S \in \mathbb{S}$  satisfies  $A(v, S) = x$ ,
  - $V(y, n|S|) = V(y, (n+1)|S|)$  if  $y \in N$  and  $S \in \mathbb{S}$  satisfies  $A(v, S) = y$ ,
  - $V(z, n|S|) < V(z, (n+1)|S|)$  if  $z \in G$  and  $S \in \mathbb{S}$  satisfies  $A(v, S) = z$ .

*Proof.* Suppose the set of outcomes  $\mathbb{O}$  is either finite or countably infinite.

Suppose that  $\succsim$  satisfies sensitivity, cancellation, Archimedean, monotonicity, sorting, and approachability. Then, by Theorem 5,  $\succsim$  has a strength & weight representation  $v, V$ . By Proposition 13 and by (d) in Proposition 12,  $v, V$  satisfies all the conditions in the the second statement (ii) of Theorem 7. The first statement in Theorem 7 thus implies the second one. We omit the straightforward proof that the second statement in Theorem 7 implies the first one.  $\square$