Representativeness and Similarity

Chen Zhao*

The University of Hong Kong

February 26, 2018

Abstract

We provide a framework for analyzing a range of well-documented non-Bayesian behaviors including base rate neglect, conjunction fallacy and disjunction fallacy. The model that we propose formally links the concept of similarity in theoretical psychology with belief updating. We follow Kahneman and Tversky (1974) and assume that when attempting to respond to the question "how likely is A given B", people mistakenly respond to the question "how similar are A to B". With a similarity-based updating rule the posterior of $A \cup C$ given B may be less than the posterior of A given B, simply because $A \cup C$ differs more from B than A does when $B \cap C = \emptyset$. Our axioms yield a Cobb-Douglas weighted geometric mean of $\mu(A|B)$ and $\mu(B|A)$ as the behavioral conditional probability of A given B, where μ is the correct subjective probability and $\mu(\cdot|\cdot)$ is the Bayesian conditional of μ. That is, our decision makers confuse these two conditional probabilities but have correct unconditional beliefs. This combination of correct priors and incorrect updating occurs often since in many experiments subjects are explicitly given the relevant prior probabilities.

*I am grateful to Roland Benabou, Faruk Gul, Wolfgang Pesendorfer and participants in the Princeton Microeconomic Theory Student Seminar for their insightful comments. All errors are mine.

1 Introduction

It is well-documented in the experimental literature that people's inferences sometimes deviate significantly from what Bayesian theory predicts. Among these deviations, the salient ones include base rate neglect, conjunction fallacy and disjunction fallacy.

Agents subject to *Base rate neglect* undervalue the base rate and update too much on the signal. In Kahneman and Tversky (1982), subjects are presented with the following problem:

> A cab was involved in a hit-and-run accident at night. Two cab companies, the Green and the Blue, operate in the city. You are given the following data: (i) 85% of the cabs in the city are Green and 15% are Blue, (ii) A witness identified the cab as a Blue cab. The court tested his ability to identify cabs under the appropriate visibility conditions. When presented with a sample of cabs (half of which were Blue and half of which were Green) the witness made correct identifications in 80% of the cases and erred in 20% of the cases. Question: What is the probability that the cab involved in the accident was Blue rather than Green?

While the Bayesian conditional probability is 0.41, the median and mode answer given by the subjects is 0.8, the credibility of the witness. The vast majority state that the cab was more likely to be blue. These responses show that subjects tend to place too small a weight on the base rate 15/85 and thereby rely too much on the signal.[1]

Agents who commit the *conjunction fallacy* assign a higher probability to event $A \cap B$ than to A. In Kahneman and Tversky (1983), subjects are given a description (D) of a stereotypical leftist, named Linda. They are asked to choose the more probable option over:

(T) Linda is a bank teller;

(T∧F) Linda is a bank teller and is active in feminist movements.

[1] There are, also, many extreme examples of base rate neglect in the legal process. In these examples, the very low probability of a particular set of evidence (B) given innocence (A) is cited as proof that given the evidence, the individual is very unlikely be innocent (A given B). For a detailed discussion see Fienberg (1989); DeGroot, Fienberg, and Kadane (1994) and Tribe (1971).

T$\wedge$F, albeit a sub-event of T, is selected by 85% of the subjects as the more probable one. Moreover, the authors report that the incidence of such a fallacy is fairly high (36%) even among statistically sophisticated subjects.

In a similar vein, Bar-Hillel and Neter (1993) define the *disjunction fallacy* as assigning higher probability to A than to $A \cup B$.[2] In the experiments, subjects are given a description of Danielle, who is sensitive and writes poetry secretly. 50% of the subjects are willing to bet on her being a literature major rather than a humanities major even though almost all subjects seem to know that the department of literature is a part of the faculty of humanities.

These systematic violations of Bayesianism are not mutually exclusive. A common theme of these experiments is that subjects is generally not sensitive enough to prior frequencies and tend to over-update on signals. Despite such a clear connection, a unified theory that reconciles the results seems elusive, if not nonexistent. Existing models of updating heuristics typically assume that the decision maker fails to access the full and correct prior when applying Bayes' formula, due to imperfect memory or recall (Mullainathan (2002b); Wilson (2014); Gennaioli and Shleifer (2010); Bordalo, Coffman, Gennaioli, and Shleifer (2016)), or categorical thinking (Mullainathan (2002a); Mullainathan, Schwartzstein, and Shleifer (2008)), or having an incorrect model in mind (Barberis, Shleifer, and Vishny (1998); Rabin and Schrag (1999); Rabin (2002); and Schwartzstein (2014)). None of these contributions offer a satisfactory explanation to experiments in which the relevant prior probabilities are fully specified and available to subjects. The fact that subjects have correct unconditional beliefs simply leaves little freedom to updating heuristics that maintain Bayes' formula.

In this paper, we provide a unified micro-foundation for behavioral subjects who tend to commit the fallacies above but have correct unconditional beliefs. Our decision maker assigns subjective probabilities rationally to all events but she adopts a behavioral updating rule. In particular, the decision maker mistakes the question"How likely is A given B?" for the question "How representative is A of B (i.e., how similar is A to B)?". We base this

[2] See, also Fischhoff, Slovic, and Lichtenstein (1978).

behavioral twist on Kahneman and Tversky (1974), who propose that,

> *"in answering such [conditional probability] questions, people typically rely on the representativeness heuristic, in which probabilities are evaluated by the degree to which A is representative of B, that is, by the degree to which A resembles B."*

Hence, we assume that our decision maker assesses the degree to which A is similar to B when asked to evaluate the conditional probability of A given B. In the Linda problem, T$\wedge$F resembles a leftist more than T does; in the Danielle problem, a literature student is more representative of sensitive poets than a humanities student is. Although it seems that no judgements of representativeness is involved in the taxicab problem, under our characterization of similarity, the event "the involved cab was blue" shares higher similarity with the event "witness identifed the cab as blue" than "the involved cab was green" does.

Our primitive is a binary relation between *two pairs* of events of the state space Ω: $(A, B) \succeq (C, D)$ means that A is more similar to B than C is to D; that is, the decision maker assigns higher probability to A given B than to C given D. We impose axioms on $\succeq$ to identify a similarity index (or equivalently, a similarity-based updating formula), $S(A, B)$, such that $(A, B) \succeq (C, D)$ if and only if $S(A, B) \geq S(C, D)$.[3]

Our characterization of $S(A, B)$ are twofold. First, our axioms yield a unique subjective probability measure μ and a nondecreasing aggregator f such that

$$S(A, B) = f(\mu(A|B), \mu(B|A)).$$

where $\mu(\cdot|\cdot)$ is given by Bayes' rule. That is, the decision maker confuses the correct conditional $\mu(A|B)$ with the reversed conditional $\mu(B|A)$. Second, we introduce a robustness condition which ensures that independent events are not informative when conditioned on.

[3]We follow the psychology literature and allow $S(A, B)$ to be asymmetric; that is, $S(A, B)$ need not equal $S(B, A)$. For example, we say that "an ellipse is similar to a circle" but not "a circle is similar to an ellipse." See Tversky (1977) for a detailed analysis.

With this additional condition we obtain the following representation:

$$S(A,B) = \mu(A|B)^{\alpha}\,\mu(B|A)^{1-\alpha}$$

where $\alpha \in (0,1]$ measures the decision maker's deviation from Bayesianism. In particular, when $\alpha = 1$, the decision maker is Bayesian. When $\alpha \to 0$, the decision maker fully confuses the two conditional probabilities.

The subjective probability measure μ in our representation is ordinally equivalent to $S(\cdot,\Omega)$; that is, μ describes the decision maker's probabilistic rankings over unconditional events. Hence, if the decision maker does not err in comparing unconditional events, or is given directly the prior probabilities, μ is indeed the correct *objective* prior.

The Cobb-Dougas updating rule $S(A,B)$, in its likelihood ratio form, naturally explains base rate neglect. Consider evaluating whether A or B is more likely given C, we have

$$\frac{S(A,C)}{S(B,C)} = \frac{\mu(A|C)^{\alpha}\mu(C|A)^{1-\alpha}}{\mu(B|C)^{\alpha}\mu(C|B)^{1-\alpha}} = \frac{\mu(C|A)}{\mu(C|B)} \cdot \left(\frac{\mu(A)}{\mu(B)}\right)^{\alpha}.$$

Since $\alpha \leq 1$, $(\mu(A)/\mu(B))^{\alpha}$ is closer to 1 than the correct prior odds ratio $\mu(A)/\mu(B)$ is. Hence, α specifies exactly the extent to which the decision maker neglects the prior odds ratio. In the taxicab problem, the likelihood ratio of the witness' signal (correct over incorrect) is $80/20 = 4$ where as the base rate (blue over green) is $15/85 = 0.176$. A decision maker with an $\alpha < 0.8$ would conclude, like the experimental subjects, that the hit and run cab was more likely to be blue than green.

Moreover, our similarity representation, $S(\cdot,\cdot)$, is also consistent with experimental evidence on conjunction fallacy and disjunction fallacy. In the Linda problem, although $\mu(\mathrm{T}\wedge\mathrm{F}|\mathrm{D}) < \mu(\mathrm{T}|\mathrm{D})$, arguably $\mu(\mathrm{D}|\mathrm{T}\wedge\mathrm{F}) > \mu(\mathrm{D}|\mathrm{T})$; that is, Linda is more likely to match the stereotypical description D if she is a feminist bank teller than if she is just a bank teller. Then, if α is small, an α-mixture of $\mu(\mathrm{T}\wedge\mathrm{F}|\mathrm{D})$ and $\mu(\mathrm{D}|\mathrm{T}\wedge\mathrm{F})$ may well be higher than an α-mixture of $\mu(\mathrm{T}|\mathrm{D})$ and $\mu(\mathrm{D}|\mathrm{T})$. In that case a decision maker would choose the conjunctive description T∧F as the more "probable" option. A similar logic applies to the results in

Bar-Hillel and Neter (1993).

As our representation generalizes Baye's rule, each axiom that we propose is consistent with Bayesian thinking. The key postulate which generates deviation from Bayes' rule is *monotonicity*: it allows the decision maker to specify $\mu(A \cup C|B) < \mu(A|B)$ if $B \cap C = \emptyset$, simply because $A \cup C$ differs more from B than A does. This axiom reflects exactly Tversky (1977)'s insight that similarity depends on both common and distinctive features, and that it should be diminished by the addition of distinctive features. We describe our axiomatic system in details in section 3.

Our approach in the paper incorporates results from Savage (1954) and Villegas (1964) to obtain a nonatomic subjective probability. Then, we exploit the nonatomicity and apply a Debreu-type ordinal representation argument to prove the existence of a similarity index. Finally, we embed the similarity relation into a two-state Anscombe-Aumann framework with a novel mixture operation and prove that such an index must be a weighted geometric average of the conditional probabilities.[4]

In section 2, we discuss the related literatue. In section 3, we present our systems of axioms and show that they are equivalent to our (f, μ) representation. We also show that with an additional robustness condition, f must be of the Cobb-Douglas form. In section 4, we summarize our method of axiomatizing a general class of means by applying results in Anscombe and Aumann (1963). Section 5 concludes.

2 Related Literature

Psychologists have proposed many similarity indices in the context of pattern recognition. Within the literature, this paper is related to Tversky (1977) and Krantz, Luce, Suppes, and Tversky (1999). Tversky (1977) models the similarity relation as an additive conjoint

[4]We feel that this paper is itself a semester long course on introductory decision theory.

structure and proves that there exists a representation

$$S(A,B)=\theta f(A\cap B)-\alpha f(A\backslash B)-\beta f(B\backslash A)$$

for some $\theta,\alpha,\beta\geq 0$ and an interval scale f.[5] In a similar vein, Krantz, Luce, Suppes, and Tversky (1999) axiomatize the following similarity index:

$$S(A,B)=\phi_1(A\cap B)-\phi_2(A\backslash B)-\phi_3(B\backslash A)$$

where ϕ_1,ϕ_2,ϕ_3 are interval scales. They also provide conditions under which ϕ_i's are additive.

In our context, similarity is a substitute for conditional probability. Therefore, it is natural to assume that the decision maker assigns maximum similarity to any event with itself. Such a property precludes the linear representations in both Tversky (1977) and Krantz, Luce, Suppes, and Tversky (1999): both representations imply that $S(A,A)$ is larger than $S(B,B)$ if A is more likely than B. In fact, the main difficulty of the proofs in this paper arises from abandoning the linear structure. Without linearity, we are forced to use a Debreu-type ordinal representation argument for identification of a similarity index.

3 Model

In this section we first describe the primitives of our model. Then we introduce our axioms and show that they are equivalent to the (f,μ) representation. After that, we identify the robustness condition which ensures that f is of the Cobb-Douglas form.

Let $\mathcal{E}$ be a σ-algebra defined on state space Ω and $\mathcal{N}\subset\mathcal{E}$ be a σ-ideal. $\mathcal{E}$ is the set of all events and $\mathcal{N}$ is the collection of all null events. Let $\succeq$ be a binary relation defined on $\mathcal{E}^2\backslash\mathcal{N}^2$ that summarizes the decision maker's *similarity* assessments. Read $(A,B)\succeq(C,D)$ as "A

[5] In the paper Tversky also proposes, albeit without axiomatization, the well-known Tversky index:

$$S_T(A,B)=\frac{f(A\cap B)}{f(A\cap B)+\alpha f(A\backslash B)+\beta f(B\backslash A)}$$

which encompasses a large class of similarity indices in the literature. Using our framework, we provide an axiomatic foundation for S_T when $\alpha+\beta=1$ in the appendix.

is more similar to B than C is to D". We do not consider similarity rankings between null events since they are irrelevant in our updating context. We follow the psychology literature and allow $\succeq$ to be asymmetric; that is, $(A,B) \sim (B,A)$ is *not* true in general.[6] For each ordered pair (A,B), we will call A the *stimulus* and B the *standard*.

We assume that the decision maker mistakenly ranks the conditional likelihood of events according to $\succeq$; that is, if $(A,B) \succeq (C,D)$, she concludes that A given B is more likely than C given D. We base this behavioral twist on Kahneman and Tversky (1974), who propose that,

> *"in answering such [conditional probability] questions, people typically rely on the representativeness heuristic, in which probabilities are evaluated by the degree to which A is representative of B, that is, by the degree to which A resembles B."*

Hence, we assume that our decision maker assesses the degree to which A is similar to B when asked to evaluate $\Pr(A|B)$. In the Linda problem, T∧F resembles a leftist more than T does; in the Danielle problem, a literature student is more representative of sensitive poets than a humanities student is. Although it seems that no judgments of representativeness is involved in the taxicab problem, under our characterization of similarity, the event "the involved cab was blue" shares higher similarity with the event "witness identifed the cab as blue" than "the involved cab was green" does.

As similarity is a substitute for conditional probability in our context, we assume that any nonnull event and itself are maximally similar. In addition, a pair of events has the minimum possible degree of similarity if and only if they have a null intersection.

Definition. $(\Omega, \mathcal{E}, \mathcal{N}, \succeq)$ *is a* ***similarity structure*** *if*

(i) $\succeq$ *is complete and transitive;*

(ii) $(A,A) \succeq (B,C)$ *for all* A,B,C*;*

(iii) $C \cap D \in \mathcal{N} \iff (A,B) \succeq (C,D)$ *for all* A,B.

[6] For example, we say that "an ellipse is similar to a circle" but not "a circle is similar to an ellipse." See Tversky (1977) for a detailed analysis.

An ordered pair of events (A, B) partitions $A \cup B$ into three subsets: the *stimulus distinction* $A \backslash B$, the *standard distinction* $B \backslash A$ and the *intersection* $A \cap B$. We will call the collection of these events the *contrast partition* generated by (A, B). For expositional purpose, we will explicitly list each element of the contrast partition subsequently and substitute (A, B) with $A \backslash B\, A \cap B\, B \backslash A$. In other words, ACB denotes the pair of events $(A \cup C, B \cup C)$ where A, B, C are *pairwise disjoint.*

We consider only nondegenerate similarity structures; that is, Ω cannot be null and there exists at least one pair of events that has an intermediate degree of similarity between the maximum and minimum possible value.

Definition. *A similarity structure* $(\Omega, \mathcal{E}, \mathcal{N}, \succeq)$ *is* ***nondegenerate*** *if there exist* $\widetilde{A}, \widetilde{B}, \widetilde{C}$ *such that* $\emptyset\Omega\emptyset \succ \widetilde{A}\widetilde{C}\widetilde{B} \succ \widetilde{A}\emptyset\widetilde{B}$.

In our original notation, a similarity structure $(\Omega, \mathcal{E}, \mathcal{N}, \succeq)$ is nondegenerate if there exists a *pairwise disjoint* collection $\{\widetilde{A}, \widetilde{B}, \widetilde{C}\}$ such that $(\Omega, \Omega) \succ (\widetilde{A} \cup \widetilde{C}, \widetilde{B} \cup \widetilde{C}) \succ (\widetilde{A}, \widetilde{B})$.

3.1 Similarity Representation

In this subsection we present our axioms and show that the axioms are equivalent to the (f, μ) representation. Our axioms are centered at Tversky (1977)'s insight that similarity depends on both common and distinctive features, and that it should be enhanced by the addition of common features as well as by the reduction of distinctive features.

First note that since Bayes' formula is a special case of our representation, *all* of our axioms are consistent with the Bayesian model. Among them, Axiom 1-3 are natural adaptations of standard axioms by Savage (1954) and Villegas (1964)[7] on subjective probability. In the statement of our axioms, we will explicitly specify each element of the contrast partition for simplicity of exposition; that is, ACB denotes the pair of events $(A \cup C,\ B \cup C)$ where A, B, C are *pairwise disjoint.*

[7]See Axioms S1-S5 in Appendix A for their list of axioms.

Axiom 1. *(Additivity) If* $(A \cup B \cup C \cup D) \cap E = \emptyset$*, then*

$$ACB \succeq ADB \iff A(C \cup E)B \succeq A(D \cup E)B.$$

Axiom 1 is a direct adaptation of the standard additivity axiom in Savage (1954) and Villegas (1964).[8] In particular, our additivity axiom states that controlling for both distinctions (A and B), adding the same common states (E) to both pairs of events does not change their relative similarity ranking. This form of additivity ensures that similarity is enhanced by the addition of common features. To see that, set $D = \emptyset$. By definition of our similarity structure, $ACB \succeq A\emptyset B$. Then it follows from Axiom 1 that $A(C \cup E)B \succeq AEB$; that is, $(A \cup E) \cup C$ is more similar to $(B \cup E) \cup C$ than $A \cup E$ is to $B \cup E$.

The following two axioms are mainly for technical purpose and are adapted from Villegas (1964)'s nonatomicity and monotone continuity.[9]

Axiom 2. *(Nonatomicity) Let* $A'CB' \succ ADB$*. Then respectively,*

$$\begin{array}{lcl} A_1DB \succ ADB \text{ for some } A_1 & & \exists \widehat{A} \subset A \text{ s.t. } A'CB' \succ \widehat{A}DB \succ ADB; \\ ADB_1 \succ ADB \text{ for some } B_1 & \text{implies} & \exists \widehat{B} \subset B \text{ s.t. } A'CB' \succ AD\widehat{B} \succ ADB; \\ A'C_1B' \succ A'CB' \text{ for some } C_1 & & \exists \widehat{C} \subset C \text{ s.t. } A'CB' \succ A'\widehat{C}B' \succ ADB. \end{array}$$

Our version of nonatomicity states that generically, given two pairs of events with different degrees of similarity, one can shrink the intersection of the more-similar pair, or the distinctions of the less-similar pair, to achieve an intermediate degree of similarity. As long as the decision maker is sensitive to stimulus distinctions in the scenario, i.e., $A_1DB \succ ADB$ for some A_1, then it is always possible to shrink the stimulus distinction A to achieve an intermediate pair. In a similar vein, the remaining two cases concern respectively the standard distinction and the intersection.

The following axiom ensures that $\succeq$ is Archimedean and the subjective probability that we will identify is countably additive. It is a direct strengthening of Villegas (1964)'s monotone continuity.

[8] See Axiom S3 in Appendix A.

[9] See Axioms S4-S5 in Appendix A.

Axiom 3. *(Monotone Continuity)* $AC_nB \succeq A'C'B' \succeq A_nCB_n$ *and* $A_{n+1}\subset A_n$, $B_{n+1}\subset B_n$, $C_{n+1}\subset C_n$ *implies* $A\left(\bigcap_n C_n\right)B \succeq A'C'B' \succeq (\bigcap_n A_n)C(\bigcap_n B_n)$.

Axioms 4-6 have no counterpart in Savage's subjective probability theorem but are consistent with rational Bayesian updating.

Axiom 4. *(Separability)* $ACB \succ ADB$ *implies* $A'CB' \succ A'DB'$ *or* $\emptyset\Omega\emptyset \sim A'CB' \sim A'DB'$.

The separability axiom states that when comparing the similarity of ACB versus ADB, the decision maker cancels A and B from both pairs and is essentially ranking C versus D. Compared to standard Debreu-type separability axioms, this axiom is weaker in two aspects. First, we impose cancellation only when the two pairs share both distinctions. Second, on the implication side we allow for a second scenario $\emptyset\Omega\emptyset \sim A'CB' \sim A'D'B$. This scenario arises naturally when, for example, $A' = B' = \emptyset$. In this case the decision maker assigns maximum similarity to $A'CB'$ if C is any nonnull event.

We now introduce our main axiom, *monotonicity*. The key notion behind monotonicity is that states in the intersection affect the similarity of the pair in the opposite way as do the states in the distinctions. We know from Axiom 4 that when comparing ACB with ADB, the decision maker is essentially ranking intersections C and D. Monotoncity then requires that if C contributes more similarity than D as an intersection, then C must also cost more similarity as a distinction. Furthermore, if C contributes strictly more similarity than D as an intersection, then C should impose a strictly larger cost on similarity than D as a *standard* distinction. We do not impose strict monotonicity on *stimulus* distinctions since we inherit the assymmetry between standard distinction and stimulus distinction from Bayes' rule.

Axiom 5. *(Monotonicity)* $ACB \succeq ADB$ *implies* $ABD \succeq ABC$ *and* $DAB \succeq CAB$*. If* $B \notin \mathcal{N}$*, then* $ACB \succ ADB$ *implies* $ABD \succ ABC$*.*

Monotonicity is the main axiom that generates deviations from Bayes' rule. To see that, consider evaluating the probability of $C\cup A$ conditioning on $A\cup B$ where A, B, C are pairwise disjoint; that is, C is the stimulus distinction, A is the intersection and B is the standard distinction. According to Bayes' rule, $\Pr(C \cup A | A \cup B) = \Pr(A) / \Pr(A \cup B)$. It follows that for a Bayesian decision maker the probability of the stimulus distinction C is irrelevant to the posterior assessment, i.e. a Bayesian decision maker always has $CAB \sim \emptyset AB$. In contrast, our monotonicity axiom allows that $\emptyset AB \succ CAB$, reflecting Tversky (1977)'s insight that similarity should be enhanced by the reduction of distinctive features. In fact, if we requires $\emptyset AB \sim CAB$ for any A, B, C *in addition* to Axiom 1-6, the Bayesian model emerges as the unique representation of the decsion maker's updating behavior. We will call this condition Axiom B.

Axiom B. $\emptyset AB \sim CAB$.

Next we introduce our final axiom, *scale invariance.*

Axiom 6. *(Scale Invariance) Let A, A', B, B', C, C' be pairwise disjoint. If $ACB \sim A'C'B'$ and $\emptyset CB \sim \emptyset C'B'$, then $(A \cup A')(C \cup C')(B \cup B') \sim ACB$.*

In standard Bayesian updating, if $\Pr(A_1|A_2) = \Pr(B_1|B_2)$ and $A_1 \cap (B_1 \cup B_2) = B_1 \cap (A_1 \cup A_2) = \emptyset$, then $\Pr(A_1 \cup B_1 | A_2 \cup B_2) = \Pr(A_1|A_2)$. Axiom 6 exactly reduces to this condition if the decision maker is a Bayesian, i.e. if Axiom B holds. In case that the decision maker is not a Bayesian, we weaken the condition by requiring $\emptyset CB \sim \emptyset C'B'$ in addition to $ACB \sim A'C'B'$; that is, the two pairs need to be *truly* in proportion according to a rational Bayesian.

For any probability measure μ, define the Bayesian conditional $\mu(\cdot|\cdot)$ as follows:

$$\mu(A|B) = \begin{cases} \frac{\mu(A \cap B)}{\mu(B)} & \text{if } \mu(B) > 0, \\ 0 & \text{if } \mu(B) = 0. \end{cases}$$

When $\mu(B) = 0$, we set $\mu(A|B) = 0$ for the ease of exposition. Given a similarity structure $(\Omega, \mathcal{E}, \mathcal{N}, \succeq)$, we say that a function $S : \mathcal{E}^2 \backslash \mathcal{N}^2 \to \mathbb{R}$ is a similarity index of the structure if

$$(A, B) \succeq (C, D) \iff S(A, B) \geq S(C, D).$$

Our first theorem shows that if a similarity structure is nondegenerate and obeys Axiom 1-6, then it has a similarity index that is a function of the Bayesian conditionals $\mu(A|B)$ and $\mu(B|A)$ for some probability measure μ. In other words, if the decision maker adopts a similarity-based updating procedure that satisfies our axioms, then there exists a subjective probability μ such that her posterior assessments can be expressed as a function of the relevant two Bayesian conditionals.

Theorem 1. *A nondegenerate similarity structure* $(\Omega, \mathcal{E}, \mathcal{N}, \succeq)$ *satisfies Axiom 1-6 if and only if there exist a nonatomic probability* μ *and a non-decreasing function* $f : [0,1]^2 \longrightarrow [0,1]$ *which is continuous on* $(0,1]^2$ *such that*

$$(A,B) \succeq (C,D) \iff f\left(\mu\left(A|B\right), \mu\left(B|A\right)\right) \geq f\left(\mu\left(C|D\right), \mu\left(C|D\right)\right)$$

where (i) $\mu(A) = 0$ *if and only if* $A \in \mathcal{N}$ *and (ii)* $f(\cdot, y)$ *is strictly increasing if* $y > 0$.

We call the pair (f, μ) a similarity representation of $(\Omega, \mathcal{E}, \mathcal{N}, \succeq)$. The next corollary shows that μ is indeed the decision maker's subjective beliefs on unconditional events.

Corollary 1. *Suppose* (f, μ) *is a similarity representation of a nondegenerate similarity structure* $(\Omega, \mathcal{E}, \mathcal{N}, \succeq)$*, then*

$$\mu(A) \geq \mu(B) \iff (A, \Omega) \succeq (B, \Omega).$$

Hence, the representation in Theorem 1 describes a decision maker, when asked to evaluate $\mu(A|B)$, confuses $\mu(A|B)$ with $\mu(B|A)$. In legal processes, the very low probability of a particular set of evidence (B) given innocence (A) is often cited as proof that given the evidence, the individual is very unlikely be innocent (A given B).[10] Another example is that people often mistakenly treat the accuracy of a medical test as the probability of sickness given a positive result. In fact, these mistakes are made exactly because people are not sensitive enough to prior frequencies. To see that, consider evaluating the likelihood ratio of

[10] For a detailed discussion see Fienberg (1989); DeGroot, Fienberg, and Kadane (1994) and Tribe (1971).

"sick" over "not sick" given that the medical test emerges positive. For a Bayesian decision maker,

$$\frac{\mu(\text{sick}|+)}{\mu(\text{not sick}|+)} = \frac{\mu(+|\text{sick})}{\mu(+|\text{not sick})} \cdot \frac{\mu(\text{sick})}{\mu(\text{not sick})}.$$

If the decision maker is not sensitive enough to the prior frequencies, she concludes that it is more likely for her to be sick since the test is more likely to emerge positive if she is sick. Hence, she behaves as if she mistakes $\mu(\text{sick}|+)$ with $\mu(+|\text{sick})$.

Hence, Theorem 1 provides the micro-foundation for the connection between representativeness heuristics and the observed irresponsiveness to prior frequencies. The next corollary further strengthens this connection by asserting that the only difference between a behavioural decision maker as in Theorem 1 and a rational Bayesian is whether she cares about the stimulus distinction; that is, whether she updates based on representativeness and similarity.

Corollary 2. *A nondegenerate similarity structure* $(\Omega, \mathcal{E}, \mathcal{N}, \succeq)$ *satisfies Axiom 1-6 and Axiom B if and only if there exist a nonatomic probability* μ *such that*

$$(A, B) \succeq (C, D) \iff \mu(A|B) \geq \mu(C|D).$$

where $\mu(A) = 0$ *if and only if* $A \in \mathcal{N}$.

In the proof of Theorem 1, we first incorporate results from Savage (1954) and Villegas (1964) to extract a nonatomic subjective probability from $\succeq$. Then, we exploit the nonatomicity of μ and apply a Debreu-type ordinal representation argument to prove the existence of an aggregator f and therefore a similarity index. See the appendix for detailed arguments.

Our next theorem show that although the aggregator f is only identified ordinally, the subjective prior μ is unique.

Theorem 2. *Let* (f, μ) *be a similarity representation of a nondegenerate similarity structure* $(\Omega, \mathcal{E}, \mathcal{N}, \succeq)$*. Then* (f', μ') *is a similarity representation of* $(\Omega, \mathcal{E}, \mathcal{N}, \succeq)$ *if and only if* $\mu = \mu'$

and there is a strictly increasing function g, continuous on $f((0,1]^2)$, such that $f' = g \circ f$.

Proof. Suppose (f', μ') is a similarity representation. By the corollary we know that $\mu'(A) \geq \mu'(B)$ if and only if $(A, \Omega) \succeq (B, \Omega)$. Hence $\mu(A) \geq \mu(B)$ if and only if $\mu'(A) \geq \mu'(B)$. By the nonatomicity of μ, it is clear that $\mu = \mu'$. The second part is standard and therefore its proof is omitted. □

Hence, Theorem 1 provides a micro-foundation for behavioral subjects who know the correct prior but confuse $\mu(A|B)$ with $\mu(B|A)$. In the next subsection, we characterize a special class of aggregators and then discuss the experimental evidence in detail.

3.2 The Cobb-Douglas Index

In this subsection we take the similarity representation (f, μ) as given and provide a robustness condition which ensures that f has the Cobb-Douglas form. In particular, we assume that the relative similarity ranking between two pairs of events does not change when we intersect both standards or both stimuli with an independent set. Formally, we define independence and then robustness as follows.

Definition. *$A, B \in \mathcal{E}$ are* ***independent****, denoted $A \perp B$, if $\mu(A \cap B) = \mu(A)\mu(B)$.*

For any collection of subsets $\mathcal{A} \subset \mathcal{E}$, we say that $A \perp \mathcal{A}$ if $A \perp B$ for any $B \in \mathcal{A}$. Also, let $\sigma(\mathcal{A})$ be the smallest σ-algebra that contains all of the elements in $\mathcal{A}$.

Definition. *$\succeq$ is said to be* ***robust*** *if $(A, B) \succeq (A', B)$ implies $(A \cap C, B) \succeq (A' \cap C, B)$ and $(A, B \cap C) \succeq (A', B \cap C)$ for any $C \perp \sigma(A, A', B)$.*

With robustness, we can embed our similarity representation in a two-state Anscombe-Aumann framework with a nonstandard mixture operation. The two states of the world can be interpreted as the state in which B caused A and the state in which A caused B. Conditional probabilities $\mu(A|B)$ and $\mu(B|A)$ are viewed as lotteries in the corresponding state. Then, for each pair of events A, B, $(\mu(A|B), \mu(B|A))$ is an Anscombe-Aumann act. Our aggregator f, in turn, implies a complete and transitive ranking over all these Anscombe-Aumann acts. Applying the results in Anscombe and Aumann (1963), yields the representation below.

Theorem 3. *Suppose (f, μ) is a similarity representation of nondegenerate similarity structure $(\Omega, \mathcal{E}, \mathcal{N}, \succeq)$. Then $\succeq$ is robust if and only if there is a unique $\alpha \in (0, 1]$ such that*

$$(A, B) \succeq (C, D) \iff \mu\left(A|B\right)^{\alpha} \mu\left(B|A\right)^{1-\alpha} \geq \mu\left(C|D\right)^{\alpha} \mu\left(D|C\right)^{1-\alpha}.$$

Consider evaluating whether A or B is more likely conditioning on C. Given the Cobb-Douglas index, we have

$$\frac{S(A,C)}{S(B,C)} = \frac{\mu(A|C)^{\alpha}\mu(C|A)^{1-\alpha}}{\mu(B|C)^{\alpha}\mu(C|B)^{1-\alpha}} = \frac{\mu(C|A)}{\mu(C|B)} \cdot \left(\frac{\mu(A)}{\mu(B)}\right)^{\alpha}.$$

Since $\alpha \in (0, 1]$, $(\mu(A)/\mu(B))^{\alpha}$ is closer to 1 than the correct prior odds ratio $\mu(A)/\mu(B)$. That is, α specifies the extent to which the decision maker neglects the prior odds ratio.[11]

In the taxicab problem, the likelihood ratio of the witness' signal (correct over incorrect) is $80/20 = 4$ where as the base rate (blue over green) is $15/85 = 0.176$. A decision maker with an $\alpha < 0.8$ would conclude, like the experimental subjects, that the hit and run cab was more likely to be blue than green.

In the Linda problem, although $\Pr(\text{T}\wedge\text{F}|\text{D}) < \Pr(\text{T}|\text{D})$, arguably we have $\Pr(\text{D}|\text{T}\wedge\text{F}) > \Pr(\text{D}|\text{T})$; that is, Linda is more likely to match the stereotypical description D if she is a feminist bank teller than if she is just a bank teller. Then, if α is low, an α-mixture of $\Pr(\text{T}\wedge\text{F}|\text{D})$ and $\Pr(\text{D}|\text{T}\wedge\text{F})$ may well be higher than an α-mixture of $\Pr(\text{T}|\text{D})$ and $\Pr(\text{D}|\text{T})$. In that case a decision maker would choose the conjunctive description T$\wedge$F as the more "probable" option. A similar logic applies to the results in Bar-Hillel and Neter (1993).

[11] Note that our similarity index is ordinal in nature; that is, the decision maker compares the similarity ratio $S(A, C)/S(B, C)$ with 1. If $S(A, C)/S(B, C)$ is larger than 1, she concludes that A is more likely; if the similarity ratio is less than 1, then she concludes that B is more likely.

4 Method for Axiomatizing Quasi-linear Means

In this section we summarize the method used in this paper for axiomatizing means. Let X be a convex non-singleton subset of $\mathbb{R}$. For any function $g : X^n \longrightarrow X$ we say that g is diagonally-increasing if $x > y$ implies $g(x, x, \ldots, x) > g(y, y, \ldots, y)$.

Let $f : X^n \longrightarrow X$ be a nondecreasing, diagonally-increasing and continuous function. Let $\boldsymbol{x} = (x_1, x_2, \ldots, x_n), \boldsymbol{y} = (y_1, y_2, \ldots, y_n)$ be generic elements of X^n. We would like to find conditions which ensures that f is ordinally equivalent to a ϕ-mean, i.e., $\forall\ \boldsymbol{x}, \boldsymbol{y} \in X^n$,

$$f(\boldsymbol{x}) \geq f(\boldsymbol{y}) \Longleftrightarrow \phi^{-1}\left(\sum_{i=1}^{n} \alpha_i \phi(x_i)\right) \geq \phi^{-1}\left(\sum_{i=1}^{n} \alpha_i \phi(y_i)\right)$$

where $\alpha_i \in [0, 1]$ and $\sum_{i=1}^{n} \alpha_i = 1$, for some $\phi : X \to \mathbb{R}$ that is strictly increasing and continuous.

This class of means is called quasi-linear means. Hong (1983) traces this notion to Kolmogorov (1930) and de Finetti (1931), and provides an axiomatic foundation for a more general, infinite-state version of our definition of ϕ-mean. In contrast, we focus on quasi-linear means of a finite vector of numbers; we appeal to Anscombe and Aumann (1963) and the mixture space theorem in order to provide a much simpler condition that generates quasi-linear means.

Definition. *Let $\phi : X \to \mathbb{R}$ be strictly increasing and continuous. We say that $I : [0, 1] \times X^2 \to X$ is a* ***ϕ-mixture operation*** *if*

$$I_a(x, y) = \phi^{-1}(a\phi(x) + (1 - a)\phi(y)).$$

Abusing the notation a little bit, let

$$I_a(\boldsymbol{x}, \boldsymbol{y}) = (I_a(x_1, y_1), I_a(x_2, y_2), \ldots, I_a(x_n, y_n)).$$

That is, when applied to vectors, I is a dimention-by-dimension ϕ-mixture operation.

Theorem 4. *Let X be a convex non-singleton subsect of $\mathbb{R}$. Further let $f : X^n \to X$ be* *nondecreasing, diagonally-increasing and continuous,* $\phi : X \to \mathbb{R}$ *be strictly increasing and continuous, and* $I : [0,1] \times X^2 \to X$ *be a* ϕ*-mixture operation. Then* f *is ordinally equivalent to a* ϕ*-mean if and only if*

$$f(\boldsymbol{x}) > f(\boldsymbol{y}) \Longrightarrow f(I_a(\boldsymbol{x}, \boldsymbol{z})) > f(I_a(\boldsymbol{y}, \boldsymbol{z}))$$

$\forall\ \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in X^n$ *and* $a \in (0,1)$.

The theorem states that all ϕ-means boil down to a condition which resembles the independence in von Neumann and Morgenstern (1944) and Herstein and Milnor (1953). We provide the following examples to illustrate this condition.

Arithmetic mean. Let $\phi(x) = x$. Then the condition reduces to

$$\begin{aligned} & f(\boldsymbol{x}) > f(\boldsymbol{y}) \\ \Longrightarrow\quad & f(ax_1 + (1-a)z_1, \ldots, ax_n + (1-a)z_n) > f(ay_1 + (1-a)z_1, \ldots, ay_n + (1-a)z_n) \end{aligned}$$

for any $a \in (0,1)$ and $\boldsymbol{z}$.

Geometric mean. Let $\phi(x) = \ln x$. The condition reduces to

$$f(\boldsymbol{x}) > f(\boldsymbol{y}) \Longrightarrow f(x_1^a z_1^{1-a}, \ldots, x_n^a z_n^{1-a}) > f(y_1^a z_1^{1-a}, \ldots, y_n^a z_n^{1-a})$$

for any $a \in (0,1)$ and $\boldsymbol{z}$. In Appendix B, this condition is applied to prove Theorem 3.

Harmonic mean. Let $\phi(x) = -1/x$. The condition reduces to

$$f(\boldsymbol{x}) > f(\boldsymbol{y}) \Longrightarrow f\left(\frac{1}{\frac{a}{x_1} + \frac{1-a}{z_1}}, \ldots, \frac{1}{\frac{a}{x_n} + \frac{1-a}{z_n}}\right) > f\left(\frac{1}{\frac{a}{y_1} + \frac{1-a}{z_1}}, \ldots, \frac{1}{\frac{a}{y_n} + \frac{1-a}{z_n}}\right)$$

for any $a \in (0,1)$ and $\boldsymbol{z}$. In Appendix C, we apply this condition to provide a microfoundation for the well-known Tversky index of similarity.

In fact, Theorem 4 is the parameterized version of a slight generalization of Anscombe and Aumann (1963). See Appendix D for our version of the Anscombe-Aumann theorem. In their original paper, an act specifies a simple lottery in each state and the mixture operation is the usual weighted arithmetic average. By contrast, our acts are Savage acts: each vector

$(x_1, \ldots, x_n)$ specifies directly a prize in each state. Moreover, our mixture operation I is a general state-by-state mixture operation.

We interpret $I_a(x, y)$ as the certainty equivalent to objectively randomizing between constant acts $(x, x, \ldots, x)$ and $(y, y, \ldots, y)$ with probability a and $1 - a$ respectively. Then the condition in Theorem 4 corresponds to independence with respect to such objective randomization. Gul (1992) also extracts subjective probability from Savage acts in a finite state space but his independence axiom requires no objective randomization.

4.1 Proof of Theorem 4

We prove it using the Anscombe-Aumann Theorem. First of all we need the following claim which says that I is indeed a mixture operation.

Claim 1. *If I is a ϕ-mixture operation, it is a mixture operation. In particular, (i) $I_1(x, y) = x$; (ii) $I_a(x, y) = I_{1-a}(y, x)$; (iii) $I_a(I_b(x, y), y) = I_{ab}(x, y)$.*

Proof. (i) and (ii) are trivial. for (iii)

$$\begin{aligned} I_a (I_b (x, y), y) &= I_a \left(\phi^{-1} (bf (x) + (1 - b) \phi (y)), y\right) \\ &= \phi(a\phi \left(\phi^{-1} (b\phi (x) + (1 - b) \phi (y))\right) + (1 - a) \phi (y)) \\ &= \phi^{-1} (a (b\phi (x) + (1 - b) \phi (y)) + (1 - a) \phi (y)) \\ &= \phi^{-1} (ab\phi (x) + (1 - ab) \phi (y)) = I_{ab} (x, y). \end{aligned}$$

Therefore I is a mixture operation. □

On the state space X^n, define $\boldsymbol{x} \succeq_f \boldsymbol{y} \iff f(\boldsymbol{x}) \geq f(\boldsymbol{y})$. The next lemma shows that $\succeq_f$ satisfies the axioms in Anscombe and Aumann (1963).

Claim 2. *Let $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in X^n$. If $f(\boldsymbol{x}) > f(\boldsymbol{y})$ implies $f(I_a(\boldsymbol{x}, \boldsymbol{z})) > f(I_a(\boldsymbol{y}, \boldsymbol{z}))$ for $a \in (0, 1)$, then*

(i) $\succeq_f$ is a preference relation;

(ii) $\boldsymbol{x} \succ_f \boldsymbol{y}$ *and* $a \in (0,1)$ *implies* $I_a(\boldsymbol{x}, \boldsymbol{z}) \succ_f I_a(\boldsymbol{y}, \boldsymbol{z})$*;*

(iii) $\boldsymbol{x} \succ_f \boldsymbol{y} \succ_f \boldsymbol{z}$ *implies there exists* $a, b \in (0,1)$ *such that* $I_a(\boldsymbol{x}, \boldsymbol{z}) \succ_f \boldsymbol{y} \succ_f I_b(\boldsymbol{x}, \boldsymbol{z})$*;*

(iv) there exist $\boldsymbol{x}, \boldsymbol{y} \in X^n$ *such that* $\boldsymbol{x} \succ_f \boldsymbol{y}$*;*

(v) $(x_i, x_i, \ldots, x_i) \succeq_f (y_i, y_i, \ldots, y_i)$ *for* $i = 1, 2, \ldots, n$ *implies* $\boldsymbol{x} \succeq_f \boldsymbol{y}$*.*

Proof. (i),(ii),(iv),(v) are trivial. (iii) is implied by the continuity of ϕ and f. □

Therefore, by the Anscombe-Aumann Theorem[12], there exists a nonconstant function U and unique $\alpha_i \in [0,1]$ with $\sum_{i=1}^n \alpha_i = 1$ such that the function defined by

$$W(\boldsymbol{x}) = \sum_{i=1}^n \alpha_i U(x_i)$$

represents $\succeq_f$. The utility index U is linear; that is, $U(I_a(x,y)) = aU(x) + (1-a)U(y)$, and unique up to a positive affine tranformation.

Claim 3. *$U(x) = c\phi(x) + d$ for $c > 0$.*

Proof. First we prove that such U is linear.

$$\begin{aligned} U(I_a(x,y)) &= U\left(\phi^{-1}(a\phi(x) + (1-a)\phi(y))\right) \\ &= cf\left(\phi^{-1}(a\phi(x) + (1-a)\phi(y))\right) + d \\ &= aU(x) + (1-a)U(y). \end{aligned}$$

Since ϕ is strictly increasing, it represents $\succeq_f$ within $\{\boldsymbol{x} \in X^n | x_i = x_j \text{ for all } i, j\}$. Clearly ϕ is also linear with respect to I. Therefore by the uniqueness of U we have $U(x) = c\phi(x) + d$ for $c > 0$. □

The last step is to recover ϕ-mean by performing a monotone transformation on $W(\boldsymbol{x})$. We have $W(\boldsymbol{x}) = c\sum_{i=1}^n \alpha_i \phi(x_i) + d$. Because ϕ^{-1} is strictly increasing

$$\phi^{-1}\left(\frac{W(\boldsymbol{x}) - d}{c}\right) = \phi^{-1}\left(\sum_{i=1}^n \alpha_i \phi(x_i)\right)$$

[12] The original paper applies to the case where $X = [0,1]^n$, but with the mixture space theorem it holds true as long as X is a mixture space. Also, we adopt a different monotonicity condition. See Appendix D for a proof.

also represents $\succeq_f$ and we proved our theorem.

5 Conclusion

In this paper, we provided a framework for analyzing a range of well-documented non-Bayesian behaviors including base rate neglect, conjunction fallacy and disjunction fallacy. We assumed that our decision maker mistakenly assesses the similarity of A to B when evaluating the probability of A given B.

Our similarity-based updating procedure allows $A \cup C$ given B to be less likely than A given B if $B \cap C = \emptyset$, simply because the pair of events $A \cup C$ and B differ more from each other. By allowing for this type of similarity-based departure from Bayesian updating, the axioms that we proposed yield an updating formula that is a Cobb-Douglas weighted geometric mean of $\Pr(A|B)$ and $\Pr(B|A)$, where Pr is the correct rational subjective probability and $\Pr(\cdot|\cdot)$ is given by Bayes' rule. That is, we have provided a micro-foundation for behavioral subjects who confuse these two conditional probabilities but have correct unconditional beliefs.

In the proof, we applied the Anscombe and Aumann Theorem to show that the aggregator of the conditional probabilities must be of the Cobb-Douglas form. We showed that this method could be applied to micro-found a large class of means.

References

Aczel, J. (1966): *Lectures on functional equations and their applications*, New York: Academic Press.

Anscombe, F. J. and R. J. Aumann (1963): "A Definition of Subjective Probability," *Ann. Math. Statist.*, 34, 199–205.

Bar-Hillel, M. and E. Neter (1993): "How alike is it versus how likely is it: A disjunction fallacy in probability judgments." *Journal of Personality and Social Psychology*, 65, 1119.

Barberis, N., A. Shleifer, and R. Vishny (1998): "A Model of Investor Sentiment," *Journal of Financial Economics*, 49, 307–343, reprinted in Richard Thaler, ed., Advances in Behavioral Finance Vol. II, Princeton University Press and Russell Sage Foundation, 2005.

Bordalo, P., K. Coffman, N. Gennaioli, and A. Shleifer (2016): "Stereotypes," *The Quarterly Journal of Economics*, 131, 1753–1794.

de Finetti, B. (1931): "Sul Concetto di Media," *Giornale dell' Instituo Italiano degli Attuari*, 2, 369–396.

Debreu, G. (1954): "Representation of a preference ordering by a numerical function," *Decision processes*, 3, 159–165.

DeGroot, M., S. Fienberg, and J. Kadane (1994): *Statistics and the Law*, New York: Wiley.

Fienberg, S. (1989): *The Evolving role of statistical assessments as evidence in the courts*, New York: Springer-Verlag.

Fischhoff, B., P. Slovic, and S. Lichtenstein (1978): "Fault trees: Sensitivity of estimated failure probabilities to problem representation." *Journal of Experimental Psychology: Human Perception and Performance*, 4, 330.

Gennaioli, N. and A. Shleifer (2010): "What Comes to Mind," *Quarterly Journal of Economics*, 125, 1399–1433.

Gilboa, I. and D. Schmeidler (1989): "Maxmin expected utility with non-unique prior," *Journal of Mathematical Economics*, 18, 141–153.

Gul, F. (1992): "Savage's Theorem with a Finite Number of States," *Journal of Economic Theory*, 57, 369–396.

Herstein, I. N. and J. Milnor (1953): "An Axiomatic Approach to Measurable Utility," *Econometrica*, 21, 291–297.

Hong, C. S. (1983): "A Generalization of the Quasilinear Mean with Applications to the Measurement of Income Inequality and Decision Theory Resolving the Allais Paradox," *Econometrica*, 51, 1065–1092.

Kahneman, D. and A. Tversky (1972): "Subjective probability: A judgment of representativeness," *Cognitive psychology*, 3, 430–454.

——— (1974): "Judgment under Uncertainty: Heuristics and Biases," *Science*, 185, 1124–1131.

——— (1982): "Evidential impact of base rates," in *Judgment under uncertainty: Heuristics and biases*, ed. by D. Kahneman, A. Tversky, and P. Slovic, Cambridge: Cambridge University Press, 153–160.

——— (1983): "Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment." *Psychological review*, 90, 293–315.

Kolmogorov, A. (1930): “Sur la Notion de la Moyenne,” *Rendiconti Accademia dei Lincei*, 6, 388–391.

Krantz, D., D. Luce, P. Suppes, and A. Tversky (1999): “Foundations of measurement, Vol. I: Additive and polynomial representations,” .

Kruse, R. L. and J. J. Deely (1969): “Joint Continuity of Monotonic Functions,” *The American Mathematical Monthly*, 76, 74–76.

Mullainathan, S. (2002a): “A Memory-Based Model of Bounded Rationality,” *The Quarterly Journal of Economics*, 117, 735–774.

——— (2002b): “Thinking through Categories,” *Working Paper, MIT.*

Mullainathan, S., J. Schwartzstein, and A. Shleifer (2008): “Coarse Thinking and Persuasion,” *The Quarterly Journal of Economics*, 123, 577–619.

Rabin, M. (2002): “Inference by Believers in the Law of Small Numbers,” *The Quarterly Journal of Economics*, 117, 775–816.

Rabin, M. and J. L. Schrag (1999): “First Impressions Matter: A Model of Confirmatory Bias,” *The Quarterly Journal of Economics*, 114, 37–82.

Savage, L. J. (1954): *The Foundations of Statistics*, New York: Wiley.

Schwartzstein, J. (2014): “Selective Attention and Learning,” *Journal of the European Economic Association*, 12, 1423–1452.

Tribe, L. H. (1971): “Trial by Mathematics: Precision and Ritual in the Legal Process,” *Harvard Law Review*, 84, 1329–1393.

Tversky, A. (1977): “Features of Similarity,” *Psychological Review*, 84, 327–352.

Villegas, C. (1964): “On Qualitative Probability σ-Algebras,” *Ann. Math. Statist.*, 35, 1787–1796.

von Neumann, J. and O. Morgenstern (1944): *Theory of Games and Economic Behavior*, Princeton: Princeton University Press.

Wilson, A. (2014): "Bounded Memory and Biases in Information Processing," *Econometrica*, 82, 2257–2294.

A Proof of Theorem 1

In this section, we present our proof of Theorem 1. We first incorporates results from Savage (1954) and Villegas (1964) to obtain a nonatomic subjective probability. Then, we exploit the nonatomicity and apply a Debreu-type ordinal representation argument to prove the existence of an aggregator and therefore a similarity index.

First of all, we introduce Savage (1954) and Villegas (1964)'s result on subjective probability representation, which we will cite from time to time in the proof. Let $\succeq^*$ be a binary relation defined on a σ-algebra $\mathcal{A}$ of subsets of Ω.

Definition. *$\succeq^*$ is a* ***Savagian qualitative probability*** *on $\mathcal{A}$ if it is a preference relation satisfying the following axioms,*

Axiom S1. $\Omega \succ^* \emptyset$.

Axiom S2. $A \succeq^* \emptyset$.

Axiom S3. If $C \cap (A \cup B) = \emptyset$, then $A \succeq^* B \iff A \cup C \succeq^* B \cup C$.

Axiom S4. $A \succ^* B$ implies there is $A' \subset A$ such that $A \succ^* A' \succ^* B$.

Axiom S5. $A_n \succeq^* B$ and $A_{n+1} \subset A_n$ implies that $\bigcap_n A_n \succeq^* B$.

Savage (1954) shows that if a qualitative probability (i.e. a preference relation defined on a σ-algebra that satisifes Axioms S1-S3) is fine[13] and tight[14], then there exists a unique finitely additive probability representation. Villegas (1964) then identifies S4 as a necessary and sufficient condition for such probability representation to be countably additive. Moreover, he notes that S5 is enough to ensure fineness and tightness under S4.[15]

Lemma A. *(Savage (1954);Villegas (1964)) $\succeq^*$ is a Savagian qualitative probability if and*

[13] $\succeq$ is fine if given any $A \succ \emptyset$ there is a finite partition of Ω, $\{\Omega_n\}$, such that $A \succ \Omega_n$ for all n.

[14] $\succeq$ is tight if $A \sim B$ for all A, B such that $A \cup A' \succeq B$ and $B \cup B' \succeq A$ for all nonnull A', B' with $A \cap A' = B \cap B' = \emptyset$.

[15] In particular, with S1-S5 we have enough structure to partition any event into two equally probable ones so that we are able to assign dyadic probabilities to such partitions. Then a continuity argument derived from S4 completes the subjective probability assignment.

only if there is a unique nonatomic probability μ *such that*

$$A \succeq^* B \Longleftrightarrow \mu(A) \geq \mu(B).$$

Since μ is nonatomic and countably additive, it is also *convex-ranged*; that is, for any $a \in (0,1)$ and A such that $\mu(A) > 0$, there is $B \subset A$ such that $\mu(B) = a\mu(A)$. We will exploit this property throughout the paper to identify suitable events.

Throughout this subsection, it is assumed that $(\Omega, \mathcal{E}, \mathcal{N}, \succeq)$ is a nondegenerate similarity structure that satisfies Axiom 1-6. Since $\mathcal{N}$ is a σ-ideal, $ACB \in \mathcal{E}^2 \backslash \mathcal{N}^2$ if not all of A, B, C are null. Write $\overline{C} \equiv \Omega \backslash C$. Since $\Omega \notin \mathcal{N}$, if $C \in \mathcal{N}$, then $\overline{C} \notin \mathcal{N}$. Also note that by (iii) in the definition of a similarity structure, if $C \in \mathcal{N}$ then $ACB \sim A\emptyset B$. Conversely, if $ACB \sim A\emptyset B$ for some A, B, by nondegeneracy and separability, $C \in \mathcal{N}$.

We first consider the case when maximum similarity is assigned.

Lemma 1. $A, B \in \mathcal{N}$ *implies* $ACB \sim \emptyset C \emptyset$*.*

Proof. If $A, B \in \mathcal{N}$, then $CAB \sim C\emptyset B$ and $\emptyset BC \sim \emptyset\emptyset C$. Therefore by monotonicity we have $ACB \sim \emptyset CB$ and $\emptyset CB \sim \emptyset C \emptyset$ and we are done. □

Lemma 2. *If* $C' \subset C$*, then* $ACB \succeq AC'B, C'AB \succeq CAB$ *and* $ABC' \succeq ABC$*.*

Proof. By Lemma 1 the case when $A, B \in \mathcal{N}$ is trivial. Now suppose $A\emptyset B$ is in our domain. By definition $A(C \backslash C')B \succeq A\emptyset B$. Then Axiom 1 implies that $ACB \succeq AC'B$. Then monotonicty proves the result. □

Lemma 3. $ACB \sim \emptyset C \emptyset$ *implies* $B \in \mathcal{N}$*.*

Proof. Suppose $A \cup B$ are nonnull but $\emptyset C \emptyset \sim ACB$. Lemma 2 implies that $\emptyset C \emptyset \succeq AC\emptyset \succeq ACB \sim \emptyset C \emptyset$. Clearly $C \notin \mathcal{N}$. Then by monotonocity, $A\emptyset C \sim ABC$. Suppose there is A', C' such that $A'BC' \succ A'\emptyset C'$, then since $\emptyset\Omega\emptyset \succ ABC$, by separability, $ABC \succ A\emptyset C$ delivering the desired contradiction. Hence B is null. □

For A such that both A and $\overline{A}$ are nonnull, define $C \succeq_A D$ if $\emptyset CA \succeq \emptyset DA$. Note that $\succeq_A$ is not complete in the space $\mathcal{E}$. However, it is complete and transitive in $\mathcal{E} \cap 2^{\overline{A}}$; that is, measurable subsets of $\overline{A}$. Then by Axiom 1,2 and 3, $\succeq_A$ is a Savagian qualitative probability. Then by Lemma A there exists a unique nonatomic probability μ_A such that $\mu_A(C) \geq \mu_A(D)$ if and only if $C \succeq_A D$. Note that $\mu_A(B) = 0$ if only if $B \in \mathcal{N}$ and $B \subset A$.

Now, define $\succeq^*$ as follows. On $\mathcal{R} \equiv \{A \in \mathcal{E} | \overline{A} \notin \mathcal{N}\}$, define $C \succeq^* D$ if there exist nonnull events A, A', B with $A \sim_B A'$ such that $\emptyset CA \succeq \emptyset DA'$. Then, add $\mathcal{E} \backslash \mathcal{R}$ as the single highest equivalence class.

Next, we prove that $\succeq^*$ is a Savagian qualitative probability. We proceed by first introducing the following two lemmas.

Lemma 4. *If $ACB \sim ADB$ and B is nonnull, then $A'CB' \sim A'DB', CA'B' \sim DA'B'$ and $A'B'C \sim A'B'D$.*

Proof. We prove that $A'CB' \sim A'DB'$ and the rest is given by monotonicity. Since B is nonnull we know that $\emptyset\Omega\emptyset \succ ACB$. Therefore if $A'CB' \succ A'DB'$, then by separability we have $ACB \succ ADB$, a contradiction. Hence $A'DB' \succeq A'CB'$. Similarly, $A'CB' \succeq A'DB'$. □

Lemma 5. *For pairwise disjoint $A, B, C \in \mathcal{R} \backslash \mathcal{N}$, let $A_n \subset A, B_n \subset B, C_n \subset C$ and $\mu_{\overline{A}}(A_n) = \mu_{\overline{B}}(B_n) = \mu_{\overline{C}}(C_n) = 2^{-n}$, then $ACB \sim A_nC_nB_n$, $\emptyset CB \sim \emptyset C_nB_n$ and $AC\emptyset \sim A_nC_n\emptyset$.*

Proof. Consider any partitions $\{A_n^m\}, \{B_n^m\}, \{C_n^m\}$ of respectively A, B, C such that for all m, $\mu_{\overline{A}}(A_n^m) = \mu_{\overline{B}}(B_n^m) = \mu_{\overline{C}}(C_n^m) = 2^{-n}$. By nonatomicity such partitions always exist. By the previous lemma, $\emptyset C_n^m B_n^m \sim \emptyset C_n^{m'} B_n^m \sim \emptyset C_n^{m'} B_n^{m'}$. Similarly $A_n^m C_n^m B_n^m \sim A_n^m C_n^{m'} B_n^m \sim A_n^{m'} C_n^{m'} B_n^m \sim A_n^{m'} C_n^{m'} B_n^{m'}$. Inductively applying scale invariance gives the result. □

Lemma 5 states that we could always shrink the events proportionately without changing the similarity assessment. This is essential since if C and D are so large that $\overline{C \cup D}$ is null, it is not possible to find A, B such that $ACB \nsim ADB$. In that case we cannot extract the qualitative probability preference between C and D. Lemma 5, however, gives us the luxury to shrink C and D so that $\overline{C \cup D}$ is no longer null.

The next lemma shows that $\succeq^*$ is well-defined.

Lemma 6. *Let $C, D \in \mathcal{R}$. If $C \succeq^* D$, then for any nonnull events $A \subset \overline{C}, A' \subset \overline{D}, B$ with $A \sim_B A'$, $\emptyset CA \succeq \emptyset DA'$.*

Proof. By definition there are nonnull events $\widehat{A}, \widehat{A}', \widehat{B}$ with $\widehat{A} \sim_{\widehat{B}} \widehat{A}'$ such that $\emptyset C\widehat{A} \succeq \emptyset D\widehat{A}'$. If D is null the result is trivial. If D is nonnull so is C. Let $\{C_n\}$ and $\{D_n\}$ be sequences of sets such that $C_0 = C, D_0 = D$, $C_{n+1} \subset C_n, D_{n+1} \subset D_n$ and

$$\frac{\mu_{\overline{C}}(C_n)}{\mu_{\overline{C}}(C_{n+1})} = \frac{\mu_{\overline{D}}(D_n)}{\mu_{\overline{C}}(D_{n+1})} = 2.$$

If $\widehat{A} \cap D$ is null, in other words $\mu_{\widehat{B}}(\widehat{A}) = \mu_{\widehat{B}}(\widehat{A} \backslash D)$, then

$$\emptyset C\widehat{A} \sim \emptyset C(\widehat{A} \backslash D) \succeq \emptyset D\widehat{A}' \sim \emptyset D(\widehat{A} \backslash D).$$

Let $n = 0$ and $\widehat{A}^* = \widehat{A} \backslash D$. If $\widehat{A} \cap D$ is nonnull, pick n large enough such that there is $\widehat{A}^* \subset (\widehat{A} \cap D) \backslash D_n, \widehat{A}'^* \subset \widehat{A}'$ such that

$$\mu_{\overline{\widehat{A}}}(\widehat{A}^*) = \mu_{\overline{\widehat{A}'}}(\widehat{A}'^*) = \frac{1}{2^n}.$$

This is possible by the convexed-rangeness of the probabilities extracted above. Hence we have $\emptyset C_n\widehat{A}^* \succeq \emptyset D_n\widehat{A}'^* \sim \emptyset D_n\widehat{A}^*$. Using the same procedure we obtain $\emptyset C_m A^* \sim \emptyset CA$ and $\emptyset D_m A^* \sim \emptyset DA'$. Assume that $m = n$. This is wlog since we could always shrink the larger further with the preference unchanged. It is clear by separability that $\emptyset C_m A^* \succeq \emptyset D_m A^*$ and we complete the proof. □

With the above lemma, $C \succ^* D$ if there exist A, A', B nonnull with $A \sim_B A'$ such that $\emptyset CA \succ \emptyset DA'$. Moreover, given the procedure in the proof, if $C \succeq^* D$ we could wlog assume that there is nonnull A such that $\emptyset CA \succeq \emptyset DA$. Then, we show that $\succeq^*$ is complete and transitive.

Lemma 7. $\succeq^*$ *is complete and transitive.*

Proof. It suffices to show that $\succeq^*$ is complete and transitive in $\mathcal{R}$.

(i) Completeness. It suffices to show that for any C, D in $\mathcal{R}$ there exists $A \subset \overline{C}, A' \subset \overline{D}$ and $B \subset \overline{A \cup A'}$ such that $A \sim_B A'$.

Case 1: $C \cup D$ is null. It is clear that C, D are both null. Then $\emptyset CA \sim \emptyset DA'$ for any A, A'. Consdier $\widetilde{A}, \widetilde{B}, \widetilde{C}$ such that $\emptyset\Omega\emptyset \succ \widetilde{A}\widetilde{C}\widetilde{B} \succ \widetilde{A}\emptyset\widetilde{B}$. Let $E = \widetilde{C} \cap \overline{C \cup D}$. It is easy to see that E and $\overline{E}$ are both nonnull. Therefore $E \sim_{\overline{E}} E$ and we are done.

Case 2: $C \cup D$ and $\overline{C \cup D}$ are both nonnull. Then $\overline{C \cup D} \sim_{C \cup D} \overline{C \cup D}$ and we are done.

Case 3: $C \cup D$ is nonnull but $\overline{C \cup D}$ is null. Clearly $C \backslash D$ and $D \backslash C$ are both nonnull since $C, D \in \mathcal{R}$. Pick $E \subset C\backslash D, E' \subset D \backslash C$ such that $\mu_{\overline{C\backslash D}}(E) = \mu_{\overline{D\backslash C}}(E') = \frac{1}{2}$. Clearly $C\backslash(D \cup E)$ is nonnull since $\mu_{\overline{C\backslash D}}(C\backslash(D\cup E)) = \frac{1}{2}$. If $E \sim_{C\backslash(D\cup E)} E'$ we are done. If $E \succ_{C\backslash(D\cup E)} E'$ by the convex-rangedness of $\mu_{C\backslash(D\cup E)}$ there is $E'' \subset E$ such that $E'' \sim_{C\backslash(D\cup E)} E'$.

(ii) Transitivity. Let $C \succeq^* D$ and $D \succeq^* E$. The case when at least one of them is null is trivial. So assume that C, D, E are all nonnull. By definition there are $A \sim_B A', \widehat{A} \sim_{\widehat{B}} \widehat{A}'$ such that $\emptyset CA \succeq \emptyset DA'$ and $\emptyset D\widehat{A} \succeq \emptyset E\widehat{A}'$. By the procedure in the previous lemma we could wlog assume that $A = A'$ and $\widehat{A} = \widehat{A}'$. Also assume that $\mu_{\overline{D}}(A) > \mu_{\overline{D}}(\widehat{A})$. Pick $A^* \subset A$ such that $\mu_{\overline{D}}(A^*) = \mu_{\overline{D}}(\widehat{A})$. Then by Lemma 4 and 6,

$$\emptyset CA^* \succeq \emptyset DA^* \sim \emptyset D\widehat{A} \succeq \emptyset E\widehat{A}$$

and we are done. $\square$

The next lemma proves that $\succeq^*$ is a qualitative probability defined on $\mathcal{E}$.

Lemma 8. *(i)* $\Omega \succ^* \emptyset$*; (ii)* $C \succeq^* \emptyset$*; (iii)* $(C \cup D) \cap E = \emptyset$ *implies* $[C \succeq^* D \Longleftrightarrow C \cup E \succeq^* D \cup E]$.

Proof. (i) and (ii) are by construction. (iii) If $C \cup E, D \cup E \in \mathcal{R}$, we appeal to the procedure in the proof of Lemma 6 and the additivity axiom. If $C \notin \mathcal{R}$ or $D \notin \mathcal{R}$ both directions are trivial. Now assume that $C \in \mathcal{R}$ and $D \in \mathcal{R}$. Suppose $C \cup E \sim^* \Omega$. The "$\Longrightarrow$" direction is trivial. For the other direction, since $D \backslash C$ is null, we have $C \succeq^* C \cap D \sim^* D$. Suppose $C \cup E \in \mathcal{R}$. The "$\Longleftarrow$" direction is clear since $D \cup E \in \mathcal{R}$ and it reduces to the first case that we have considered. For the other direction, we show that if $C \succeq^* D$ and $C \cup E \in \mathcal{R}$ it must also be that $D \cup E \in \mathcal{R}$. Suppose $D \cup E \sim^* \Omega$. We know that $C \backslash (D \cup E) = C \backslash D$ is null. Also, since $C \cup E \in \mathcal{R}$, $D \backslash (C \cup E) = D \backslash C$ is nonnull. Then we must have $D \succ^* C$ delivering the desired contradiction. □

Then, we proceed by showing that $\succeq^*$ is indeed a Savagian qualitative probability.

Lemma 9. *(i) $C \succ^* D$ implies there exist $C' \subset C$ such that $C \succ^* C' \succ^* D$.*

(ii) $C_n \succeq^* D$ and $C_{n+1} \subset C_n$ for all n implies that $\bigcap_n C_n \succeq^* D$.

Proof. (i) If $C \in \mathcal{R}$, it is implied by the nonatomicity axiom. If $C \notin \mathcal{R}$, then by construction $D \in \mathcal{R}$. If D is null, set $C' = C \cap \widetilde{C}$. Since $\widetilde{C}$ is nonnull and $\overline{C}$ is null, C' must be nonnull. Moreover since $\widetilde{C} \in \mathcal{R}$, $\overline{C'}$ cannot be null. Hence $C \succ^* C' \succ^* D$. Now suppose D is nonnull, then $C \cap D$ is nonnull. Since $D \backslash C$ and $\overline{D} \backslash C$ are null we have

$$\emptyset D \overline{D} \sim \emptyset (C \cap D) \overline{D} \sim \emptyset (C \cap D) \left(C \cap \overline{D}\right).$$

Pick $E \subset C \cap \overline{D}$ such that E and $C \cap \overline{D} \backslash E$ are both nonnull. Let $C' = (C \cap D) \cup E$ and we are done.

(ii) It suffices to prove that if $C_n \in \mathcal{E} \backslash \mathcal{R}$ and $C_{n+1} \subset C_n$, then $\bigcap_n C_n \in \mathcal{E} \backslash \mathcal{R}$. It suffices to prove that if D_n are null and $D_n \subset D_{n+1}$, $\bigcup_n D_n$ is null. This is directly given by the fact that $\mathcal{N}$ is a σ-ideal. □

Hence, by Lemma A, there exists a unique nonatomic probability μ defined on $\mathcal{E}$ such that $\mu(C) \geq \mu(D)$ if and only if $C \succeq^* D$. Clearly, $\mu(A) = 0$ if and only if $A \in \mathcal{N}$.

Lemma 10. *$ACB \sim A'C'B'$ if $\mu(A) = \mu(A'), \mu(B) = \mu(B')$ and $\mu(C) = \mu(C')$.*

Proof. Assume that A, B, C are all nonnull. The cases when any of them is null is similar. By Lemma 5, wlog we could assume that A, B, C, A', B', C' are mutually exclusive. Then we know that

$$ACB \sim A'CB \sim A'C'B \sim A'C'B'$$

by monotonicity and Lemma 4. □

Lemma 10 implies that if $\succeq$ has a representation, then the representation can only depends on $\mu(A)$, $\mu(B)$ and $\mu(C)$. Since we have abandoned linearity, we are not able to apply the techniques from linear conjoint measurement. Instead, a Debreuian argument is needed for identification of the representation. In particular, we need to show that there exists a countable order dense set. We define order denseness as follows.

Definition. *For any binary relation $\succeq$ on X, we say that the subset Y is $\succeq$-order dense if for any $x, y \in X$ such that $x \succ y$, there exists $z \in Y$ such that $x \succeq z \succeq y$.*

Then we introduce Debreu (1954)'s result as a lemma. Then we show that a representation for our similarity rankings exists.

Lemma B. *(Debreu (1954)) For any set X and binary relation $\succeq$ on X, there exists a function U that represents $\succeq$ if and only if $\succeq$ is a preference relation and X has a countable $\succeq$-order dense subset.*

Lemma 11. *$\mathcal{E}^2 \backslash \mathcal{N}^2$ has a countable $\succeq$-order dense subset.*

Proof. By the convex-rangedness of μ, we construct the order dense subset as follows. Pick A_0, B_0, C_0 mutually exclusive such that $A_0 \cup B_0 \cup C_0 = \Omega$ and $\mu(A) = \mu(B) = \mu(C) = 1/3$. Let $\{A_{mn}\}, \{B_{mn}\}, \{C_{mn}\}$ be subsets of respectively A, B, C, such that

$$\mu(A_{mn}) = \mu(B_{mn}) = \mu(C_{mn}) = \frac{m}{n}$$

for $m = 0, 1, 2, \ldots, \lfloor n/3 \rfloor$ and $n = 1, 2, 3, \ldots$

Next, we show that $\{ACB | A \in \{A_{mn}\}, B \in \{B_{mn}\}, A \in \{C_{mn}\}\}$ is a $\succeq$-order dense subset.

Let $A'C'B' \succ ACB$. Wlog assume that $\mu(A), \mu(B), \mu(C), \mu(A'), \mu(B'), \mu(C') \leq 1/3$. If $\emptyset\Omega\emptyset \sim A'C'B'$ we have $A'C'B' \sim A_{01}C_{13}B_{01} \succ ACB$ and we are done.

Suppose $\emptyset\Omega\emptyset \succ A'C'B'$, then by the nonatomicity axiom, there is $\widehat{C} \subset C'$ such that $A'C'B' \succ A'\widehat{C}B' \succ ACB$. Pick m_1, n_1 such that

$$\mu(C') > \frac{m_1}{n_1} > \mu(\widehat{C}).$$

Pick $C^* \subset C'$ such that $\mu(C^*) = m_1/n_1$. It follows that

$$A'C'B' \succ A'C^*B' \succ A'\widehat{C}B'.$$

If $\emptyset C^*B' \sim A'C^*B'$ then set $A^* = \emptyset$ and $m_2 = 0, n_2 = 1$. Suppose $\emptyset C^*B' \succ A'C^*B'$, then by nonatomicity, there is $\widehat{A} \subset A'$ such that

$$A'C'B' \succ \widehat{A}C^*B' \succ A'C^*B'.$$

Pick m_2, n_2 such that

$$\mu(A') > \frac{m_2}{n_2} > \mu(\widehat{A})$$

and $A^* \subset A'$ such that $\mu(A^*) = m_2/n_2$. Then by monotonicity

$$A'C'B' \succ \widehat{A}C^*B' \succeq A^*C^*B' \succeq A'C^*B' \succ ACB.$$

If $A^*C^*\emptyset \sim A^*C^*B'$ then set $B^* = \emptyset$ and $m_3 = 0, n_3 = 1$. Suppose $A^*C^*\emptyset \succ A^*C^*B'$, again by the nonatomicity axiom there is $\widehat{B} \subset B'$ such that

$$A'C'B' \succ A^*C^*\widehat{B} \succ A^*C^*B' \succ ACB.$$

Then similarly pick $\mu(B^*) = m_3/n_3$ such that

$$A'C'B' \succ A^*C^*\widehat{B} \succeq A^*C^*B^* \succeq A^*C^*B' \succ ACB.$$

Therefore

$$A'C'B' \succ A^*C^*B^* \succ ACB$$

and

$$A^*C^*B^* \sim A_{m_2n_2}C_{m_1n_1}B_{m_3n_3}.$$

Thus, we have proved the result. $\square$

Hence $\succeq$ has a numerical representation $S(A,B)$. By Lemma 10 we know that $S(A,B) = h(\mu(A\backslash B), \mu(A\cap B), \mu(B\backslash A))$. Then we characterize how h should behave.

Definition. *A function $f: D \subset \mathbb{R}^n \to \mathbb{R}$ is separately continuous if for any fixed n-tuple $(x_1, x_2, \ldots, x_n) \in D$ and for every $i \in \{1, 2, \ldots, n\}$, the mapping $t \mapsto f(x_1, \ldots, x_{i-1}, t, x_{i+1}, \ldots, x_n)$ is continuous for all t such that $(x_1, \ldots, x_{i-1}, t, x_{i+1}, \ldots, x_n) \in D$.*

Lemma C. *(Kruse and Deely (1969)) Let $f(x_1, x_2, ..., x_{n-1}, y)$ be a real-valued function defined on an open set G in $\mathbb{R}^n, n \geq 2$. Suppose f is separately continuous and is monotone in each x_i separately, $1 \leq i \leq n-1$. Then f is continuous on G.*

By Lemma C the following characterizations of h is true. Let Δ^n be the n-simplex in $\mathbb{R}^n$.

Lemma 12. *There exists a function $h: \Delta^3\backslash\{(0,0,0)\} \longrightarrow [0,1]$ such that*

$$ACB \succeq A'C'B' \Longleftrightarrow h(\mu(A), \mu(C), \mu(B)) \geq h(\mu(A'), \mu(C'), \mu(B')).$$

where

(i) $h(a,0,b) = 0$, $h(0,c,0) = 1$;

(ii) $h(a,\cdot,b)$ is nondecreasing;

(iii) $h(\cdot,c,b)$ and $h(a,c,\cdot)$ are nondecreasing;

(iv) h is continuous at (a,c,b) if $0 < c < 1$;

(v) $h(a,c,b) = h(ka,kc,kb)$ for $k \in (0,1)$ if $0 < c < 1$.

Proof. (i) is simple rescaling. (ii) is given by additivity. (iii) is given by monotonicity.

Next we prove (iv) by construction. By (ii), fixed any a, b, $h(a,\cdot,b)$ is monotone and hence continuous almost everywhere. Let $\{x_n\}$ be the discontinuity points of $h(a,\cdot,b)$. we prove that there cannot exist $(a', b', c') \in \Delta^3$ such that $h(a, x_n^+, b) \geq h(a', c', b') \geq h(a, x_n^-, b)$ with $h(a', b', c') \neq h(a, x_n, b)$. Consider first if

$$h\left(a, x_n^+, b\right) \geq h(a', c', b') > h(a, x_n, b) \geq h\left(a, x_n^-, b\right).$$

Pick any decreasing sequence $\{x_{n,m}\} \in [0, 1-a-b]$ such that $x_{n,m} \to x_n$ as $m \to \infty$. There are A, B, A', B', C' such that $\mu(A) = a, \mu(B) = b, , \mu(A') = a', \mu(B') = b'$ and $\mu(C') = c'$. Also pick $\{C_m\}$ such that $C_{m+1} \subset C_m$ and $\mu(C_m) = x_{n,m}$. Hence

$$\lim_{m \to \infty} \mu(C_m) = \mu\left(\bigcap_m C_m\right) = x_n.$$

Therefore $AC_mB \succeq A'C'B'$ however $A'C'B' \succ A(\bigcap_m C_m)B$, a contradiction to Axiom 6. For the other case, that is when

$$h\left(a, x_n^+, b\right) \geq h(a, x_n, b) > h(a', c', b') \geq h\left(a, x_n^-, b\right),$$

it suffices to prove $A'C'B' \succeq AC_mB$ and $C_m \subset C_{m+1}$ implies $A'C'B' \succeq A(\bigcup_m C_m)B$. Suppose $A(\bigcup_m C_m)B \succ A'C'B'$. There exists $C \subset \bigcup_m C_m$ such that $A(\bigcup_m C_m)B \succ ACB \succ A'C'B'$. Therefore there is m^* such that

$$\mu\left(\bigcup_m C_m\right) > \mu(C_{m^*}) > \mu(C).$$

By (ii)

$$A(\bigcup_m C_m)B \succeq AC_{m^*}B \succeq ACB \succ A'C'B'$$

which is a contradiction.

By (iii), $h(\cdot, c, b)$ is monotone therefore is continuous almost everywhere. Let $\{y_n\}$ be the discontinuity points of $h(\cdot, c, b)$. The next step is prove that there cannot exist $(a', b', c') \in \Delta^3$ such that $h(y_n^-, c, b) \geq h(a', c', b') \geq h(y_n^+, c, b)$ with $h(a', c', b') \neq h(y_n, c, b)$. Suppose

$$h\left(y_n^-, c, b\right) \geq h(a', c', b') > h(y_n, c, b) \geq h\left(y_n^+, c, b\right).$$

By the nonatomicty axiom there is $\widehat{y}_n < y_n$ such that

$$h\left(a', c', b'\right) > h\left(\widehat{y}_n, c, b\right) > h\left(y_n, c, b\right)$$

which is a contradiction. Suppose

$$h\left(y_n^-, c, b\right) \geq h\left(y_n, c, b\right) > h\left(a', c', b'\right) \geq h\left(y_n^+, c, b\right).$$

Pick any increasing sequence $\{y_{n,m}\} \in [0, 1-c-b]$ such that $y_{n,m} \to y_n$ as $m \to \infty$. There are B, C, A', B', C' such that $\mu\left(B\right) = b, \mu\left(C\right) = c, , \mu\left(A'\right) = a', \mu\left(B'\right) = b'$ and $\mu\left(C'\right) = c'$. Also pick $\{A_m\}$ such that $A_{m+1} \subset A_m$ and $\mu\left(A_m\right) = y_{n,m}$. It follows that

$$(\bigcap_m A_m)CB \succ A'C'B' \succ A_m CB.$$

which is a contradiction to countable-addivity. Similar results obtain for $h\left(a, c, \cdot\right)$.

Consider $h\left(0, \frac{1}{2}, \cdot\right)$ and $h\left(0, \cdot, \frac{1}{2}\right)$. By the arguments above, these two functions capture all the discontinuity of range in any $h(\cdot, c, b)$, $h(a, \cdot, b)$ or $h(a, c, \cdot)$. Let $\{a_n\}$ be the set of discontinuity points of $h\left(0, \frac{1}{2}, \cdot\right)$ and $\{b_n\}$ be that of $h\left(0, \cdot, \frac{1}{2}\right)$. Then construct h^* as follows. Let

$$h^1\left(a, c, b\right) = \begin{cases} h\left(0, \frac{1}{2}, a_n^+\right) & \text{if } h\left(a, c, b\right) = h\left(0, \frac{1}{2}, a_n\right) \text{ for some } n, \\ h\left(0, b_n^-, \frac{1}{2}\right) & \text{if } h\left(a, c, b\right) = h\left(0, b_n, \frac{1}{2}\right) \text{ for some } n, \\ h\left(a, c, b\right) & \text{otherwise.} \end{cases}$$

Then let

$$v_1(a, c, b) = \sum_{k:\ h^1(a,c,b) > h^1\left(0, \frac{1}{2}, a_k^-\right)} \left[h^1\left(0, \frac{1}{2}, a_k^-\right) - h^1\left(0, \frac{1}{2}, a_k^+\right)\right].$$

That is, the cumulative discontinuity of $h(0, \frac{1}{2}, \cdot)$ within the range $h(0, \frac{1}{2}, \frac{1}{2})$ to $h(a, c, b)$. Also let,

$$v_2(a, c, b) = \sum_{k: h^1(a,c,b) > h^1\left(0, b_k^+, \frac{1}{2}\right)} \left[h^1\left(0, b_k^+, \frac{1}{2}\right) - h^1\left(0, b_k^-, \frac{1}{2}\right)\right].$$

That is, the cumulative discontinuity of $h(0,\cdot,\frac{1}{2})$ within the range from $h(0,0,\frac{1}{2})$ to $h(a,c,b)$.

$$h^2(a,c,b) = \begin{cases} h^1(a,c,b) - v_2(a,c,b) & \text{if } h^1(a,c,b) \le h^1\left(0,\frac{1}{2},\frac{1}{2}\right), \\ h^1(a,c,b) - v_1(a,c,b) - v_2(a,c,b) & \text{if } h^1(a,c,b) > h^1\left(0,\frac{1}{2},\frac{1}{2}\right). \end{cases}$$

Then rescale by

$$h^*(a,b,c) = \frac{h^2(a,b,c) - h^2(1,0,0)}{h^2(0,1,0) - h^2(1,0,0)}.$$

Basically h^* is produced from h by squeezing out the discontinuity points and then rescaling. By the argument above the resulting h^* is separately continuous on $\Delta^3 \backslash \{(0,0,0)\}$. Extend h^* by setting

$$h^*(a,c,b) = h^*(\max\{a,0\}, c, \max\{b,0\})$$

to $G \equiv \{(a,c,b) \in \mathbb{R}^3 | 0 < c < 1,\ a+b+c < 1,\ b+c < 1, \text{ and } a+c < 1\}$. It is clear that h^* is monotone and separately continuous on G. Therefore, by Lemma C, h^* is continuous on G. It then suffices to prove that h^* is also continuous on $\{(a,c,b) \in \Delta^3 | a+b+c=1\}$. This is true since $h^*(a,c,b) = h^*(\frac{a}{2},\frac{b}{2},\frac{c}{2})$ by Lemma 5.

By Lemma 5 and scale invariance we have that for all n and $m = 1, 2, \ldots, 2^n$.

$$h(a,c,b) = h\left(\frac{m}{2^n}a, \frac{m}{2^n}c, \frac{m}{2^n}b\right).$$

By continuity (v) follows. □

Let $f : [0,1]^2 \longrightarrow [0,1]$ be defined as follows. For $(x,y) \in (0,1]^2$ let

$$f(x,y) = h\left(\frac{\frac{1}{y}-1}{2(\frac{1}{x}+\frac{1}{y}-1)}, \frac{1}{2(\frac{1}{x}+\frac{1}{y}-1)}, \frac{\frac{1}{x}-1}{2(\frac{1}{x}+\frac{1}{y}-1)}\right).$$

Let $f(0,0)=0$, $f(x,0) = \lim_{y\to 0^+} f(x,y)$ and $f(0,y) = \lim_{x\to 0^+} f(x,y)$. Clearly f is continuous on $(0,1]^2$ since h is continuous at (a,c,b) such that $0 < c < 1$.

Lemma 13. *For $a+c>0$ and $b+c>0$, $h(a,c,b) = f(\frac{c}{b+c}, \frac{c}{a+c})$.*

Proof. If $c = 0$, then $a, b > 0$. It follows that

$$f\left(\frac{c}{b+c}, \frac{c}{a+c}\right) = 0 = h\left(a, c, b\right).$$

If $0 < c < 1$ then

$$f\left(\frac{c}{b+c}, \frac{c}{a+c}\right) = h\left(\frac{a}{2(a+c+b)}, \frac{c}{2(a+c+b)}, \frac{b}{2(a+c+b)}\right) = h(a, c, b)$$

by Lemma 12. If $c = 1$ and therefore $a = b = 0$, then $f(1, 1) = h(0, 1, 0) = h(0, c, 0)$. $\square$

Let $\mu(A|B)$ be given by Bayes' rule if $\mu(B) > 0$. For expositional purpose, let $\mu(A|B) = 0$ if $\mu(B) = 0$. Therefore, it is easy to see that $S(A, B) = f(\mu(A|B), \mu(B|A))$ represents $\succeq$. The fact that f is nondecreasing is given by the properties of h. It suffices to prove that f is strictly increasing in the first argument, or equivalently, that $h(a, c, \cdot)$ is strictly decreasing for $c > 0$. Due to additivity $h(a, \cdot, c)$ is strictly increasing for $c > 0$. Then monotonicity finishes the argument.

The "if" part of Theorem 1 is standard and therefore omitted.

B Proof of Theorem 3

We only have to prove that f has the Cobb-Douglas form on $(0,1]^2$ since $f(0,0)=0$ and it is simply impossible to have $\mu(A|B)>0$ and $\mu(B|A)=0$ for any A,B. Therefore, in the proof we will assume that the arguments of f are all within $(0,1]$. Lemma 14 is a direct implication of the robustness condition.

Lemma 14. *$f(p_1,p_2)\geq f(q_1,q_2)$ implies $f(rp_1,p_2)\geq f(rq_1,q_2)$ and $f(p_1,rp_2)\geq f(q_1,rq_2)$.*

Proof. First we prove that for any $(p_1,p_2),(q_1,q_2)$ there are A,A',B such that

$$\mu(A|B)=p_1,\quad \mu(B|A)=p_2,\quad \mu(A'|B)=q_1 \text{ and } \mu(B|A')=q_2.$$

It is equivalent to picking

$$\frac{c}{b+c}=p_1,\quad \frac{c}{a+c}=p_2,\quad \frac{c'}{b'+c'}=q_1,\quad \frac{c'}{a'+c'}=q_2 \text{ and } b'+c'=b+c.$$

Let

$$a=\left(\frac{1}{p_2}-1\right)p_1x,\quad b=(1-p_1)x,\quad c=p_1x,$$
$$a'=\left(\frac{1}{q_2}-1\right)q_1x,\quad b'=(1-q_1)x,\quad c'=q_1x,$$

and pick a positive x small enough such that both $a+b+c$ and $a'+b'+c'$ are not larger than 1. Now we prove for any such A,A',B there is $C\perp\sigma(A,A',B)$ such that $\mu(C)=r$. Let $\{A_1,...,A_n\}$ be the finest partition of Ω which is contained in $\sigma(A,A',B)$. Pick $C_k\subset A_k$ such that $\mu(C_k)=r\mu(A_k)$. Let $C=\bigcup_{k=1}^n C_k$ and we are done. Then it is clear that robustness implies the lemma. □

With the above lemma, the following claim is true.

Lemma 15. *For $a\in(0,1)$, $f(p_1,p_2)>f(q_1,q_2)$ implies $f(p_1^a,p_2^a)>f(q_1^a,q_2^a)$.*

Proof. By Lemma 14, $f(p_1,p_2)>f(q_1,q_2)$ implies that

$$f\left(p_1^2,p_2^2\right)\geq f(p_1q_1,p_2q_2)\geq f\left(q_1^2,q_2^2\right).$$

In fact for any n,

$$f\left(p_1^n, p_2^n\right) \geq f\left(p_1^{n-1}q_1, p_2^{n-1}q_2\right) \geq f\left(p_1^{n-2}q_1^2, p_2^{n-2}q_2^2\right) \geq \cdots \geq f\left(p_1q_1^{n-1}, p_2q_2^{n-1}\right) \geq f\left(q_1^n, q_2^n\right).$$

For any m, it must be the case that

$$f\left(p_1^{\frac{1}{m}}, p_2^{\frac{1}{m}}\right) > f\left(q_1^{\frac{1}{m}}, q_2^{\frac{1}{m}}\right)$$

since if otherwise inductively we would have that $f\left(p_1, p_2\right) \leq f\left(q_1, q_2\right)$, which is a contradiction. Since f is monotone and continuous on $(0,1]^2$, there are $r > s$ such that

$$f\left(p_1^{\frac{1}{m}}, p_2^{\frac{1}{m}}\right) > f(r,r) > f(s,s) > f\left(q_1^{\frac{1}{m}}, q_2^{\frac{1}{m}}\right)$$

Combining the two claims, for any m, n with $n > 0$

$$f\left(p_1, p_2\right) > f\left(q_1, q_2\right) \Longrightarrow f\left(p_1^{\frac{n}{m}}, p_2^{\frac{n}{m}}\right) \geq f(r^n, r^n) > f(s^n, s^n) \geq f\left(q_1^{\frac{n}{m}}, q_2^{\frac{n}{m}}\right).$$

Pick $t > u$ such that

$$f\left(p_1, p_2\right) > f\left(t, t\right) > f\left(u, u\right) > f\left(q_1, q_2\right).$$

It follows that

$$f\left(p_1^{\frac{n}{m}}, p_2^{\frac{n}{m}}\right) > f\left(t^{\frac{n}{m}}, t^{\frac{n}{m}}\right) > f\left(u^{\frac{n}{m}}, u^{\frac{n}{m}}\right) > f\left(q_1^{\frac{n}{m}}, q_2^{\frac{n}{m}}\right).$$

Let $\frac{n}{m} \to \alpha \in (0,1)$ we have that

$$f\left(p_1^{\alpha}, p_2^{\alpha}\right) \geq f\left(t^{\alpha}, t^{\alpha}\right) > f\left(u^{\alpha}, u^{\alpha}\right) \geq f\left(q_1^{\alpha}, q_2^{\alpha}\right)$$

where the central inequality is strict due to the fact that f is strictly increasing in its first argument. □

With Theorem 4, it suffices to prove the following lemma.

Lemma 16. *$f(p_1, p_2) > f(q_1, q_2)$ and $a \in (0,1)$ implies* $f\left(p_1^a r_1^{1-a}, p_2^a r_2^{1-a}\right) > f\left(q_1^a r_1^{1-a}, q_2^a r_2^{1-a}\right).$

Proof. By the previous lemma

$$f(p_1, p_2) > f(q_1, q_2) \Longrightarrow f(p_1^a, p_2^a) > f(q_1^a, q_2^a)$$

Pick $r > s$ such that

$$f(p_1^a, p_2^a) > f(r, r) > f(s, s) > f(q_1^a, q_2^a).$$

Then by Lemma 14

$$\begin{aligned} & f\left(p_1^a r_1^{1-a}, p_2^a\right) \geq f(r r_1^{1-a}, r) > f(s r_1^{1-a}, s) \geq f\left(q_1^a r_1^{1-a}, q_2^a\right) \\ \Longrightarrow \quad & f\left(p_1^a r_1^{1-a}, p_2^a r_2^{1-a}\right) \geq f(r r_1^{1-a}, r r_2^{1-a}) > f(s r_1^{1-a}, s r_2^{1-a}) \geq f\left(q_1^a r_1^{1-a}, q_2^a r_2^{1-a}\right) \end{aligned}$$

where the central inequality is strict due to the fact that f is strictly increasing in its first argument. □

The rest is established by Theorem 4. Clearly $\alpha > 0$ since f is strictly increasing in its first argument.

C Recovering the Tversky Index

Although the Tversky Index is popular in the psychology literature, even Tversky himself has not provided an axiomatic foundation. In this section, we utilize again the Anscombe-Aumann framework to achieve this goal. Consider an alternative definition of robustness.

Definition. *$\succeq$ is said to be **robust*** if for any nonnull C and mutually exclusive $\widehat{A}, \widehat{B}, \widehat{C}$ such that $(A \cup B \cup C) \cap (\widehat{A} \cup \widehat{B} \cup \widehat{C}) = \emptyset$,*

$$ACB \succ A'CB' \Longrightarrow (A \cup \widehat{A})(C \cup \widehat{C})(B \cup \widehat{B}) \succ (A' \cup \widehat{A})(C \cup \widehat{C})(B' \cup \widehat{B}).$$

Theorem 5. *Suppose (f, μ) is a similarity representation of nondegenerate similarity structure $(\Omega, \mathcal{E}, \mathcal{N}, \succeq)$. Then $\succeq$ is robust* if and only if there is a unique $\alpha \in (0, 1]$ such that*

$$(A, B) \succeq (C, D) \iff \frac{\mu(A \cap B)}{\alpha\mu(B) + (1 - \alpha)\mu(A)} \geq \frac{\mu(C \cap D)}{\alpha\mu(D) + (1 - \alpha)\mu(C)}.$$

Similarly to the last section, consider only $(0, 1]^2$. By Theorem 4, if suffices to prove the following lemma.

Lemma 17. *$f(p_1, p_2) > f(q_1, q_2)$ implies*

$$f\left(\frac{1}{\frac{\beta}{p_1} + \frac{1-\beta}{r_1}}, \frac{1}{\frac{\beta}{p_2} + \frac{1-\beta}{r_2}}\right) > f\left(\frac{1}{\frac{\beta}{q_1} + \frac{1-\beta}{r_1}}, \frac{1}{\frac{\beta}{q_2} + \frac{1-\beta}{r_2}}\right)$$

for all $\beta \in (0, 1)$.

Proof. Pick any $\beta \in (0, 1)$ and $a, b, c, a', b', \widehat{a}, \widehat{b}, \widehat{c}$ such that

$$p_1 = \frac{c}{a + c}, \quad p_2 = \frac{c}{b + c}, \quad q_1 = \frac{c}{a' + c}, \quad q_2 = \frac{c}{b' + c}, \quad r_1 = \frac{\widehat{c}}{\widehat{a} + \widehat{c}}, \quad r_2 = \frac{\widehat{c}}{\widehat{b} + \widehat{c}}, \quad \beta = \frac{c}{c + \widehat{c}}.$$

and $a + b + c + a' + b' + \widehat{a} + \widehat{b} + \widehat{c} \leq 1$. Then

$$\begin{aligned}
\left(\frac{1}{\frac{\beta}{p_1} + \frac{1-\beta}{r_1}}, \frac{1}{\frac{\beta}{p_2} + \frac{1-\beta}{r_2}}\right) &= \left(\frac{c + \widehat{c}}{a + \widehat{a} + c + \widehat{c}}, \frac{c + \widehat{c}}{b + \widehat{b} + c + \widehat{c}}\right), \\
\left(\frac{1}{\frac{\beta}{q_1} + \frac{1-\beta}{r_1}}, \frac{1}{\frac{\beta}{q_2} + \frac{1-\beta}{r_2}}\right) &= \left(\frac{c + \widehat{c}}{a' + \widehat{a} + c + \widehat{c}}, \frac{c + \widehat{c}}{b' + \widehat{b} + c + \widehat{c}}\right).
\end{aligned}$$

Then pick $A, B, C, A', B', \widehat{A}, \widehat{B}, \widehat{C}$ such that

$$\mu(A) = a, \ \ \mu(B) = b, \ \ \mu(C) = c, \ \ \mu(A') = a', \ \ \mu(B') = b', \ \ \mu(\widehat{A}) = \widehat{a}, \ \ \mu(\widehat{B}) = \widehat{b}, \ \ \mu(\widehat{C}) = \widehat{c}.$$

Then robustness implies the result. □

Clearly in the representation $\alpha > 0$ since f is strictly increasing in its first argument.

D A Generalization of Anscombe and Aumann (1963)

In this section we provide a slight generalization of Anscombe and Aumann (1963). Let X be a mixture space with mixture operation I. In the original paper, X is the space of simple lotteries and I is the normal weighted arithmetic average. Let $S = \{1, 2, \ldots, n\}$ be a finite set of states. An act is a function $h : S \longrightarrow X$. Let H denote the set of all acts. Abusing the notation a little, write $I_a(f, g) = (I_a(f_1, g_1), \ldots, I_a(f_n, g_n))$; that is, we are mixing acts in a state-by-state manner. Let $\succeq$ be a binary relate defined on H. Axioms 1-4 are direct translations of the classic Anscombe-Aumann axioms. Axiom 5 is from Gilboa and Schmeidler (1989). We identify constant acts by elements in X.

Axiom A1. $\succeq$ is a preference relation.

Axiom A2. $f \succ g$ and $a \in (0, 1)$ implies $I_a(f, h) \succ I_a(g, h)$.

Axiom A3. $f \succ g \succ h$ implies that there exist $a, b \in (0, 1)$ such that $I_a(f, h) \succ g \succ I_b(f, h)$.

Axiom A4. There exist f, g such that $f \succ g$.

Axiom A5. $f_j \succeq g_j$ for all $j \in S$ implies $f \succeq g$.

We say that a function U on H is linear if $U(I_a(f, g)) = aU(f) + (1 - a)U(g)$.

Theorem 6. *$\succeq$ satisfies Axiom A1-A5 if and only if there exist a non-constant linear function U on X and a probability measure μ on S such that $W(f) = \sum_i U(f_i)\mu(i)$ represents $\succeq$. This U is unique up to a positive affine transformation and μ is unique.*

First of all, consider only constant acts. By the mixture space theorem there is U such that $U(h_a(x, y)) = aU(x) + (1 - a)U(y)$. Let $U(f) = (U(f_1), \ldots, U(f_n))$. Then, consider all acts in H, also by the mixture space theorem, there is W such that $W(h_a(f, g)) = aW(f) + (1 - a)W(g)$.

We only have to proof the next lemma. The rest is standard, implied by mixture space uniqueness.

Lemma. *W is linear if and only if there exists a collection of linear functions U_j, for $j \in S$, such that $W(f) = \sum_j U_j(f_j)$.*

Proof. **Step 1:** Let $a_j \in [0,1]$ and $\sum_j a_j = 1$. If $U(f) = \sum_j a_j U(f^j)$ then $W(f) = \sum_j a_j W(f^j)$.

First of all, suppose $\#\{a_j > 0\} = 2$. We know that $U(I_a(f^1, f^2)) = aU(f^1) + (1-a)U(f^2)$. Since $U(f) = U(I_a(f^1, f^2))$, by Axiom 5, it must be the case that $W(f) = W(I_a(f^1, f^2))$. By linearity of W we prove the claim for $\#\{a_j > 0\} = 2$. An inductive argument finishes this step.

Let $x_0 \in X$ and let h_x^j be the act that yields x in state j and x_0 in every other state. Let x_0 be the constant act.

Step 2: $U(I_{\frac{1}{n}}(f, x^0)) = \frac{1}{n}U(f) + \frac{n-1}{n}U(x^0) = \sum_j \frac{1}{n}U(h_{f_j}^j)$.

Define $U_j(x) = W(h_x^j) - \frac{n-1}{n}W(x_0)$ for all $x \in X$, we have

$$\begin{aligned}
\sum_j U_j(f_j) &= \sum_j W(h_{f_j}^j) - (n-1)W(x_0) \\
&= n\sum_j \frac{1}{n}W(h_{f_j}^j) - (n-1)W(x_0) \\
&= nW(I_{\frac{1}{n}}(f, x_0)) - (n-1)W(x_0) \\
&= n\left(\frac{1}{n}W(f) + \frac{n-1}{n}W(x_0)\right) - (n-1)W(x_0) \\
&= W(f).
\end{aligned}$$

Then we show that U_j is linear.

$$\begin{aligned}
U_j(I_a(x,y)) &= W(h_{I_a(x,y)}^j) - \frac{n-1}{n}W(x_0) \\
&= W(I_a(h_x^j, h_y^j)) - \frac{n-1}{n}W(x_0) \\
&= aW(h_x^j) + (1-a)W(h_y^j) - \frac{n-1}{n}W(x_0) \\
&= aU_j(x) + (1-a)U_j(y)
\end{aligned}$$

for all $j \in S$. □