Online Appendix

Eran Shmaya [*]   Leeat Yariv[†]

Current Version: November 11, 2015

ABSTRACT.   This online appendix accompanies the paper "Experiments on Decisions Under Uncertainty: A Theoretical Framework." It provides an analysis of the consequences of restricting subjects' conjectures and a detailed model of data derived from aggregating observations. It also considers a variant of the model studied in the paper in which dimensions of signals are unordered. Last, it presents the formal proofs for the results in Section 4 of the paper.

1.   Plausibility of Conjectures

We start with the notion of conditionally independent and identically distributed (i.i.d.) signals. In a wide array of economic models, ranging from strategic voting to private value auctions, private signals are often assumed to be conditionally independent. Furthermore, in the statistics literature, the term Bayesian is frequently used to indicate conditional independence, the idea being that the sequence of signals $s_1, s_2, \ldots$ are samples from some distribution that depends on the state, or parameter, k. The statistician has some prior belief over the true parameter that governs the distribution of the sample, and she updates her belief given the observations (much like questions leading to maximum likelihood methods in econometrics, see, e.g., Greene, 1993).

In analogy to our original definitions, we say that an updating rule $\sigma : S^* \to \Delta(K)$ is *conditionally i.i.d. Bayesian* if there exists a probability measure μ over $K \times S^\infty$ such that, for every n, signals $s_1, s_2, \ldots$ are conditionally i.i.d. given the state of nature k and for every Borel subset B of K, $\sigma[s_1, \ldots, s_n](B) = \mu(B|s_1, \ldots, s_n)$.

[*]Managerial Economics and Decision Sciences (MEDS), Kellogg School of Management, Northwestern University, Evanston, IL 60208, USA. e-mail: e-shmaya@kellogg.northwestern.edu

[†]Division of the Humanities and Social Sciences, Caltech, Pasadena, CA 91125, USA. e-mail: lyariv@hss.caltech.edu

As the following example illustrates, the restrictions of Theorem 1 do not generally assure that an updating rule that is Bayesian is conditionally i.i.d.[1]

Example Assume that $K = S = \{0, 1\}$. Then $\Delta(K)$ can be identified with the interval $[0, 1]$, where an element $p \in [0, 1]$ stands for the probability measure over K that assigns probability $1 - p$ to 0 and probability p to 1. Consider the updating rule σ given by

$$\sigma\left[s_1, \ldots, s_n\right](1) = \begin{cases} 1/(n+2), & \text{if } s_1 + \cdots + s_n \text{ is even} \\ 1 - 1/(n+2), & \text{if } s_1 + \cdots + s_n \text{ is odd.} \end{cases}$$

In particular, $\sigma\left[\,\right](1) = 1/2$. This updating rule is such that a signal '0' strengthens the agent's previous opinion about the state of nature, whereas a signal '1' changes it drastically in the other direction. A Bayesian statistician, who believes the signals to be i.i.d., cannot exhibit such behavior. However, σ clearly satisfies the condition of Theorem 1 and is therefore Bayesian according to our definition.

Requiring an updating rule to be derived from a belief that signals are conditionally i.i.d. translates directly into a formula derived from Bayes rule. A necessary and sufficient condition for an updating rule to be conditionally i.i.d. Bayesian is then immediate.[2]

One natural extension of i.i.d. processes is to the class of exchangeable processes. These are processes in which the joint distribution of each set of signals does not depend on the order at which they arrive (see Feller, 1966):

---

[1] Naturally, the number of degrees of freedom allowed by a conditionally i.i.d. signal generation process is lower than the number of degrees of freedom corresponding to an arbitrary signal generation process. Indeed, recall that the number of degrees of freedom available for specifying a belief μ over $K \times S^T$ is $|K| \times |S|^T - 1$. However, the number of (non-linear) degrees of freedom corresponding to a measure μ that corresponds to i.i.d. signals given the state of nature (according to the statistics paradigm) is $(|K| - 1) + |K| \times (|S| - 1)$, which is much smaller.

[2] For the sake of illustration, suppose $K = S = \{0, 1\}$. For any updating rule $\sigma : S^* \to \Delta(K)$, denote by $L(s_1, ..., s_n) = \ln \frac{\sigma[s_1,...,s_n](1)}{\sigma[s_1,...,s_n](0)}$ the log-likelihood corresponding to reported beliefs. σ is conditionally i.i.d. Bayesian if and only if there exist $a, b \in R$, $sign(a)sign(b) = -1$, such that

$$L(s_1, ..., s_{n+1}) - L(r_1, ..., r_n) = a \quad \text{whenever} \quad \sum_{i=1}^{n+1} s_i = \sum_{i=1}^{n} r_i, \text{ and}$$

$$L(s_1, ..., s_n) - L(r_1, ..., r_n) = b - a \quad \text{whenever} \quad \sum_{i=1}^{n} s_i = \sum_{i=1}^{n} r_i + 1.$$

(1)

Definition (Exchangeability) *A measure μ over $K \times S^\infty$ is called* conditionally exchangeable *if for any $k \in K$, for any $i_1, .., i_m$ and permutation $\pi : \{i_1, ...i_m\} \rightarrow \{i_1, ...i_m\}$, and any $s, r \in S^\infty$ such that for all $i \notin \{i_1, ...i_m\}$, $s_i = r_i$ and for all $j = 1, ..., m$, $r_{i_j} = s_{\pi(i_j)}$, $\mu(k, s) = \mu(k, r)$.*

In keeping with terminology, we say that an updating rule $\sigma : S^* \rightarrow \Delta(K)$ is *conditionally exchangeable Bayesian* if there exists a conditionally exchangeable probability measure μ over $K \times S^\infty$ such that, for every n and every Borel subset B of K, $\sigma[s_1, \ldots, s_n](B) = \mu(B | s_1, \ldots, s_n)$.

If an updating rule σ is conditionally exchangeable Bayesian, it must be the case that for any n, and permutation $\pi : \{1, ..., n\} \rightarrow \{1, ..., n\}$, $\sigma[s_1, \ldots, s_n](B) = \sigma[s_{\pi(1)}, \ldots, s_{\pi(n)}](B)$ for all Borel subsets B of K. As it turns out, this condition is not sufficient.

Example (Continued) Consider the environment of the Example above and note that
$$\tau(s_1, \ldots, s_n, s_{n+1}) \neq \tau(s_1, \ldots, s_n, 1 - s_{n+1}) \text{ for every } s_1, \ldots, s_{n+1} \in S. \text{ This implies}$$
that the coefficients appearing in the proof of Theorem 1 are determined uniquely. That is, for a convex hull condition of the form:

$$\tau(s_1, ..., s_n) = \lambda(s_{n+1}; s_1, ..., s_n)\tau(s_1, ..., s_n, s_{n+1}) + \lambda(1 - s_{n+1}; s_1, ..., s_n)\tau(s_1, ..., s_n, 1 - s_{n+1}),$$

where $\lambda(s_{n+1}; s_1, ..., s_n) + \lambda(1 - s_{n+1}; s_1, ..., s_n) = 1$, it follows that

$$\lambda(s_{n+1}; s_1, ..., s_n) = \frac{\tau(s_1, ..., s_n) - \tau(s_1, ..., s_n, 1 - s_{n+1})}{\tau(s_1, ..., s_n, s_{n+1}) - \tau(s_1, ..., s_n, 1 - s_{n+1})}.$$

Consider two signals s_1, s_2. We can use the construction of Theorem 1 to derive conditions assuring that the underlying belief μ satisfies $\mu(k, s_1, s_2) = \mu(k, s_2, s_1)$ for $k = 0, 1$. These translate into:

$$\frac{\tau(\cdot) - \tau(1)}{\tau(0) - \tau(1)} \cdot \frac{\tau(0) - \tau(0, 0)}{\tau(0, 1) - \tau(0, 0)} = \frac{\tau(\cdot) - \tau(0)}{\tau(1) - \tau(0)} \cdot \frac{\tau(1) - \tau(1, 1)}{\tau(1, 0) - \tau(1, 1)}.$$

While the updating rule in the Example is consistent with Bayesian updating and satisfies $\tau(s_1, s_2) = \tau(s_2, s_1)$ for all s_1, s_2, it does not satisfy the above equality and is therefore inconsistent with conditional exchangeable Bayesian updating.

Requiring an updating rule to be derived from a belief satisfying conditional exchangeability of signals necessitates the updating rule to be independent of the order by which signals arrive, as well as satisfy more stringent conditions on the marginal effects of different signals that follow directly from the conditional exchangeability conditions, as in the Example above.[3]

2. AGGREGATE OBSERVATIONS

Assume, for example, that $A = \{a, b\}$, $S = \{u, d\}$, and $N = 1$. Let Σ be the set of all possible responses $\sigma : S \to A$ and let $\Sigma_0 \subseteq \Sigma$ be the set of all responses that can be rationalized by some restricted conjectured experiment (i.e., responses satisfying the condition of Theorem 1). Note that $|\Sigma| = 8$ and $|\Sigma_0| = 6$. We assume that subjects' responses $\sigma_1, \sigma_2, ...$ are i.i.d draws from Σ according to some unknown parameter $\mu \in \Delta(\Sigma)$. This is a parametric model, and the dimension of the parameter space $\Delta(\Sigma)$ is 7. We are interested in testing the hypothesis that $\mu(\Sigma_0) = 1$.

Assume now that we only observe, for every subject i, the choice $\sigma_i(\mathbf{s}_i)$, where $\mathbf{s}_1, \mathbf{s}_2, ...$ are fixed and exogenous. With this information, μ is only partially identified: we can only identify the proportion of subjects in the population that choose each action a after any given sequence. Formally, let $\Theta = [0, 1]^{\{\emptyset, u, d\}}$ be a three-dimensional parameter space and let $T : \Delta(\Sigma) \to \Theta$ be given by

$$T(\mu)[\mathbf{s}] = \mu \left(\{\sigma \in \Sigma \mid \sigma[\mathbf{s}] = a\} \right)$$

for every $\mathbf{s} \in \{\emptyset, u, d\}$. Then, using the observable information, we can identify $\theta = T(\mu)$. That is, a pair μ, μ' of parameters induces the same distribution over observables if and only

---

[3]Indeed, for the sake of illustration, suppose $K = S = \{0, 1\}$ and assume that $\tau(s_1, \ldots, s_n, s_{n+1}) \neq \tau(s_1, \ldots, s_n, 1 - s_{n+1})$ for every $s_1, \ldots, s_{n+1} \in S$. Then, it can be readily seen that an updating rule $\sigma : S^* \to \Delta(K)$ is conditionally exchangeable Bayesian if and only if for every sequences $(s_1, \ldots, s_n)$ and $(r_1, \ldots, r_n)$ of signals such that $\sum_{i=1}^n s_i = \sum_{i=1}^n r_i$ one has

$$\tau(s_1, ..., s_n) = \tau(r_1, ..., r_n), \quad \text{and}$$

$$\prod_{i=1}^n \frac{\tau(s_1, ..., s_{i-1}) - \tau(s_1, ..., s_{i-1}, 1 - s_i)}{\tau(s_1, ..., s_{i-1}, s_i) - \tau(s_1, ..., s_{i-1}, 1 - s_i)} = \prod_{i=1}^n \frac{\tau(r_1, ..., r_{i-1}) - \tau(r_1, ..., r_{i-1}, 1 - r_i)}{\tau(r_1, ..., r_{i-1}, r_i) - \tau(r_1, ..., r_{i-1}, 1 - r_i)}.$$

if $T(\mu) = T(\mu')$. Using the characterization in the paper, it is straightforward to verify that

$$\Theta_0 = T\left(\{\mu \mid \mu(\Sigma_0) = 1\}\right) = \left\{\theta \in [0,1]^{\{\emptyset, u, d\}} \mid \theta(\emptyset) \le \theta(u) + \theta(d) \le \theta(\emptyset) + 1\right\}.$$

So, instead of testing $\mu(\Sigma_0) = 1$ we can test $\theta \in \Theta_0$.

3. UNORDERED DIMENSIONS

Throughout the paper, we have assumed there is a natural sequencing of signals. This is why the experimental design was captured by the *number* of signals reported to the subject. This assumption is applicable in many situations (indeed, any context in which signals are tied with time), and eases the presentation. Nonetheless, in certain environments there is no natural ordering of signals and a general conjectured experiment pertains to the *dimensions* of information that are reported (or which elements of the set of signals are reported) and their correlation with the underlying states and realized signals.

Many classical experiments in Psychology exhibit this feature when they entail a description of an individual that contains a selected set of dimensions that are described. For instance, in the classical experiment of Darley and Gross (1983) on stereotypes, subjects were divided into two groups. Both groups were informed about the wealth of a fourth-grader – the first group was told the girl came from a very wealthy family, the second group was told she came from a very poor family.[4] Both groups then watched an ambiguous video of the girl taking an oral test and answering some questions correctly, some not. Subjects were ultimately asked to assess the grade level of the girl's performance. Subjects in the first group rated the girl's grade level as significantly lower than subjects in the second group.[5] While this may indicate subjects' stereotypes, it might also reflect beliefs of subjects regarding the dimensions the experimenter chooses to reveal when expecting a particular response. Even in the "Linda Experiment" (Kahneman and Tversky, 1983) described in Section 4 of the paper there was, in fact, no natural order of the items appearing in the description of Linda and our assumption that the dimensions revealed in the blurb were ordered was a simplifying assumption. The lack

---

[4] Unlike Economic experiments, deception is common in Psychology experiments.

[5] This experiment is tied to a significant amount of work in Economics on "confirmatory bias." See Rabin and Schrag (1999) and references that followed.

of natural order in revealed signals is also present in political economy experiments in which candidates select which issues to highlight and subjects consequently come to an assessment of the candidates (see, for instance, Ansolabehere and Iyengar, 1994).

In this section we extend our analysis to contexts in which there is no natural sequencing of signals. Our "anything goes" result still holds when signals are unordered and conjectured experiments are unrestricted. However, the implications imposed by the natural analogue of restricted conjectured experiments do not carry through from the sequential-signal environment. In fact, we identify stronger observational restrictions that are in the spirit of "Dutch books." Experimental observations can be explained if and only if, after observing the subjects' responses, the experimenter cannot design a sequence of bets that would lead the agent to lose money for sure.

We now require notation in which there is no natural order between dimensions, and the experimenter decides which dimensions to reveal, and not only how many of them. Formally, let A, S, N be, as in our original setting, the set of alternatives (finite, but arbitrary), the set of possible realizations of signals, and the number of dimensions. An *instance* is now given by a pair (D, δ), where $D \subseteq \{1, \ldots, N\}$ and $\delta : D \to S$. The interpretation is that the subject observes the realization of the signals pertaining only to dimensions in D. Let $\mathcal{O}$ be the set of instances. As before, experimental observations are summarized by a mapping $\sigma : \mathcal{O} \to A$.

A *conjectured experiment* is given by a triplet (α, τ, ζ) with values in $A, \mathcal{P}(N)$, and S^N respectively, where $\mathcal{P}(N)$ is the set of subsets of $\{1, \ldots, N\}$. As before, a conjectured experiment is restricted if τ is independent of (α, ζ). A conjectured experiment *explains* the experimental observations σ if for every instance (D, δ) one has

$$\mathbb{P}(\tau = D, \;\; \zeta_i = \delta(d) \text{ for every } d \in D) > 0 \quad \text{ and} \tag{2}$$

$$\sigma(\mathbf{s}) = \arg\max_a \mathbb{P}(\alpha = a | \tau = D, \;\; \zeta_d = \delta(d) \text{ for every } d \in D) \tag{3}$$

The "anything goes" result captured in Theorem 2, and its proof, are valid, *mutatis mutandis*, in the unordered model: for every $\sigma : \mathcal{O} \to A$, the experimental observations summarized by σ can be explained with an unrestricted conjectured experiment.

In the rest of this section, we focus on the case of restricted conjectured experiments.

Say that an instance $\mathbf{s}' = (D', \delta')$ *extends* an instance $\mathbf{s} = (D, \delta)$ if $D \subseteq D'$ and $\delta'(d) = \delta(d)$ for every $d \in D$. The condition in Theorem 1 can be adapted to a necessary condition for existence of an explanation by a restricted conjectured experiment in the unordered model. Namely, suppose the experimental observations are given by $\sigma : \mathcal{O} \to A$. If σ can be explained by a restricted conjectured experiment then for any instance $\mathbf{s} = (D, \delta)$, if for some $a^* \in A$, $\sigma(\mathbf{s}') = a^*$ for every instance $\mathbf{s}' = (D', \delta')$ that extends $\mathbf{s}$, then $\sigma(\mathbf{s}) = a^*$.

The proof that the above condition is necessary to the existence of an explanation by a restricted conjectured experiment follows that shown in the ordered model. However, we will soon see that in the unordered model, this condition is not sufficient. First, we require some additional definitions.

Let $\sigma : \mathcal{O} \to A$ be experimental observations. Let us say that an instance $\mathbf{s} = (D, \delta)$ *agrees* with a realization $(s_1, \ldots, s_N)$ of the signals if $\delta(d) = s_d$ for every $d \in D$. A *positive bet* is given by a triplet $(z, \mathbf{s}, c)$ such that z is a positive real number, $\mathbf{s} = (D, \delta)$ is an instance, and $c \in A$ is such that $c \neq \sigma(\mathbf{s})$. A bet of this form is activated when $\mathbf{s}$ agrees with the realization of the signals and provides the subject z if the state of nature is $\sigma(\mathbf{s})$ and $-z$ if the state of nature is c. Since $\sigma(\mathbf{s})$ is the subjectively most probable state of nature given $\mathbf{s}$, the subject's subjective expected payoff from the bet is strictly positive. For a bet $\beta = (z, \mathbf{s}, c)$, we denote by $p(\beta, \tilde{a}, s_1, \ldots, s_N)$ the payoff under β if the state of nature is $\tilde{a}$ and the realization of the signals is $s_1, \ldots, s_N$. Thus,

$$
p(\beta, \tilde{a}, s_1, \ldots, s_N) = \begin{cases} z & \text{if } \mathbf{s} \text{ agrees with } (s_1, \ldots, s_N) \text{ and } \tilde{a} = \sigma(\mathbf{s}) \\ -z & \text{if } \mathbf{s} \text{ agrees with } (s_1, \ldots, s_N) \text{ and } \tilde{a} = c \\ 0 & \text{otherwise.} \end{cases}
$$

A *Dutch book* is given by a non-empty set B of positive bets such that

$$
\sum_{\beta \in B} p(\beta, \tilde{a}, s_1, \ldots, s_N) \leq 0
$$

for every $\tilde{a} \in A$ and every $s_1, \ldots, s_N \in S$. Thus, if the subject accepts all the bets in B, her final payoff would be non-positive for all realizations of the state and signals.

If the subject had accurate beliefs regarding the underlying process, she would never accept

a Dutch book. It follows intuitively that if the experimental observations can be explained with a restricted conjectured experiment, the subject would never accept a Dutch book. As it turns out, the absence of Dutch books is, in fact, a necessary and sufficient condition for explaining observations with restricted conjectured experiments:

Theorem (Unordered Dimensions): *Experimental observations admit an explanation by a restricted conjectured experiment if and only if they do not entail a Dutch book.*

Proof of Theorem. We will use the following version of the alternative theorem (see Border(2003), Theorem 10).

Theorem of the Alternative: *Let M be a matrix. Then exactly one of the following alternatives holds. Either $xM \leq 0$ for some row vector $x > 0$ or $My \gg 0$ for some column vector $y \geq 0$.*[6]

Fix experimental observations σ. Let M be the $(|\mathcal{O}| \cdot (|A| - 1)) \times (|A| \cdot |S|^N)$ matrix whose rows are indexed $(\mathbf{s}, c)$ for an instance $\mathbf{s} = (D, \delta)$ and $c \neq \sigma(\mathbf{s})$, columns are indexed $(a, s_1, \ldots, s_N)$ for $a \in A$ and $s_1, \ldots, s_N \in S$, and such the matrix entry $M[\mathbf{s}, c][a, s_1, \ldots, s_N]$ at row $(\mathbf{s}, c)$ and column $(a, s_1, \ldots, s_N)$ is given by

$$M[\mathbf{s}, c][a, s_1, \ldots, s_N] = \begin{cases} 1, & \text{if } \mathbf{s} \text{ agrees with } s_1, \ldots, s_N \text{ and } a = \sigma(\mathbf{s}) \\ -1, & \text{if } \mathbf{s} \text{ agrees with } s_1, \ldots, s_N \text{ and } a = c \\ 0, & \text{otherwise.} \end{cases}$$

The assertion of the theorem follows from the Theorem of the Alternative and from the following two simple lemmas.

Lemma A.1 (Explainable σ $-$ Matrix Form): *There exists $y \geq 0$ such that $My \gg 0$ if and only if σ can be explained by a restricted conjectured experiment.*

---

[6]For a vector x, $x \geq 0$ means that all coordinates of x are nonnegative, $x > 0$ means that $x \geq 0$ and $x \neq 0$ and $x \gg 0$ means that all coordinates of x are strictly positive.

Proof. Assume that $(\alpha, \tau, \zeta_1, \ldots, \zeta_N)$ is a restricted conjectured experiment that explains σ. Then for every instance $\mathbf{s}$ and every $c \neq \sigma(\mathbf{s})$ it follows from the definition of explanation and the fact that τ is independent of (α, ζ) that

$$\mathbb{P}(\alpha = \sigma(\mathbf{s})|\mathbf{s} \text{ agrees with } \zeta_1, \ldots, \zeta_N) > \mathbb{P}(\alpha = c|\mathbf{s} \text{ agrees with } \zeta_1, \ldots, \zeta_N).$$

Since by (3) $\mathbb{P}(\zeta_1 = s_1, \ldots, \zeta_N = s_N) > 0$, the last equation is equivalent to

$$\mathbb{P}(\alpha = \sigma(\mathbf{s}), \ \mathbf{s} \text{ agrees with } \zeta_1, \ldots, \zeta_N) > \mathbb{P}(\alpha = c, \ \mathbf{s} \text{ agrees with } \zeta_1, \ldots, \zeta_N).$$

Thus, for every instance $\mathbf{s}$ and every $c \neq \sigma(\mathbf{s})$, we get that

$$\sum_{s_1, \ldots, s_N \text{ and } \mathbf{s} \text{ agrees with } s_1, \ldots, s_N} \mathbb{P}(\alpha = \sigma(\mathbf{s}), \zeta_1 = s_1, \ldots, \zeta_N = s_N) - \mathbb{P}(\alpha = c, \zeta_1 = s_1, \ldots, \zeta_N = s_N) > 0.$$

Therefore, $My \gg 0$ when y is the column vector such that

$$y[a, s_1, \ldots, s_N] = \mathbb{P}(\alpha = a, \zeta_1 = s_1, \ldots, \zeta_N = s_N). \tag{4}$$

Conversely, assume that $My \gg 0$ for some $y \geq 0$. Since the set of solutions to $My \gg 0$ is open, we can assume without loss of generality that $y \gg 0$. Moreover, we can assume without loss of generality that y is normalized so that the sum of its entries is 1. Let α, ζ be random variables whose joint distribution is given by (4) and let τ be a random variable with values in $\mathcal{P}(N)$ and full support, which is independent of (α, ζ). Then the above argument can be reversed to show that the restricted conjectured experiment (α, τ, ζ) explains σ. ∎

Lemma A.2 (Duth Books – Matrix Form): *There exists $x > 0$ such that $xM \leq 0$ if and only if σ admits a Dutch book.*

Proof. Let $x > 0$ such that $xM \leq 0$. Then every coordinate $(\mathbf{s}, c)$ of x where $\mathbf{s} \in \mathcal{O}$ and $c \neq \sigma(\mathbf{s})$ such that $z = x[\mathbf{s}, c] > 0$ gives rise to a positive bet $(z, \mathbf{s}, \sigma(\mathbf{s}))$. Since $x > 0$, the set of coordinates is non-empty. Since $xM \leq 0$, the corresponding set of bets is a Dutch book. The argument is reversible and so the claim follows. ∎

The proof of the Theorem now follows directly. ∎

Dutch Books and Conditional Probabilities. There is a literature using Dutch book arguments as a tool for deriving standard conditional probability assessments (see, e.g., de Finetti, 1937 and Regazzini, 1987).[7] The results of this section do not, however, follow from the analysis presented in that long line of work. Indeed, in our setup experimental observations indicate only the (subjectively) most likely state, and so the set of bets our hypothetical bookie can choose is restricted to bets of the form $(z, \mathbf{s}, c)$, a strict subset of the set of bets the extant literature considers. Consequently, our characterization result allows for multiple posteriors that are consistent with Bayesian updating and the experimental observations, whereas previous analysis pinned down a unique conditional probability system.

4. PROOFS PERTAINING TO PARTIALLY RESTRICTED CONJECTURED EXPERIMENTS

4.1. Proof of Lemma 4. The proof follows several additional lemmas.

Lemma A.3 (Conditioning on Independent Events): *Let X, Y, Z be events in some probability space such that Y, Z are independent given the partition (X, X^c). Then*

$$\mathbb{P}(X|Y, Z) = \rho \left(\frac{\mathbb{P}(Y|X^c)}{\mathbb{P}(Y|X)}, \mathbb{P}(X|Z) \right),$$

where

$$\rho(r, q) = \frac{q}{q + r \cdot (1 - q)}. \tag{5}$$

Proof. One has

$$\mathbb{P}(X|Y, Z) = \frac{\mathbb{P}(X, Y|Z)}{\mathbb{P}(Y|Z)} = \frac{\mathbb{P}(X|Z)\mathbb{P}(Y|X, Z)}{\mathbb{P}(X|Z)\mathbb{P}(Y|X, Z) + \mathbb{P}(X^c|Z)\mathbb{P}(Y|X^c, Z)} =$$
$$\frac{\mathbb{P}(X|Z)\mathbb{P}(Y|X)}{\mathbb{P}(X|Z)\mathbb{P}(Y|X) + \mathbb{P}(X^c|Z)\mathbb{P}(Y|X^c)} = \rho \left(\frac{\mathbb{P}(Y|X^c)}{\mathbb{P}(Y|X)}, \mathbb{P}(X|Z) \right),$$

where we used the facts that $\mathbb{P}(Y|X, Z) = \mathbb{P}(Y|X)$ and $\mathbb{P}(Y|X^c, Z) = \mathbb{P}(Y|X^c))$, which follow from Y and Z being independent given X and X^c. ∎

---

[7]Recently, Weinstein (2013) pointed to the link between dynamic consistency and the impossibility of Dutch books.

Lemma A.4 (Ranking of Conditional Probabilities): *Let X, Y, Z_1, Z_2 be events in some probability space such that*

1. *Y and Z_1 are independent given the partition (X, X^c).*

2. *Y and Z_2 are independent given the partition (X, X^c).*

If $P(X|Y, Z_1) > P(X|Y, Z_2)$ then $P(X|Z_1) > P(X|Z_2)$.

Proof. Let $r = \mathbb{P}(Y|X^c)/\mathbb{P}(Y|X)$ and $q_i = \mathbb{P}(X|Z_i)$ for $i = 1, 2$. By the previous lemma $\mathbb{P}(X|Y, Z_i) = \rho(r, q_i)$, where ρ is given by (5). The assertion follows from the fact that $\rho(r, q)$ is monotone in q. ∎

Proof of Lemma 4. Following the structure of Definition 5, we prove the lemma by induction on the number of remaining layers in the tree describing the experimental observations. Assume first that $\sigma(\mathbf{s}) = $ a and $\sigma(\mathbf{t}) = $ b. Then, by Definition 3, since (α, τ, ζ) explains σ, it follows that

$$\mathbb{P}(\alpha = \mathrm{a}|\tau = n, \; \zeta_1 = s_1, \ldots, \zeta_s = s_n) > 1/2 > \mathbb{P}(\alpha = \mathrm{a}|\tau = n, \; \zeta_1 = t_1, \ldots, \zeta_n = t_n).$$

Applying Lemma A.4 we get

$$\mathbb{P}(\alpha = \mathrm{a}|\zeta_1 = s_1, \ldots, \zeta_n = s_n) > \mathbb{P}(\alpha = \mathrm{a}|\zeta_1 = t_1, \ldots, \zeta_n = t_n),$$

as desired. In particular, this also provides the first induction step pertaining to instances of length N (with no remaining signals that can be observed).

Assume now that $\mathbf{s}\hat{\,}s$ is revealed higher than $\mathbf{t}\hat{\,}t$ for $s, t \in S$. By the induction hypothesis, it follows that

$$\mathbb{P}(\alpha = \mathrm{a}|\zeta_1 = s_1, \ldots, \zeta_n = s_n, \zeta_{n+1} = s) > \mathbb{P}(\alpha = \mathrm{a}|\zeta_1 = t_1, \ldots, \zeta_n = t_n, \zeta_{n+1} = t), \quad (6)$$

for every $s, t \in S$. From Lemma 1 it follows that

$$
\begin{aligned}
&\mathbb{P}(\alpha = \mathrm{a}|\zeta_1 = s_1, \ldots, \zeta_n = s_n) \\
&\quad \in \mathrm{Conv}\{\mathbb{P}(\alpha = \mathrm{a}|\zeta_1 = s_1, \ldots, \zeta_n = s_n, \zeta_{n+1} = s)|s \in S\} \text{ and} \\
&\mathbb{P}(\alpha = \mathrm{a}|\zeta_1 = t_1, \ldots, \zeta_n = t_n) \\
&\quad \in \mathrm{Conv}\{\mathbb{P}(\alpha = \mathrm{a}|\zeta_1 = t_1, \ldots, \zeta_n = t_n, \zeta_{n+1} = t)|t \in S\}.
\end{aligned}
\tag{7}
$$

From (6) and (7) we get

$$
\mathbb{P}(\alpha = \mathrm{a}|\zeta_1 = s_1, \ldots, \zeta_n = s_n) > \mathbb{P}(\alpha = \mathrm{a}|\zeta_1 = t_1, \ldots, \zeta_n = t_n),
$$

as desired.   ■

4.2.   Preliminaries.   Before turning to the proof of Theorem 3, we require some results on interval orders, which we now present.

We use the standard terminology of a *partial order* $\leq$ over a set X being a reflexive, transitive, and antisymmetric relation. A function $f : X \to \mathbb{R}$ is called *strictly monotone* if $f(x) < f(y)$ whenever $x < y$ for every $x, y \in X$. If $\leq, \leq'$ are partial orders over X, we say that $\leq'$ is an *extension of* $\leq$ if $x \leq' y$ whenever $x \leq y$. If $U, V \subseteq X$ then we write $U < V$ if $x < y$ for every $x \in U$ and $y \in V$. If $x \in X$ and $V \subseteq X$ we write $x < V$ when $\{x\} < V$.

A partial order $\leq$ is called a *linear order* if, for every $x, y \in X$, either $x \leq y$ or $y \leq x$.

A partial order $\leq$ over a finite set X is called an *interval order* (see Fishburn, 1985) if there is an assignment of closed real intervals $I_x = [l(x), r(x)]$, where $l(x), r(x)$ are real numbers and $l(x) \leq r(x)$, to the elements x of X such that $x \leq y$ if and only if I_y is to the right of I_x (i.e., $r(x) \leq l(y)$). Such an assignment is called an *interval representation* of $\leq$. If $\leq$ is an interval order then it admits an interval representation such that I_x and I_y have no common endpoints for any distinct x and y in X. Note that if all the intervals I_x in a representation of $\leq$ are distinct and degenerate then $\leq$ is a linear order, and every linear order admits such a representation.

A *partition* of a set X is a collection $\mathcal{A}$ of non-empty mutually disjoint subsets of X such that $X = \bigcup_{U \in \mathcal{A}} U$. Elements of $\mathcal{A}$ are called *atoms*.

A partial order $\leq$ over a set X is called a *ranking* if its elements can be partitioned into *ranks* $X_1, \ldots, X_m$ such that two elements are incomparable if and only if they belong to the same rank. Every ranking is an interval order.

Lemma A.5 (Representation): *Let $(X, \leq)$ be a finite set equipped with an interval order, $\mathcal{A}$ a partition of X, and $F : \mathcal{A} \to (0, 1)$ a real-valued one-to-one function such that $0 < F(U) < 1$ for every atom U of A. Assume that the following condition is satisfied:*

$$\text{If } V, U \text{ are atoms of } \mathcal{A} \text{ and } V < U \text{ then } F(V) < F(U). \tag{8}$$

Then there exists a strictly monotone, one-to-one function $f : X \to R$ such that

$$\min\{f(x)|x \in U\} \leq F(U) \leq \max\{f(x)|x \in U\} \tag{9}$$

for every atom U of A, and the inequalities in (9) are strict whenever $|U| > 1$.

Proof. We first prove the lemma under the stronger assumption that $\leq$ is a linear order over X. Under this assumption, we can assume without loss of generality that $X = \{1, \ldots, n\}$ with the standard order over numbers. For every $r \in \mathbb{R}$ let

$$U(r) = \min\{F(U)|U \text{ is an atom of } \mathcal{A}, F(U) > r\},$$

where we define the minimum over the empty set $\varnothing$ to be 1. From the definition of $U(r)$ it follows that

$$r < U(r) \quad \text{and} \tag{10}$$

$$\text{if } r < F(U) \text{ then } U(r) \leq F(U), \tag{11}$$

for every $r \in \mathbb{R}$ and every atom U of $\mathcal{A}$.

For $x \in X$, let $\pi(x)$ be the atom of $\mathcal{A}$ that contains x. Call elements $x, y \in X$ *siblings* if $\pi(x) = \pi(y)$.

We now define $f : X \to \mathbb{R}$ inductively so that f is strictly monotone and the following

condition is satisfied for every $x \in X$ and every atom U of $\mathcal{A}$:

$$\text{If } x < U \text{ then } f(x) < F(U) \tag{12}$$

Let $f(0) = 0$ and let $z \geq 1$. Suppose we have already defined $f(1) < \cdots < f(z-1)$, such that (12) is satisfied for $x = 1, \ldots, z-1$.

Case 1. $[|\pi(z)| = 1]$ *Choose $f(z) = F(\pi(z))$. Note that in this case $z - 1 < \{z\} = \pi(z)$, where the inequality follows from (12) for $x = z - 1$ and $U = \{z\}$. Therefore, $f(z-1) < F(\pi(z)) = f(z)$. In addition, if U is an atom of $\mathcal{A}$ and $z < U$ then $\pi(z) = \{z\} < U$ and therefore $F(\pi(z)) < F(U)$ by (8). Thus, (12) is satisfied for $x = z$.*

Case 2. $[|\pi(z)| > 1$ and $z = \max \pi(z)]$ *Let $r = f(z-1) \vee F(\pi(z))$. We choose $f(z)$ such that $r < f(z) < U(r)$. In particular $f(z-1) < f(z)$. Let U be an atom of $\mathcal{A}$ such that $z < U$. Then, $z - 1 < U$ and, therefore, $f(z-1) < F(U)$ by (12) with $x = z - 1$. Also, since $z = \max \pi(z)$, it follows that $\pi(z) < U$ and, therefore, $F(\pi(z)) < F(U)$ by (8). This implies that $r < F(U)$ and so $f(z) < U(r) \leq F(U)$, where the second inequality follows from (11). Thus, (12) is satisfied for $x = z$.*

Case 3. $[|\pi(z) > 1|$ and $z < \max \pi(z)]$ *Choose $f(z)$ so that $f(z-1) < f(z) < U(f(z-1))$. Let U be an atom of $\mathcal{A}$ such that $z < U$. Then, $z - 1 < U$ and therefore $f(z-1) < F(U)$ by (12) with $x = z - 1$. Hence, $f(z) < U(f(z-1)) \leq F(U)$, where the second inequality follows from (11).*

We now claim that the function f defined above satisfies (9). Indeed, let U be an atom of $\mathcal{A}$. Suppose first that $|U| = 1$, so that $U = \{z^*\}$ for some $z^* \in X$. Then $\pi(z^*) = U$ and therefore $f(z^*) = F(U)$ by Case 1 in the construction of f. In particular, (9) is satisfied with equalities. Suppose now that $|U| > 1$ and let $z_{\max}$ and $z_{\min}$ be the maximal and minimal elements of U. Then $\pi(z_{\min}) = \pi(z_{\max}) = U$. From Case 2 above, $F(U) < f(z_{\max})$. In addition, $z_{\min} - 1 < U$ so that $f(z_{\min} - 1) < F(U)$ by (12), and therefore $f(z_{\min}) < U(f(z_{\min} - 1)) \leq F(U)$, where the first inequality follows from Case 3 in the construction of f and the second inequality follows from (11). In particular, (9) is satisfied with strict inequalities. Finally, note that f

is one-to-one because it is strictly monotone. The proof of the lemma for linear orders is now complete.

We now turn to environments in which the order relation $\leq$ over X is an interval order. We will show that $\leq$ can be extended to a linear order $\leq'$ over X that satisfies (8). Then, by the previous argument, there exists a ($\leq'$-strictly monotone and therefore) $\leq$-strictly monotone function f that satisfies (9). Let $I_x = [l(x), r(x)]$ be a representation of $\leq$ and assume, without loss of generality, that I_x and I_y have no common endpoints whenever $x, y \in X$ and $x \neq y$.

For an atom U of $\mathcal{A}$ let $L(U) = \min\{l(x)|x \in U\}$ and $R(U) = \max\{r(x)|x \in U\}$. Fix $x^* \in X$ such that I_{x^*} is not degenerate. We will find $p^* \in I_{x^*}$ such that the extension $\leq'$ that is induced by the interval representation I'_x given by

$$
I'_x = \begin{cases} [p^*, p^*], & \text{if } x = x^* \\ I_x, & \text{otherwise} \end{cases}
$$

satisfies (8). If $r(x^*) < R(\pi(x^*))$, we choose $p^* = l(x^*)$. Then I' and I induce the same order over atoms of $\mathcal{A}$ and therefore (8) holds. Otherwise, if $l(x^*) > L(\pi(x^*))$, we choose $p^* = r(x^*)$ and again I' and I induce the same order over atoms of $\mathcal{A}$. If $r(X^*) = R(\pi(x^*))$ and $l(x^*) = L(\pi(x^*))$ we proceed as follows: Let

$$
\begin{aligned}
p_{\max} &= \min\{R(V)|V \in \mathcal{A}, F(\pi(x^*)) \leq F(V)\} \text{ and} \\
p_{\min} &= \max\{L(W)|W \in \mathcal{A}, F(W) \leq F(\pi(x^*))\},
\end{aligned} \tag{13}
$$

where the minimum and maximum over the empty set $\varnothing$ are taken as 1 and 0 respectively. We claim that $p_{\min} < p_{\max}$. Indeed, let $V, W \in \mathcal{A}$ such that $F(W) \leq F(\pi(x^*)) \leq F(V)$. We distinguish between two cases:

Case 4. *[$V = W$] Since F is one-to-one it follows that $W = \pi(x^*) = V$. Since I_{x^*} is not degenerate it follows that*

$$
L(W) \leq l(x^*) < r(x^*) \leq R(V).
$$

Case 5. *[$V \neq W$] By (8) there exists $y \in V$ and $z \in W$ such that $y \not< z$ (i.e., it is not the case that $y < z$.) Since $V \neq W$ it follows that $y \neq z$. Since I_x is a representation of $\leq$, it follows that $l(z) < r(y)$. In particular, $L(W) < R(V)$.*

Thus, $L(W) < R(V)$ in both cases. As V, W were arbitrary, it follows that $p_{\min} < p_{\max}$ as desired. Let p^* be chosen so that $p_{\min} < p^* < p_{\max}$ and $I_x \neq [p^*, p^*]$ for every $x \in X$. We claim that $\leq'$ satisfies (8). For atoms U, V of $\mathcal{A}$ which are different from $\pi(x^*)$, $\leq'$ satisfies (8) because $\leq$ does. Assume now that $U = \pi(x^*)$, $V \in \mathcal{A}$, and $V <' U$. We have to show that $F(V) < F(U)$. Indeed, if $F(U) \leq F(V)$ then $p_{\max} \leq R(V)$ by (13), which leads to a contradiction, since $R(V) \leq p^* < p_{\max}$, where the first inequality follows from the definition of $\leq'$. By a similar argument, (8) is satisfied when $U = \pi(x^*)$, $V \in \mathcal{A}$, and $U <' V$.

We showed that if $\leq$ is an interval order over X with representation I_x that satisfies (8) and $x^* \in X$, then $\leq$ can be extended to an interval order over X with representation I_x' that satisfies (8) such that $I_x' \subseteq I_x$ for every $x \in X$ and I_{x^*} is degenerate. Going over all the elements of X, we get an extension $\leq'$ of $\leq$ which satisfies (8) such that all the intervals in the representation $\leq'$ are degenerate. Therefore, $\leq'$ is a linear order. ∎

Remark The assumption that $\leq$ is an interval order is essential in Lemma A.5. As a counterexample, let $X = \{a, b, c, d\}$ with $a < b$ and $c < d$ and let $\mathcal{A} = \{\{a\}, \{c\}, \{b, d\}\}$ and $F : \mathcal{A} \to (0, 1)$ be such that $F(\{b, d\}) < F(\{a\}) < F(\{c\})$. Then F is one-to-one and (8) is trivially satisfied, but there exists no strictly monotone $f : X \to R$ satisfying (9).

4.3. Proof of Theorem 3. Let $\sigma : S^{\leq N} \to \{a, b\}$ represent experimental observations that satisfy the condition of Theorem 3. We assume, without loss of generality, that $\sigma(e) = a$.

Let $\pi : S^n \to S^{n-1}$ be the *parent function* of the tree of instances: $\pi(s_1, \ldots, s_n) = (s_1, \ldots, s_{n-1})$. Let $\leq^n$ stand for the relation "revealed higher" over S^n: For two nodes $\mathbf{s}, \mathbf{t} \in S^n$, $\mathbf{t} \leq^n \mathbf{s}$ whenever $\mathbf{s}$ is revealed higher than $\mathbf{t}$.

We claim first that $\leq^n$ is an interval order for every n. We prove this assertion by induction over $N - n$ (the remaining layers in the tree). For $n = N$ the order over S^n is, in fact, a ranking, the ranks being the sets $\{\mathbf{s} \in S^N | \sigma(\mathbf{s}) = a\}$ and $\{\mathbf{s} \in S^N | \sigma(\mathbf{s}) = b\}$. Assume now that $\leq^n$ is represented by $I_{\mathbf{s}}^n = [l^n(\mathbf{s}), r^n(\mathbf{s})]$. Without loss of generality, suppose also that

$0 < l^n(\mathbf{s}) \le r^n(\mathbf{s}) < 1$ for every $\mathbf{s} \in S^n$. For $\mathbf{s} \in S^{n-1}$ let $I_\mathbf{s}^{n-1} = [l^{n-1}(\mathbf{s}), r^{n-1}(\mathbf{s})]$, where

$$l^{n-1}(\mathbf{s}) = \lambda(\mathbf{s}) + \min\{l^n(\mathbf{s}') | \mathbf{s}' \in S^n \text{ and } \pi(\mathbf{s}') = \mathbf{s}\}, \text{ and}$$

$$r^{n-1}(\mathbf{s}) = \lambda(\mathbf{s}) + \max\{r^n(\mathbf{s}') | \mathbf{s}' \in S^n \text{ and } \pi(\mathbf{s}') = \mathbf{s}\}, \text{ where}$$

$$\lambda(\mathbf{s}) = \begin{cases} 1, & \text{if } \sigma(\mathbf{s}) = a \\ 0, & \text{if } \sigma(\mathbf{s}) = b \end{cases}.$$

Then $r^{n-1}(\mathbf{t}) \le l^{n-1}(\mathbf{s})$ if one of the following is satisfied:

- $\sigma(\mathbf{s}) = a$ and $\sigma(\mathbf{t}) = b$.

- $\sigma(\mathbf{s}) = \sigma(\mathbf{t})$ and $\mathbf{s}' <^n \mathbf{t}'$ for every child $\mathbf{s}'$ of $\mathbf{s}$ and every child $\mathbf{t}'$ of $\mathbf{t}$.

It follows from the recursive definition of $\le^n$ that I^{n-1} is a representation of $\le^{n-1}$.

We now construct an assignment $p_\mathbf{s} \in (0, 1)$ of probabilities for every $\mathbf{s} \in S^{\le N}$ such that

$$p_e > 1/2, \tag{14}$$

$$p_\mathbf{s} \in \mathrm{ri}\,(conv\{p(\mathbf{s}') | \mathbf{s}' \text{ is a child of } \mathbf{s}\}) \text{ for every } \mathbf{s} \in S^{<N}, \text{ and} \tag{15}$$

$$p \text{ is strictly-} \le^n \text{ monotone.} \tag{16}$$

To construct p, we go over the nodes from the root to the leafs. Choose p_e arbitrarily so that $p_e > 1/2$. Assume we defined $p_\mathbf{s}$ for $\mathbf{s} \in S^{n-1}$ such that $\mathbf{s} \mapsto p_\mathbf{s}$ is $\le^{n-1}$-strictly monotone. The set S^n is equipped with an interval order $\le^n$. The parent function $\pi : S^n \to S^{n-1}$ induces a partition over S^n (the atoms of the partition are $\pi^{-1}(\mathbf{s})$ for $\mathbf{s} \in S^{n-1}$). Let $U = \pi^{-1}(\mathbf{s})$ and $V = \pi^{-1}(\mathbf{t})$ be two such atoms. If $V <^n U$ then, by Definition 5, it follows that $\mathbf{t} <^{n-1} \mathbf{s}$ and therefore $p_\mathbf{t} < p_\mathbf{s}$. Thus, Condition (8) of Lemma A.5 is satisfied and p can be defined over S^n such that (15) and (16) are satisfied. Note that (16) and the definition of $\le^n$ imply that

$$p_\mathbf{s} > p_\mathbf{t} \text{ whenever } \sigma(\mathbf{s}) = a, \sigma(\mathbf{t}) = b \text{ and } d(\mathbf{s}) = d(\mathbf{t}). \tag{17}$$

We now claim that for every $1 \le n \le N$, there exists some $0 < r_n < \infty$ such that

$$\rho(r_n, p_\mathbf{s}) > 1/2 \text{ for every } \mathbf{s} \in S^n \text{ such that } \sigma(\mathbf{s}) = \text{a, and}$$
$$\rho(r_n, p_\mathbf{t}) < 1/2 \text{ for every } \mathbf{t} \in S^n \text{ such that } \sigma(\mathbf{t}) = \text{b,} \tag{18}$$

where ρ is given by (5). Indeed, fix $1 \le n \le N$ and let q be a real number such that

$$q < p_{\mathbf{s}} \text{ for every } \mathbf{s} \in S^n \text{ such that } \sigma(\mathbf{s}) = \text{a, and}$$

$$p_{\mathbf{t}} < q \text{ for every } \mathbf{t} \in S^n \text{ such that } \sigma(\mathbf{t}) = \text{b.}$$

(19)

The existence of such a q follows from (17). Since the function ρ is continuous and monotone in the first argument, and since $\lim_{r \to 0} \rho(q, r) = 0$ and $\lim_{r \to \infty} \rho(q, r) = 1$, it follows that there exists some r_n such that $\rho(r_n, q) = 1/2$. Since the function ρ is monotone in the second argument, (19) implies that

$$\rho(r_n, p_{\mathbf{s}}) > \rho(r_n, q) = 1/2 \text{ for every } \mathbf{s} \in S^n \text{ such that } \sigma(\mathbf{s}) = \text{a, and}$$

$$\rho(r_n, p_{\mathbf{t}}) < \rho(r_n, q) = 1/2 \text{ for every } \mathbf{t} \in S^n \text{ such that } \sigma(\mathbf{t}) = \text{b,}$$

as desired. We now define $r_0 > 0$ such that (18) is also satisfied for $n = 0$, and moreover, $1 \in \text{ri}(\text{conv}\{r_n | n = 0, \dots, N\})$. To achieve this, choose $r_0 > 1$ arbitrarily if $r_n < 1$ for some $n \in \{1, \dots, N\}$; choose $r_0 < 1$ such that $\rho(r_0, p_e) > 1/2$ if $r_n \ge 1$ for every $n \in \{1, \dots, N\}$ and $r_n > 1$ for some $n \in \{1, \dots, N\}$; and choose $r_0 = 1$ if $r_n = 1$ for every $n \in \{1, \dots, N\}$. Since $p_e > 1/2$, such a choice can be made and, furthermore, (18) is satisfied (recall that $\sigma(e) = a$). Let $\lambda_n > 0$ be such that $\sum_{n=0}^{N} \lambda_n = 1$ and $\sum_{n=0}^{N} \lambda_n r_n = 1$.

We now construct (α, τ, ζ) such that τ and ζ are independent given α,

$$\mathbb{P}\left(\alpha = \text{a} | \zeta_i = s_i \text{ for } 1 \le i \le n\right) = p_{\mathbf{s}},$$

(20)

for every $\mathbf{s} = (s_1, \dots, s_n) \in S^{\le N}$, and

$$\frac{\mathbb{P}(\tau = n | \alpha = \text{b})}{\mathbb{P}(\tau = n | \alpha = \text{a})} = r_n$$

(21)

for every $n \in \{1, \dots, N\}$. The existence of such a triplet follows from the following argument. By Lemma 1 there exists random variables (α, ζ) over some probability space with values in U and S^N respectively such that (20) is satisfied. Possibly augmenting the underlying probability space, we introduce the random variable τ with values in $\mathbf{N}$ such that τ is independent of ζ

given α,

$$\mathbb{P}\left(\tau = n | \alpha = \mathrm{a}\right) = \lambda_n, \text{ and}$$

$$\mathbb{P}\left(\tau = n | \alpha = \mathrm{b}\right) = \lambda_n r_n.$$

Then (21) is satisfied. By Lemma A.3, (20), and (21), we get that

$$\mathbb{P}\left(\alpha = \mathrm{a} | \tau = n, \zeta_i = s_i \text{ for } 1 \leq i \leq n\right) = \rho(r_n, p_{\mathbf{s}}).$$

Finally, from the latter equation and (19) it follows that

$$\mathbb{P}\left(\alpha = \mathrm{a} | \tau = n, \zeta = \mathbf{s}\right) > 1/2 \text{ for every } \mathbf{s} \in S^n \text{ such that } \sigma(\mathbf{s}) = \mathrm{a}, \text{ and}$$

$$\mathbb{P}\left(\alpha = \mathrm{b} | \tau = n, \zeta = \mathbf{t}\right) < 1/2 \text{ for every } \mathbf{t} \in S^n \text{ such that } \sigma(\mathbf{t}) = \mathrm{b},$$

as desired. ∎

REFERENCES

[1] Stephen Ansolabehere and Shanto Iyengar (1994), "Riding the Wave and Issue Ownership: The Importance of Issues in Political Advertising and News," *Public Opinion Quarterly*, **58**, 335-357.

[2] John M. Darley and Paget H. Gross (1983), "A Hypothesis Confirming Bias in Labeling Effects," *Journal of Personality and Social Psychology*, **44(1)**, 20-33.

[3] William Feller (1966), *An Introduction to Probability Theory and its Applications*, Wiley Series in Probability and Mathematical Statistics.

[4] Bruno de Finetti (1937), "La Prevision: ses lois logiques, se sources subjectives," *Annales de l'Institut Henri Poincare*, **7**, 1-68. Translated into English and reprinted in Kyburg and Smokler, Studies in Subjective Probability. Huntington, NY: Krieger; 1980.

[5] William H. Greene (1993), *Econometric Analysis*, Macmillan Publishing Company.

[6] Eugenio Regazzini (1987), "De Finetti's Coherence and Statistical Inference," *The Annals of Statistics*, **15(2)**, 845-864.