Comparison of Decisions Under Unknown Experiments*

Andrew Caplin† and Daniel Martin‡

December 16, 2020

Abstract

We take the perspective of an econometrician who wants to determine which of two experiments provides higher expected utility but only knows the decisions under each experiment. To compare these decisions, the econometrician must make inferences about what the experiment might have been for each set of decisions. We provide a necessary and sufficient condition that identifies when every experiment consistent with one set of decisions has a higher value of information than every experiment consistent with the other set of decisions.

Key words: Information economics, revealed preferences, experiments
JEL codes: D60, D83, D91

*We thank Edwin Muñoz Rodriguez for excellent research assistance and Jose Apesteguia, Victor Aguiar, Miguel Ballester, Henrique de Oliveira, Alex Frankel, Ben Handel, Emir Kamenica, Matt Kovach, Josh Schwartzstein, Charlie Sprenger, Bruno Strulovici, Dmitry Taubinsky, Gerelt Tserenjigmid, and seminar audiences at Alicante, Chicago, and Northwestern for valuable feedback. All mistakes are our own.

†Department of Economics, New York University.

‡Kellogg School of Management, Northwestern University.

1 Introduction

There are many facts about the world (or "states of the world") that can be payoff-relevant for decision-makers. For example, their payoffs can depend on the fundamentals of a stock, the effectiveness of a vaccine, characteristics of a health plan, and so on. These facts can be presented to decision-makers in a number of different ways by better-informed parties. For instance, advisors, firms, news networks, and governments can choose to selectivity allocate information about the facts, as in Bayesian persuasion (e.g., Kamenica and Gentzkow 2011) or voluntary information disclosure (e.g., Milgrom 1981).[1] They can also choose the format that this information takes, making it easier or harder for decision-makers to understand.[2] In both of these cases, it has been well-documented that the way the better-informed party chooses to present the facts can strongly influence how well informed decision-makers are when they make their choices.

We take the perspective of an econometrician who wants to compare different ways of presenting the facts based on how valuable that information was for decision-makers. For example, the econometrician might want to determine whether the advice from one financial advisor helped decision-makers make better portfolio allocations than the advice provided by a different advisor, whether watching one news program helped decision-makers choose better health behaviors than watching another news program (e.g., Bursztyn, Rao, Roth, and Yanagizawa-Drott 2020), or whether one description of fees lead to better health plan choices made than a different description of fees (e.g., Bhargava, Loewenstein, and Sydnor 2017).

We model a presentation of the facts as an *experiment* (a joint distribution of signals and states). Traditionally, experiments have been used to model physical activities where observing signals is easy (e.g., drilling for oil or performing a medical test). However, in our application it is more challenging for the econometrician to observe the experiment itself. For instance, it can be hard to know what an advisor said to their clients if there are privacy concerns, advice is proprietary, or it is challenging to codify the advice provided. Further, even if we know exactly the information they provided to their clients, it might be challenging to know what the clients understood about the facts based on that information. Yet it is often possible for the econometrician to observe the actions taken under each experiment. For instance, many data sets contain the stocks that were bought, the vaccines that were taken, the health plans that were selected, and so on.

[1] For reviews of the Bayesian persuasion literature see Kamenica (2019), the voluntary disclosure literature see Dranove and Jin (2010), and the disclosure experimental literature see Jin, Luca, and Martin (2015).

[2] For example, see Hastings and Tejeda-Ashton (2008), Choi, Laibson, and Madrian (2009), Abeler and Jager (2015), Carrera and Villas-Boas (2015), Ericson and Starc (2016), Esponda and Vespa (2016), Jin, Luca, and Martin (2018), and Carpenter et al. (2019).

Because of this, we assume the econometrician only knows the actions taken under each experiment, and nothing about the experiment itself (either the signal structure or the signal realizations). For example, all the econometrician might know about a particular experiment is that it results in the following joint distribution over actions (a, b) and states (ω_1, ω_2):

$$P_g = \begin{array}{cc} \begin{array}{cc} \omega_1 & \omega_2 \end{array} & \\ \begin{pmatrix} .4 & .1 \\ .1 & .4 \end{pmatrix} & \begin{array}{c} a \\ b \end{array} \end{array}$$

In practice, this joint distribution could be the frequency a stock is bought when it has certain fundamentals, a vaccine is taken when it has certain effectiveness, a PPO health plan is chosen when it has certain benefits, a TV model is purchased when a competing model is on sale, a loan is made to someone who will default, a test is ordered for someone who actually has a disease, etc.[3]

Thus, the econometrician wants to be able to determine which of two experiments provides higher expected utility for the decision-maker based just on the joint distributions of actions and states under each experiment. For example, does the experiment that produced P_g provide higher expected utility than the experiment that produced P_h?

$$P_g = \begin{array}{cc} \begin{array}{cc} \omega_1 & \omega_2 \end{array} & \\ \begin{pmatrix} .4 & .1 \\ .1 & .4 \end{pmatrix} & \begin{array}{c} a \\ b \end{array} \end{array} \quad \& \quad P_h = \begin{array}{cc} \begin{array}{cc} \omega_1 & \omega_2 \end{array} & \\ \begin{pmatrix} .15 & .05 \\ .35 & .45 \end{pmatrix} & \begin{array}{c} a \\ b \end{array} \end{array}$$

When there is an objectively correct choice, such as a clearly dominant option, these assessments are easy. For example, Handel (2013) and Bhargava, Loewenstein, and Sydnor (2017) leverage settings where some health care plans dominate others. Outside of these special cases, choice options have trade-offs, and decision-makers have private knowledge about their tastes and requirements that allow them to negotiate these trade-offs.

To compare decisions more generally, the econometrician must make inferences about what the experiment might have been for each set of decisions. For this, we use the same maintained assumptions as Blackwell (1953): experiment π_g is *consistent* with P_g if P_g maximizes expected utility among the joint distributions of actions and states feasible under experiment π_g.

Using the set of consistent experiments, we define a binary relation $\succsim_V$ which enables the econometrician can rank decisions. We say $P_g \succsim_V P_h$ if for every utility function u, every

[3]In all of these cases, the state impacts utility, the decision maker may not be fully informed about the state, and the econometrician knows the state.

experiment consistent with P_g has a higher value of information than every experiment consistent with P_h.

To characterize the relation $\succsim_V$, we leverage two features of the problem. First, for a given utility function u, every experiment consistent with a given P_g has the same value of information, which is the expected utility provided by P_g for that utility function. Second, there are utility functions for which no experiments are consistent with P_g or P_h. Because the condition for $P_g \succsim_V P_h$ is trivially satisfied for such utility functions, the econometrician does not need to consider them when comparing P_g and P_h.

Thus, the econometrician can safely conclude that $P_g \succsim_V P_h$ if they can rule out enough utility functions to ensure that P_g provides higher expected utility than P_h for all remaining utility functions. While it is possible to rule out some utility functions when just a single set of decisions is observed, observing two sets of decisions allows more utility functions to be ruled out. We provide examples illustrating that utility functions can be restricted sufficiently to conclude that P_g provides higher expected utility than P_h.

We establish the general logic by first showing that restricting utility functions has a clean geometric structure in the space of outcome lotteries. We then build on this structure to produce a necessary and sufficient condition for $P_g \succsim_V P_h$. The necessity and sufficiency of this condition follows as a direct consequence of the separating hyperplane theorem. In technical terms, this condition requires that a vector representing the difference in outcomes between P_g and P_h falls in the cone generated by the restrictions for a utility function to be consistent with P_g and P_h. This corresponds to solving a system of linear equations, so it is simple to check, and MATLAB programs that implement it are provided.[4]

Finally, we draw a point of connection between $\succsim_V$ and Blackwell's Theorem by defining a *revealed experiment* for each distribution of actions and states. This is the experiment if we assume there is a single posterior at which each action is taken, so it can inferred directly from the distribution of actions and states.[5] The revealed experiment for P_h being a garbling of the revealed experiment for P_g is sufficient to reveal that $P_g \succsim_V P_h$, but this condition is unnecessarily strong. A garbling of revealed experiments ensures one revealed experiment has a higher utility function for *all* possible utility functions, not just those utility functions where there exist experiments consistent with P_g and P_h.

The rest of the paper is organized as follows. We start Section 2 by providing our framework and formally defining our relation. We then show how utility functions can be ruled out when there is a single set of decisions and the econometrician does not know the outcomes from taking each action in each state. Next, we show how having a second set

[4] Programs available at https://github.com/danieljosephmartin/framinginformationwelfare.git.

[5] The revealed experiment reflects the minimal informativeness the decision-maker can have given the distribution of actions and states.

of decisions or knowing outcomes can allow the econometrician to rule out enough utility functions for decisions to be ordered according to the relation.

Section 3 introduces our geometric representation of ruling out utility functions for which no consistent experiments exist. Leveraging this representation, we identify a testable condition that is both necessary and sufficient for decisions to be ranked according to $\succsim_V$. We end this section by showing how this condition can be amended when outcomes are known and when there is clear dominance between some outcomes.

Section 4 shows that the garbling of revealed experiments is sufficient for decisions to be ranked according to $\succsim_V$. We then revisit a preceding example to illustrate how decisions can be ranked according to $\succsim_V$ even if neither revealed experiment is a garbling of the other. Section 5 concludes by discussing related literature.

2 Framework

For each presentation of the facts, we assume the decision-maker (DM) starts with a strictly positive prior over a finite set of states Ω given by $\mu \in \Delta(\Omega)$. The DM receives a signal realization, and as is now standard, a signal realization is represented by the posterior belief $\gamma \in \Delta(\Omega)$ it generates (see Kamenica and Gentzkow 2011). This process is summarized by experiment π which is a joint distribution over states Ω and posteriors $\Delta(\Omega)$, and for notational simplicity we assume the experiment has finite support over $\Delta(\Omega)$.

Given a posterior belief generated by a experiment, we assume the DM implements a decision rule $\sigma : \Delta(\Omega) \to \Delta(A)$, where A is a finite set of actions. The DM receives outcome $x(a, \omega) \in X$ when action $a \in A$ is chosen in state $\omega \in \Omega$, and the decision rule maximizes expected utility based on a utility function $u : X \to \mathbb{R}$.[6]

We consider an econometrician who wants to compare the expected utility provided by two experiments (two presentations of the facts). We assume the unconditional probability of states is the same across the experiments, and our interpretation of this assumption is that a presentation of the facts cannot change the facts themselves. For instance, a financial advisor cannot change the actual financial conditions of the institutions their clients might invest in. In the case of choosing which way to describe fees, the fees themselves cannot be changed. This assumption is not required for our characterization, but it simplifies our analysis and the connection to Blackwell's approach.

All the econometrician knows about these experiments is the joint distributions over actions and states they generate. We refer to an arbitrary joint distribution of actions and states as $P_f \in \{P_g, P_h\}$ and denote $P_f(a, \omega)$ as the probability of choosing action a and being

[6] The outcome space is also finite, and we denote its cardinality as M. It has generic element x or x_m.

in state ω for P_f.[7]

As with other stochastic choices, each P_g and P_h can be interpreted as watching the DM face a decision infinitely often. In practice, one might estimate it from repeated but finite choice data or by looking at a population rather than an individual, as in the literature on discrete choice following McFadden (1973). For notational simplicity, we assume that all outcomes can be obtained by taking some action in some state, and that for each distribution of actions and states, each action is chosen in some state and an action is chosen in each state.[8] This joins a growing literature that considers stochastic choice to be essential for studying information and utility (e.g., Manzini and Mariotti 2014, Apesteguia and Ballester 2018).

For what follows, it is not necessary for the econometrician to also know the outcome received from taking each action in each state, as it is without loss of generality for the econometrician to arbitrarily assign a distinct outcome to every action in every state. However, the presence of an outcomes space allows us to accommodate cases where the econometrician knows that utility is equal across some states and actions. For example, the econometrician might know that an action is "safe" because it yields the same outcome in every state.

2.1 Comparison of Decisions

The econometrician would like to compare joint distributions of actions and states P_g and P_h based on the value of the information provided by the experiments that generated them. Without knowing anything about the structure of these experiments, the econometrician must determine the experiments that are consistent with P_g and P_h.

We define consistency using the same maintained assumptions as in Blackwell (1953). An experiment π_f is consistent with P_f if P_f maximizes expected utility among distributions of actions and states feasible under π_f.

A joint distribution of actions and states P_f is *feasible* under π_f if there exists a decision rule $\sigma_f : \Delta(\Omega) \to \Delta(A)$ such that

$$P_f(a, \omega) = \sum_{\gamma \in supp(\pi_f)} \pi_f(\gamma, \omega)\sigma_f(a|\gamma)$$

where the set of possible posterior beliefs is given by $supp(\pi_f)$. Given u, the highest expected

[7] P_g and P_h are state-dependent stochastic choice (SDSC) data, which were proposed for information-theoretic revealed preferences by Caplin and Martin (2015).

[8] Our results would still go through without these assumptions, but doing so would require carefully specifying the support of each distribution of actions and states and adding technical regularity conditions, which would necessitate several pieces of additional notation while adding little additional economic insight.

utility for experiment π_f is

$$V(u, \pi_f) = \max_{P \in \Phi(\pi_f)} \sum_{a \in A} \sum_{\omega \in \Omega} P(a, \omega) u(x(a, \omega))$$

where $\Phi(\pi_f)$ is the set of all distributions of actions and states feasible under π_f. $V(u, \pi_f)$ is also known as the *value of information* for experiment π_f given utility function u.[9] Thus, P_f is *consistent* with π_f if

$$P_f \in \operatorname*{argmax}_{P \in \Phi(\pi_f)} \sum_{a \in A} \sum_{\omega \in \Omega} P(a, \omega) u(x(a, \omega))$$

To allow the econometrician to rank distributions of actions and states based on the value of information for consistent experiments, we formally define the relation $\succsim_V$ as

$$P_g \succsim_V P_h$$

if for every u,

$$V(u, \pi_g) \geq V(u, \pi_h)$$

for every π_g consistent with P_g and every π_h consistent with P_h.

There are two features of the problem that help us in characterizing this relation. First, for a given u, every experiment π_f consistent with P_f has the same value of information, which is the expected utility provided by P_f:

$$V(u, \pi_f) = \sum_{\omega \in \Omega} \sum_{a \in A} P_f(a, \omega) u(x(a, \omega))$$

Second, we only need to consider those utility functions for which there exist experiments consistent with P_g and P_h, as the condition for $P_g \succsim_V P_h$ is trivially satisfied for such utility functions. Putting this together, $P_g \succsim_V P_h$ if and only if for all u for which there are experiments consistent with P_g and P_h,

$$\sum_{\omega \in \Omega} \sum_{a \in A} P_g(a, \omega) u(x(a, \omega)) \geq \sum_{\omega \in \Omega} \sum_{a \in A} P_h(a, \omega) u(x(a, \omega))$$

2.2 Ruling Out Utility Functions

To operationalize this restatement of $P_g \succsim_V P_h$, the econometrician needs to identify those u for which there are experiments consistent with P_g and P_h, or equivalently, to rule out those u for which there does not exist an experiment consistent with P_g or P_h.

[9] An alternative way to define the value of information is as the improvement over the utility from taking actions at prior beliefs (see Lara and Gossner 2020 and Frankel and Kamenica 2018). Since the prior is fixed across experiments in our framework, this definition would provide the same relative welfare assessments.

For utility function u, there does not exist an experiment consistent with $P_f \in \{P_g, P_h\}$ if it is possible to improve utility by making a *wholesale* switch from any chosen action to another action. If the DM can improve utility by switching to an action $b \in A$ at all posteriors where they would have chosen action some $a \in A$, this means that whatever decision rule pairs with a experiment to make P_f feasible cannot maximize expected utility.

This is demonstrated in the following simple example where utility is higher when action b is taken in state ω_1 and action a is taken in state ω_2, but action a is taken more often in state ω_1 and action b is taken more often in state ω_2:

$$u = \begin{array}{cc} \begin{array}{cc} \omega_1 & \omega_2 \end{array} & \\ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} & \begin{array}{c} a \\ b \end{array} \end{array} \;\&\; P_f = \begin{array}{cc} \begin{array}{cc} \omega_1 & \omega_2 \end{array} & \\ \begin{pmatrix} .4 & .1 \\ .1 & .4 \end{pmatrix} & \begin{array}{c} a \\ b \end{array} \end{array}$$

For this utility function, P_f is not consistent with any experiment because the DM can improve utility by making a wholesale switch to choosing b when a was chosen. When the DM chooses a, they get a utility of 1 with unconditional probability .1, but if they had chosen b instead, the DM could have gotten a utility of 1 with unconditional probability .4.

Caplin and Martin (2015) formalize this logic by introducing the No Improving Action Switches (NIAS) condition, which is a system of linear inequalities ensuring it is better not to make a wholesale switch from any chosen action a to any other action b. Utility function u satisfies NIAS for P_g and P_h if

$$\sum_{\omega \in \Omega} P_f(a, \omega) u(x(a, \omega)) \geq \sum_{\omega \in \Omega} P_f(a, \omega) u(x(b, \omega))$$

for all $P_f \in \{P_g, P_h\}$ and $a, b \in A$. The NIAS inequality for choosing a over b in P_f indicates that choosing a instead of b is optimal on average at the choice probabilities where a is chosen in P_f, given the utility of the outcomes from choosing a instead of b.

As illustrated above, if a utility function u does not satisfy NIAS, then there does not exist an experiment consistent with P_f. Caplin and Martin (2015) show that the reverse is true as well. If u satisfies NIAS, then there always exists an experiment consistent with P_f for that u.[10] Thus, the set of all u satisfying NIAS for P_g and P_h is precisely the set of u the econometrician should consider when comparing P_g and P_h.

We establish for the first time here a general feature of NIAS that enhances its analytical and computation tractability. The following lemma indicates which NIAS inequalities must hold with equality and which must hold strictly. In words, it states that for $P_f \in \{P_g, P_h\}$ the NIAS inequality for choosing a over b holds with equality if and only if the outcomes

[10]For instance, if u satisfies NIAS, then the revealed experiment for P_f is consistent with P_f for that u.

associated with that NIAS inequality (the additional probability of each outcome gained by not switching from a to b) can be expressed as a non-positive combination of the outcomes associated with other NIAS inequalities.

Lemma 1. *For every u that satisfies NIAS,*

$$\sum_{\omega \in \Omega} P_f(a,\omega) u(x(a,\omega)) = \sum_{\omega \in \Omega} P_f(a,\omega) u(x(b,\omega))$$

for $a, b \in A$ if and only if there exists a collection of N triples with generic element (P_n, a_n, b_n) having $P_n \in \{P_g, P_h\}$, $a_n \in A$, $b_n \in A$, and $(P_n, a_n, b_n) \neq (P_f, a, b)$ and non-positive weights $w_1, ..., w_N$ such that for every $x \in X$

$$\begin{aligned} &\sum_{\omega \in \Omega} P_f(a,\omega)(\mathbf{1}_{\{x(a,\omega)=x\}} - \mathbf{1}_{\{x(b,\omega)=x\}}) \\ &= \sum_{n=1}^{N} w_n * \left(\sum_{\omega \in \Omega} P_n(a_n,\omega)(\mathbf{1}_{\{x(a_n,\omega)=x\}} - \mathbf{1}_{\{x(b_n,\omega)=x\}}) \right) \end{aligned} \tag{1}$$

where $\mathbf{1}_{\{x(a,\omega)=x\}}$ is an indicator function that takes a value of 1 when the outcome from taking action a in state ω yields outcome x.

Proof. See Appendix. □

2.3 Ruling Out Utility Functions to Rank Decisions

The following two examples demonstrate that NIAS can rule out enough utility functions to allow the value of information to be ranked between two distributions of actions and states.

2.3.1 Tracking Problems

We first consider "tracking" decision problems in which the decision maker receives a state-specific outcome x_k if their action matches state ω_k and outcome x_B if they fail to match the action to the state. For this class of decision problems, the map $x(a,\omega)$ between actions, states, and outcomes is known by the econometrician and is given by

$$x(a_j, \omega_k) = \begin{cases} x_k & j = k \\ x_B & j \neq k \end{cases}$$

For the 3 action and 3 state version of this tracking problem, the map between actions, states, and outcomes can be represented as a matrix where actions a_1 to a_3 are given in the rows and states ω_1 to ω_3 are given in the columns:

$$\begin{array}{c c} \begin{array}{ccc} \omega_1 & \omega_2 & \omega_3 \end{array} & \\ \begin{pmatrix} x_1 & x_B & x_B \\ x_B & x_2 & x_B \\ x_B & x_B & x_3 \end{pmatrix} & \begin{array}{c} a_1 \\ a_2 \\ a_3 \end{array} \end{array}$$

We represent distributions of actions and states P_f and P_g as matrices where actions a_1 to a_3 are given in the rows and states ω_1 to ω_3 are given in the columns:

$$P_f = \begin{array}{c c} \begin{array}{ccc} \omega_1 & \omega_2 & \omega_3 \end{array} & \\ \begin{pmatrix} \frac{20}{100} & 0 & 0 \\ 0 & \frac{22}{100} & \frac{18}{100} \\ 0 & \frac{18}{100} & \frac{22}{100} \end{pmatrix} & \begin{array}{c} a_1 \\ a_2 \\ a_3 \end{array} \end{array} \quad \& \quad P_g = \begin{array}{c c} \begin{array}{ccc} \omega_1 & \omega_2 & \omega_3 \end{array} & \\ \begin{pmatrix} \frac{10}{100} & \frac{20}{100} & \frac{20}{100} \\ \frac{5}{100} & \frac{20}{100} & 0 \\ \frac{5}{100} & 0 & \frac{20}{100} \end{pmatrix} & \begin{array}{c} a_1 \\ a_2 \\ a_3 \end{array} \end{array}$$

In the analysis that follows, we will show that these distributions of actions and states reveal the outcomes from matching actions to states are "good" and the outcome from not matching actions is "bad". Formally, this means that for all utility functions that rationalize P_g and P_h, $u(x_k) \geq u(x_B)$ for all $k \in \{1, 2, 3\}$. Given this, P_g will be revealed to have a higher value of information because the DM matches actions to states more often in every state.

However, while the DM will be revealed to be perfectly informed when taking action a_1 for P_g, the DM will be revealed to be somewhat better informed about the matching state when taking actions a_2 and a_3 for P_g. As a result, the DM will not be revealed to have better informed actions for either P_f or P_g. In terms of signal structures, it is as if P_g is generated by a DM with a signal structure that is perfectly informative about whether the state is ω_1, but is not as informative about the other states as the signal structure that generated P_h.

Without loss of generality we set $u(x_B) = 0$, so we can compute the value of information for P_g as

$$\frac{20}{100}u(x_0) + \frac{22}{100}u(x_1) + \frac{22}{100}u(x_2)$$

and for P_g as

$$\frac{10}{100}u(x_0) + \frac{20}{100}u(x_1) + \frac{20}{100}u(x_2)$$

Clearly, if $u(x_1)$, $u(x_2)$, or $u(x_3)$ are revealed to be non-negative for all rationalizing utility functions, then P_g is revealed to have a higher value of information.

The fact that these utilities are non-negative can be established through the NIAS inequalities for P_g. The NIAS inequality for P_g for a_1 chosen over action a_2 gives $u(x_1) \geq 0$ because

$$\begin{aligned}\sum_{\omega\in\Omega} P_f(a_1,\omega)u(x(a_1,\omega)) &\geq \sum_{\omega\in\Omega} P_f(a_1,\omega)u(x(a_2,\omega)) \\ \frac{20}{100}u(x_1) &\geq \frac{20}{100}u(x_B) = 0\end{aligned}$$

Likewise, the NIAS inequality for P_g for a_2 chosen over action a_1 gives $u(x_2) \geq 0$, and the NIAS inequality for a_3 chosen over action a_1 gives $u(x_3) \geq 0$.

2.3.2 Problems with Distinct Outcomes

Next, we consider a common class of decision problems in which every action yields a distinct outcome in every state, so that $x(a,\omega) \neq x(b,\nu)$ if $a \neq b$ or $\omega \neq \nu$. As noted previously, this case covers the situation where the econometrician does not know the map between actions, states, and outcomes.

For the 3 action and 3 state version of this tracking problem, the map between actions, states, and outcomes can be represented as a matrix where actions a_1 to a_3 are given in the rows and states ω_1 to ω_3 are given in the columns:

$$\begin{array}{cc} \begin{array}{ccc}\omega_1 & \omega_2 & \omega_3\end{array} & \\ \begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33}\end{pmatrix} & \begin{array}{c} a_1 \\ a_2 \\ a_3\end{array}\end{array}$$

One example of this is given by

$$P_f = \begin{array}{cc} \begin{array}{ccc}\omega_1 & \omega_2 & \omega_3\end{array} & \\ \begin{pmatrix} \frac{24}{72} & 0 & 0 \\ 0 & \frac{16}{72} & \frac{8}{72} \\ 0 & \frac{8}{72} & \frac{16}{72}\end{pmatrix} & \begin{array}{c} a_1 \\ a_2 \\ a_3\end{array}\end{array} \quad \& \quad P_g = \begin{array}{cc} \begin{array}{ccc}\omega_1 & \omega_2 & \omega_3\end{array} & \\ \begin{pmatrix} \frac{12}{72} & \frac{6}{72} & \frac{6}{72} \\ \frac{6}{72} & \frac{5}{72} & \frac{13}{72} \\ \frac{6}{72} & \frac{13}{72} & \frac{5}{72}\end{pmatrix} & \begin{array}{c} a_1 \\ a_2 \\ a_3\end{array}\end{array}$$

Like the tracking example, these distributions of actions and states will reveal that the DM prefers the outcome obtained when choosing action a_1 when the state is ω_1, prefers the outcomes obtained when choosing a_2 and a_3 in the other states, and is perfectly informed when taking action a_1. Once again, it is as if for P_g the DM gets a signal realization that is perfectly informative of whether the state is ω_1, so knows to take action a_1 if the state is ω_1 and not to choose action a_1 otherwise.

However, unlike the tracking example, this P_g and P_h reveal that the utility obtained from taking actions a_2 and a_3 is the same in every state.[11] This follows from the fact that the NIAS inequalities for a_2 chosen over a_3 and a_3 chosen over a_2 hold with equality for both P_g and P_h, which is a consequence of Lemma 1. Lemma 1 states that an NIAS inequality is equal to zero if and only if that NIAS inequality can be expressed as a non-positive combination of other NIAS inequalities. For example, the NIAS inequality for a_2 chosen over a_3 for P_g is $\frac{16}{72}(u(a_2,\omega_2) - u(a_3,\omega_2)) + \frac{8}{72}(u(a_2,\omega_3) - u(a_3,\omega_3)) \geq 0$.[12] The negative of this can be obtained by simply adding together the outcome lotteries from the NIAS inequalities for a_3 chosen over a_2 for P_g, for a_2 chosen over a_3 for P_g, and for a_3 chosen over a_2 for P_g.

Given that the NIAS inequalities for a_2 chosen over a_3 and a_3 chosen over a_2 hold with equality for P_g, the utility differences between a_2 and a_3 in ω_2 and the utility differences between a_2 and a_3 in ω_3 are both equal to 0 because those NIAS inequalities say

$$\frac{16}{72}\left(u(a_2,\omega_2) - u(a_3,\omega_2)\right) + \frac{8}{72}\left(u(a_2,\omega_3) - u(a_3,\omega_3)\right) = 0$$

and

$$-\frac{8}{72}\left(u(a_2,\omega_2) - u(a_3,\omega_2)\right) - \frac{16}{72}\left(u(a_2,\omega_3) - u(a_3,\omega_3)\right) = 0$$

which is only possible if $u(a_2,\omega_2) - u(a_3,\omega_2) = 0$ and $u(a_2,\omega_3) - u(a_3,\omega_3) = 0$. Likewise, given that the NIAS inequalities for a_2 chosen over a_3 and a_3 chosen over a_2 hold with equality for P_g, the utility difference between a_2 and a_3 in ω_1 is also equal to 0. Thus, the utility from taking a_2 is the same as the utility from taking a_3 in every state.

Given this, the value of information is higher for P_g if

$$\begin{aligned} & \frac{12}{72}(u(a_1,\omega_1) - u(a_2,\omega_1)) \\ + \; & \frac{6}{72}(u(a_2,\omega_2) - u(a_1,\omega_2) + u(a_2,\omega_3) - u(a_1,\omega_3)) \geq 0 \end{aligned}$$

To show that this holds, we first note that $u(a_1,\omega_1) \geq u(a_2,\omega_1)$ (the DM preferring to take action a_1 in state ω_1) follows directly from the NIAS inequality for a_1 chosen over a_2 for P_g. Second, because a_2 and a_3 give the same utility in every state, the NIAS inequalities for a_2 chosen over a_1 and a_3 chosen over a_1 for P_g yield

$$\frac{16}{72}(u(a_2,\omega_2) - u(a_1,\omega_2)) + \frac{8}{72}(u(a_2,\omega_3) - u(a_1,\omega_3)) \geq 0$$

and

$$\frac{8}{72}(u(a_2,\omega_2) - u(a_1,\omega_2)) + \frac{16}{72}(u(a_2,\omega_3) - u(a_1,\omega_3)) \geq 0$$

[11] This example can also be generalized to any version of this problem with arbitrarily many actions and at least as many states as actions.

[12] Given that there are no common outcomes across states and actions in this decision problem, we will shorten $u(x(a,\omega))$ to $u(a,\omega)$.

Adding these together gives

$$u(a_2, \omega_2) - u(a_1, \omega_2) + u(a_2, \omega_3) - u(a_1, \omega_3) \geq 0$$

With this, we have that P_g provides a higher value of information.

3 Characterizing the Relation

Is there a general approach to checking whether there are enough restrictions on u to ensure $P_g \succsim_V P_h$? We produce a necessary and sufficient condition for $P_g \succsim_V P_h$ by moving fully to the space of probabilities and probability differences over outcomes. There are three features that make this space important. First, it allows geometric representation of the NIAS inequalities. Second, it allows identification of all utility functions that satisfy these inequalities. Third, it identifies all difference in outcome lotteries that are guaranteed to raise utility (which allows us to identify $\succsim_V$). Because of this geometric representation, we can reduce $P_g \succsim_V P_h$ to a single system of linear equations.

3.1 Ruling Out Utility Functions Geometrically

First, each NIAS inequality can be represented as an M-dimensional vector $\vec{d}_f(a, b)$ that gives the outcome lottery gained from not making a wholesale switch from action a to action b for P_f. In other words, the additional probability of receiving each outcome from not making this wholesale switch. Element m of this vector gives the additional probability of receiving outcome x_m in X from not making a wholesale switch from action a to action b for P_g, which is

$$\sum_{\omega \in \Omega} P_f(a, \omega)(\mathbf{1}_{\{x(a,\omega)=x_m\}} - \mathbf{1}_{\{x(b,\omega)=x_m\}})$$

where $\mathbf{1}_{\{x(a,\omega)=x_m\}}$ is an indicator function that takes a value of 1 when the outcome from taking action a in state ω yields outcome x_m. The convex cone D formed by all NIAS inequalities is

$$D = \{\alpha_1 \vec{d}_{f_1}(a_1, b_1) + ... + \alpha_N \vec{d}_{f_N}(a_N, b_N) | \alpha_n \in \mathbb{R}_+, f_n \in \{g, h\}, a_n, b_n \in A\}$$

A utility function can be represented as an M-dimensional vector $\vec{u}$ where element m gives the utility of outcome x_m. A utility vector $\vec{u}$ satisfies NIAS if $\vec{d} \bullet \vec{u} \geq 0$ for every vector $\vec{d} \in D$. We call the convex cone formed by all $\vec{u}$ that satisfy NIAS the *NIAS Utility Cone.*

We illustrate this with simple decision problem that has a safe action, where the map between states, actions, and outcomes is given by

$$\begin{array}{ccc} \omega_1 & \omega_2 & \\ \multicolumn{2}{c}{\begin{pmatrix} x_1 & x_2 \\ x_3 & x_3 \end{pmatrix}} & \begin{array}{c} a \\ b \end{array} \end{array}$$

Imagine the following distribution of actions and states, which appeared in the introduction.

$$P_g = \begin{array}{ccc} \omega_1 & \omega_2 & \\ \multicolumn{2}{c}{\begin{pmatrix} .4 & .1 \\ .1 & .4 \end{pmatrix}} & \begin{array}{c} a \\ b \end{array} \end{array} \qquad \& \qquad P_h = \begin{array}{ccc} \omega_1 & \omega_2 & \\ \multicolumn{2}{c}{\begin{pmatrix} .15 & .05 \\ .35 & .45 \end{pmatrix}} & \begin{array}{c} a \\ b \end{array} \end{array}$$

Choosing a in P_g gets x_1 and x_2 with unconditional probabilities .4 and .1. Choosing b gets x_3. Hence sticking with a over b yields,

$$\vec{d_g}(a, b) = (.4, .1, -.5)$$

Likewise sticking with b over a in P_g gets x_3 rather than x_1 with unconditional probability .1 and x_2 with unconditional probability .4:

$$\vec{d_g}(b, a) = (-.1, -.4, .5)$$

Analogously,

$$\vec{d_h}(a, b) = (.15, .05, -.2)$$
$$\vec{d_h}(b, a) = (-.35, -.45, .8)$$

NIAS identifies rationalizing utility functions as all that have (weakly) positive dot products with all of these vectors. This can be visualized in two dimensions by normalizing $u(x_3) = 0$. With this normalization, the (x_1, x_2) space can illustrate both D and the NIAS Utility Cone, which is given in Figure 1

3.2 Ranking Decisions Geometrically

Let $\vec{d}(g, h)$ be an M-dimensional vector that gives the outcome lottery gained from encountering P_g instead of P_h. In other words, the additional probability of receiving each outcome from P_g. For outcome x_m in X, this is

$$\sum_{a\in A}\sum_{\omega\in\Omega}(P_g(a, \omega) - P_h(a, \omega))\mathbf{1}_{\{x(a,\omega)=x\}}$$

P_g being revealed to have a higher value of information ($P_g \succsim_V P_h$) is equivalent to $\vec{d}(g, h) \bullet \vec{u} \geq 0$ for all $\vec{u}$ in the NIAS Utility Cone. Thus, $P_g \succsim_V P_h$ if and only if the vector

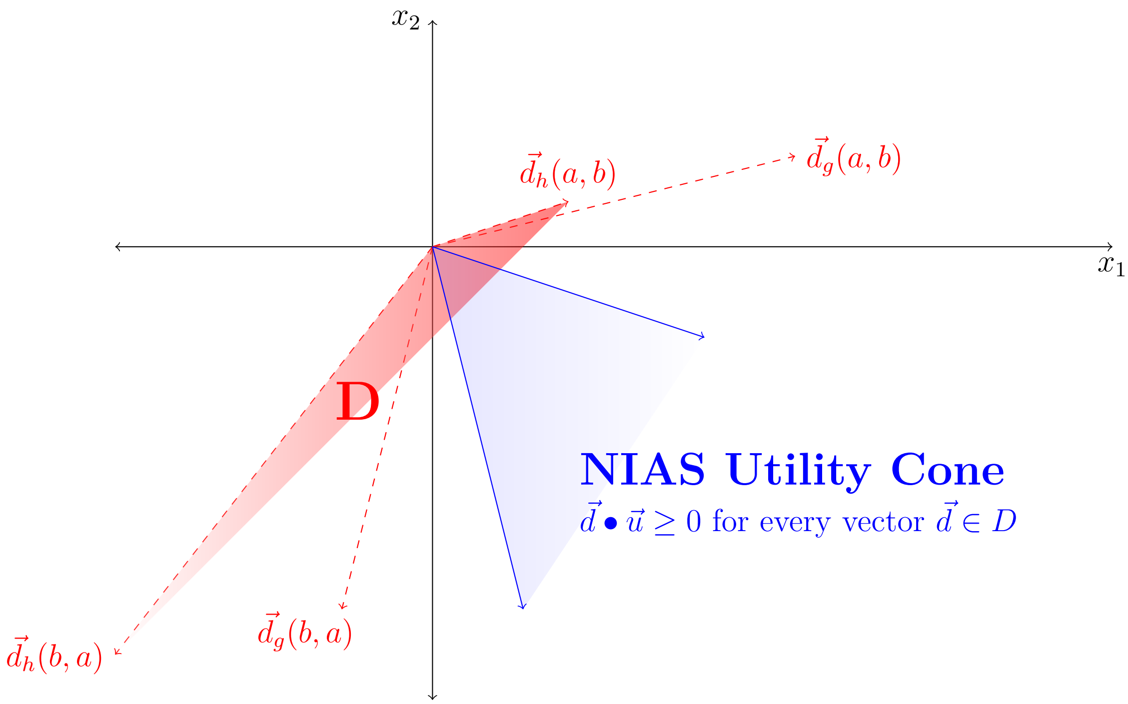

Figure 1: Illustration of the geometric structure of ruling out utility functions using NIAS.

$\vec{d}(g,h)$ is in D because a vector is in D if and only if it has a non-negative dot product with all $\vec{u}$ in the NIAS Utility Cone.

Finally, because $\vec{d}(g,h)$ is in D if and only if it is a non-negative weighted average of vectors in D, a necessary and sufficient condition for $P_g \succsim_V P_h$ corresponds to the outcome lottery gained from P_g being a non-negative weighted average of the outcome lotteries gained from not making wholesale switches from any action for either P_f or P_g.

Returning to the example, P_g yields $(.4, .1, .5)$ and P_h yields $(.15, .05, .8)$. Hence, $\vec{d}(g,h) = (.25, .05, -.3)$. As illustrated in Figure 2, $\vec{d}(g,h)$ is in D, so it has a positive dot product with the entire NIAS Utility Cone, so P_g and P_h are ranked by $\succsim_V$.

3.3 General Condition for Ranking Decisions

The condition that $\vec{d}(g,h)$ is in D, which we call *Decision Improvement without Action Switches (DISI)* for P_g over P_h, is defined for a weighting function $t_{gh} : A \times A \to \mathbb{R}_+$, which provides these non-negative weights.

Condition 1 (Decision Improvement without Action Switches (**DISI**))**.** *Weighting function*

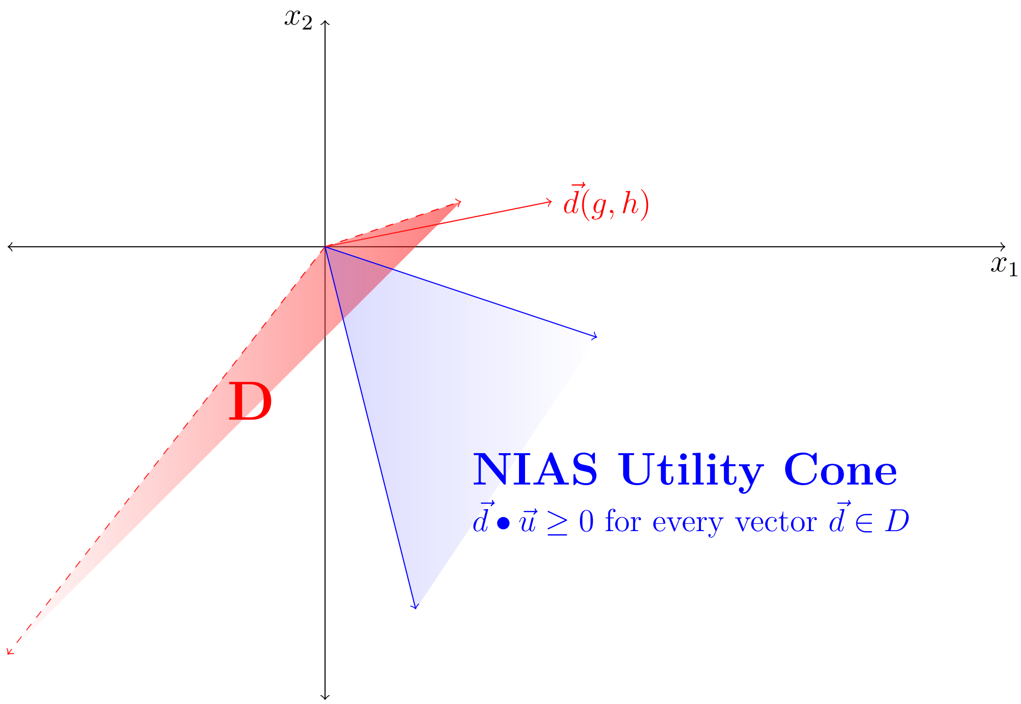

Figure 2: Illustration of the geometric structure of ranking decisions.

$t_{gh} : A \times A \to \mathbb{R}_+$ *satisfies DIAS for* P_g *over* P_h *if for every* $x \in X$

$$\sum_{P_f \in \{P_g, P_h\}} \sum_{a \in A} \sum_{b \in A} \sum_{\omega \in \Omega} P_f(a,\omega)(\boldsymbol{1}_{\{x(a,\omega)=x\}} - \boldsymbol{1}_{\{x(b,\omega)=x\}}) t_{gh}(a,b)$$
$$= \sum_{a \in A} \sum_{\omega \in \Omega} (P_g(a,\omega) - P_h(a,\omega)) \boldsymbol{1}_{\{x(a,\omega)=x\}}$$

The following theorem formally shows that DISI provides a necessary and sufficient condition to reveal that every experiment consistent with one distribution of actions and states has a higher value of information than every experiment consistent with one distribution of actions and states. After restating NIAS and DISI in terms of matrix multiplication, the proof of this theorem follows as a direct consequence of Farkas lemma (Farkas 1902).

Theorem 1. $P_g \succsim_V P_h$ *if and only if there exists a weighting function* t_{gh} *that satisfies DISI for* P_g *over* P_h.

Proof. See Appendix. □

This theorem has an economic interpretation in terms of preferences over outcome lotteries. DISI states that the difference in the outcome lotteries offered by the distributions of actions and states can be represented as the difference in two compound lotteries: one composed of the outcome lotteries from taking each action a for P_g and the other composed of outcome lotteries from taking each action b with the same probability as a. These compound lotteries have the same weights for all $P_f \in \{P_g, P_h\}$ and $a, b \in A$, which are given by a normalized version of t_{gh}. Because NIAS is satisfied, there exists a preference relation over lotteries such that every element of one compound lottery is weakly preferred to every element in the other compound lottery. Since all elements of the two compound lotteries are preference ordered, the compound lotteries are also preference ordered, which means the outcome lotteries given by each distribution of actions and states are as well.

3.4 Testability

Determining whether or not there exists a t_f that satisfies DISI corresponds to determining whether there is a solution to a system of linear equations, so it is simple to check if distributions of actions and states are welfare ranked. We provide MATLAB computer programs that determine whether a solution to this linear system exists for a given set of data.[13]

Also, there are settings where some options are clearly dominant, and NIAS and DISI can be easily amended to account for these additional restrictions. Say for example, that outcome x_1 clearly dominates outcome x_2. This restriction on utility can be incorporated into NIAS by generating an additional linear inequality given by

$$\sum_{x \in X} \left(\mathbf{1}_{x=x_1} - \mathbf{1}_{x=x_2}\right) u(x) \geq 0$$

Clearly, this restriction on the set of admissible utility functions can only reduce the set of u that satisfy NIAS.

Although DISI in not expressed in terms of utility, the dominance of outcome x_1 over outcome x_2 can be incorporated into DISI for P_g by requiring that, in addition to the weighting function t_{gh}, there is exists a non-negative t that solves

$$\begin{aligned} & \sum_{P_f \in \{P_g, P_h\}} \sum_{a \in A} \sum_{b \in A} \sum_{\omega \in \Omega} P_f(a, \omega)(\mathbf{1}_{\{x(a,\omega)=x\}} - \mathbf{1}_{\{x(b,\omega)=x\}}) t_{gh}(a, b) \\ + \quad & \left(\mathbf{1}_{x=x_1} - \mathbf{1}_{x=x_2}\right) t \\ = \quad & \sum_{a \in A} \sum_{\omega \in \Omega} (P_g(a, \omega) - P_h(a, \omega)) \mathbf{1}_{\{x(a,\omega)=x\}} \end{aligned}$$

for every $x \in X$. If t is equal zero, this reduces to the requirement for DISI, so this addition can only increase the proportion of P_g and P_h where there exists a weighting function t_{gh} that satisfies DISI. Sensibly, knowledge about dominance improves our ability to rank decisions according to their welfare.

[13] Programs available at https://github.com/danieljosephmartin/framinginformationwelfare.git.

4 Blackwell Within Our Framework

In this section, we draw a connection between our relation and Blackwell's informativeness order over experiments (Blackwell 1953). Because experiments are unknown in our framework, we define a *revealed experiment* $\bar{\pi}_f$ for each $P_f \in \{P_g, P_h\}$ in order to make this connection. This is the experiment if we assume there is a single posterior at which each action is taken, so it reflects the minimal informativeness the DM can have given their actions.

Importantly, the revealed experiment can inferred directly from the distribution of actions and states. The building block of the revealed experiment are the *revealed posteriors*, which are the average distributions of states when each actions is taken. For action a in P_f, the revealed posterior is defined by

$$\bar{\gamma}_f^a(\omega) := P_f(\omega|a)$$

So the revealed experiment can be read directly off of the data:

$$\bar{\pi}_f(\bar{\gamma}_f^a, \omega) = P_f(a, \omega)$$

For example, consider the distributions of actions and states given previously

$$P_f = \begin{array}{cc} \begin{array}{cc} \omega_1 & \omega_2 \end{array} & \\ \begin{pmatrix} .4 & .1 \\ .1 & .4 \end{pmatrix} & \begin{array}{c} a \\ b \end{array} \end{array}$$

The revealed posteriors are given by

$$\bar{\gamma}_f^a = \begin{pmatrix} .8 \\ .2 \end{pmatrix} \begin{array}{c} \omega_a \\ \omega_b \end{array} \quad \& \quad \bar{\gamma}_f^b = \begin{pmatrix} .2 \\ .8 \end{pmatrix} \begin{array}{c} \omega_a \\ \omega_b \end{array}$$

And the revealed experiment is simply

$$\bar{\pi}_f = \begin{array}{cc} \begin{array}{cc} \omega_1 & \omega_2 \end{array} & \\ \begin{pmatrix} .4 & .1 \\ .1 & .4 \end{pmatrix} & \begin{array}{c} \bar{\gamma}_f^a \\ \bar{\gamma}_f^b \end{array} \end{array}$$

The point of connection with Blackwell's Theorem is that a garbling of revealed experiments is sufficient for $\succsim_V$. By Blackwell (1953), if the revealed experiment $\bar{\pi}_h$ is a garbling of revealed experiment $\bar{\pi}_g$, then $V(u, \bar{\pi}_g) \geq V(u, \bar{\pi}_h)$ for every u. Thus, if the revealed experiment $\bar{\pi}_h$ is a garbling of revealed experiment $\bar{\pi}_g$, then clearly $V(u, \bar{\pi}_g) \geq V(u, \bar{\pi}_h)$ for every u that satisfies NIAS for P_g and P_h. By Caplin and Martin (2015), if u satisfies NIAS, $\bar{\pi}_g$ is consistent with P_g and $\bar{\pi}_h$ is consistent with P_h. Finally, because of value equivalence, this means $V(u, \pi_g) \geq V(u, \pi_h)$ for all π_g that are consistent with P_g and all π_h that are consistent with P_h, as desired.

4.1 Why Garbling is Not Necessary

However, for $P_g \succsim_V P_h$ to hold, it is not necessary that $\bar{\pi}_h$ be garbing of $\bar{\pi}_g$. Blackwell (1953) showed that $\bar{\pi}_h$ is a garbing of $\bar{\pi}_g$ if and only if $\bar{\pi}_g$ has a higher value of information than $\bar{\pi}_h$ for all possible u. But this is unnecessarily strong, as for $P_g \succsim_V P_h$ we only need that $\bar{\pi}_g$ has a higher value of information than $\bar{\pi}_h$ for all u that satisfy NIAS.

The divergence between the garbling order and $\succsim_V$ illustrates the importance of ruling out utility functions where there are no experiments consistent with P_g and P_h. It can be that $\bar{\pi}_g$ has a higher value of information than $\bar{\pi}_h$ for all u that satisfy NIAS, but $\bar{\pi}_h$ has a higher value of information than $\bar{\pi}_g$ for a u that do not satisfy NIAS.

This is easy to show in trivial cases such as when there are just two possible outcomes or all actions give the same outcome in some state. However, it can also occur outside of these trivial cases. This includes the examples given previously for decision problems where state-dependent outcomes are obtained when the action matches the state and where every action yields a distinct outcome in every state.

To illustrate this, we reconsider the 3 action and 3 state tracking problem analyzed previously:

$$\begin{array}{cc} \begin{array}{ccc} \omega_1 & \omega_2 & \omega_3 \end{array} & \\ \begin{pmatrix} x_1 & x_B & x_B \\ x_B & x_2 & x_B \\ x_B & x_B & x_3 \end{pmatrix} & \begin{array}{c} a_1 \\ a_2 \\ a_3 \end{array} \end{array}$$

As before, the distributions of actions and states were:

$$P_g = \begin{array}{cc} \begin{array}{ccc} \omega_1 & \omega_2 & \omega_3 \end{array} & \\ \begin{pmatrix} \frac{20}{100} & 0 & 0 \\ 0 & \frac{22}{100} & \frac{18}{100} \\ 0 & \frac{18}{100} & \frac{22}{100} \end{pmatrix} & \begin{array}{c} a_1 \\ a_2 \\ a_3 \end{array} \end{array} \quad \& \quad P_h = \begin{array}{cc} \begin{array}{ccc} \omega_1 & \omega_2 & \omega_3 \end{array} & \\ \begin{pmatrix} \frac{10}{100} & \frac{20}{100} & \frac{20}{100} \\ \frac{5}{100} & \frac{20}{100} & 0 \\ \frac{5}{100} & 0 & \frac{20}{100} \end{pmatrix} & \begin{array}{c} a_1 \\ a_2 \\ a_3 \end{array} \end{array}$$

We showed that the DM was revealed to have a higher value of information for this P_g than P_h by showing that for all u that satisfy NIAS, the outcomes from matching actions to states have a weakly positive utility.

But while the value of information is higher for $\bar{\pi}_g$ for all u that satisfy NIAS, there are many u that do not satisfy NIAS for which the value of information is higher for $\bar{\pi}_h$ instead of $\bar{\pi}_g$. For example, set $u(x_1) = -.0001$, $u(x_2) = 1$, and $u(x_3) = 1$.

As established before, NIAS is not satisfied for this utility function because $u(x_1) < 0$. But because NIAS is not satisfied for this u, there are no experiments that are consistent with P_g and P_h. Thus, P_g and P_h do not maximized expected utility among the set of distributions of actions and states feasible under $\bar{\pi}_h$ and $\bar{\pi}_g$.

As a consequence, to determine whether the value of information is higher for $\bar{\pi}_h$ or $\bar{\pi}_g$ for this u, we have to determine an optimal distribution of actions and states under these experiments for this u. In this case, an optimal decision rule will try to match actions in ω_2 and ω_3. Avoiding matching actions in ω_1 is much less important.

To determine an optimal decision rule, we need to look at the revealed posteriors for each revealed experiment. The revealed posteriors for P_g are

$$\gamma_g^{a_1} = \begin{array}{c} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \end{array} \begin{array}{c} \omega_1 \\ \omega_2 \\ \omega_3 \end{array} \quad \& \quad \gamma_g^{a_2} = \begin{pmatrix} 0 \\ \frac{22}{40} \\ \frac{18}{40} \end{pmatrix} \begin{array}{c} \omega_1 \\ \omega_2 \\ \omega_3 \end{array} \quad \& \quad \gamma_g^{a_3} = \begin{pmatrix} 0 \\ \frac{18}{40} \\ \frac{22}{40} \end{pmatrix} \begin{array}{c} \omega_1 \\ \omega_2 \\ \omega_3 \end{array}$$

and the revealed posteriors for P_h are

$$\gamma_h^{a_1} = \begin{pmatrix} \frac{10}{50} \\ \frac{20}{50} \\ \frac{20}{50} \end{pmatrix} \begin{array}{c} \omega_1 \\ \omega_2 \\ \omega_3 \end{array} \quad \& \quad \gamma_h^{a_2} = \begin{pmatrix} \frac{5}{25} \\ \frac{20}{25} \\ 0 \end{pmatrix} \begin{array}{c} \omega_1 \\ \omega_2 \\ \omega_3 \end{array} \quad \& \quad \gamma_h^{a_3} = \begin{pmatrix} \frac{5}{25} \\ 0 \\ \frac{20}{25} \end{pmatrix} \begin{array}{c} \omega_1 \\ \omega_2 \\ \omega_3 \end{array}$$

These revealed posteriors are presented in Figure 3, which shows the probability of just two states (ω_2 and ω_3) as the third is constrained to be the remainder. For example, the revealed posterior for action a_1 for P_g gives a likelihood of zero to these states, so it puts all likelihood on state ω_1.

Looking at the revealed posteriors, it is clear that it is an optimal decision rule to choose a_1 for $\gamma_g^{a_1}$, a_2 for $\gamma_g^{a_2}$, and a_3 for $\gamma_g^{a_3}$ and to choose a_2 for $\gamma_h^{a_1}$, a_2 for $\gamma_h^{a_2}$, and a_3 for $\gamma_h^{a_3}$. This only departs from the chosen action for one revealed posterior, but this is enough to flip the order of the value of information.

To see that the value of information is higher for $\bar{\pi}_h$ with this optimal decision rule, note that the revealed experiment is:

$$\bar{\pi}_g = \begin{array}{c} \begin{array}{ccc} \omega_1 & \omega_2 & \omega_3 \end{array} \\ \begin{pmatrix} \frac{20}{100} & 0 & 0 \\ 0 & \frac{22}{100} & \frac{18}{100} \\ 0 & \frac{18}{100} & \frac{22}{100} \end{pmatrix} \end{array} \begin{array}{c} \\ \gamma_g^{a_1} \\ \gamma_g^{a_2} \\ \gamma_g^{a_3} \end{array} \quad \& \quad \bar{\pi}_h = \begin{array}{c} \begin{array}{ccc} \omega_1 & \omega_2 & \omega_3 \end{array} \\ \begin{pmatrix} \frac{10}{100} & \frac{20}{100} & \frac{20}{100} \\ \frac{5}{100} & \frac{20}{100} & 0 \\ \frac{5}{100} & 0 & \frac{20}{100} \end{pmatrix} \end{array} \begin{array}{c} \\ \gamma_h^{a_1} \\ \gamma_h^{a_2} \\ \gamma_h^{a_3} \end{array}$$

For $\bar{\pi}_g$, the optimal decision rule above produces an expected utility of $\frac{44}{100}$, and for $\bar{\pi}_h$, the optimal decision rule above produces an expected utility of $\frac{60}{100}$.

Because this u ranks the value of information for $\bar{\pi}_h$ over the value of information for $\bar{\pi}_g$, $\bar{\pi}_h$ cannot be a garbling of $\bar{\pi}_g$. This is shown clearly in Figure 3, as the revealed posteriors

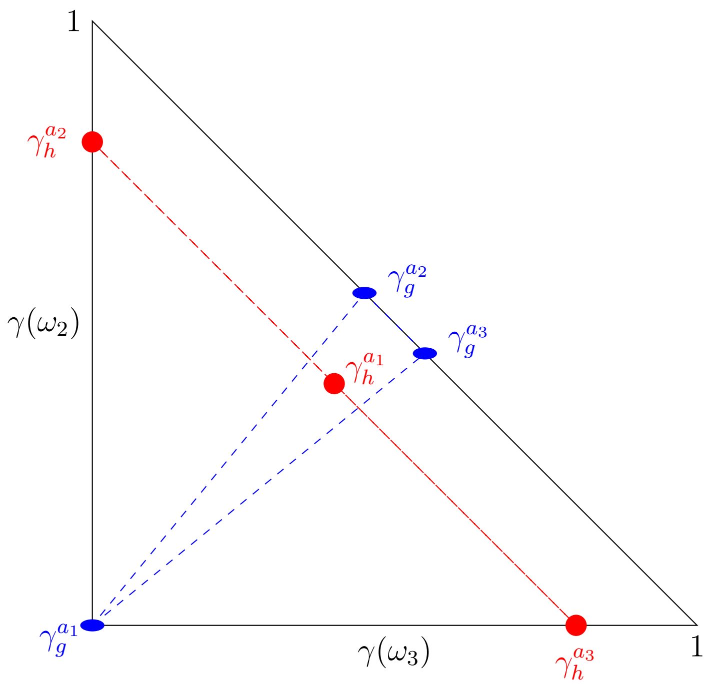

Figure 3: Revealed posteriors for tracking example. Blue ovals give revealed posteriors for P_g, red dots give revealed posteriors for P_h, and dashed lines provided the convex hull of the revealed posteriors for each experiment.

for one set of decisions do not lie in the revealed posteriors for the other. The revealed posteriors indicate that for P_g the DM is better informed on average about the matching state when taking action a_1, but less informed on average about the matching state when taking actions a_2 and a_3 for P_g.

While this u ranks the value of information for $\bar{\pi}_h$ over the value of information for $\bar{\pi}_g$, we showed previously for all u that satisfy NIAS, the reverse order holds. For all u that satisfy NIAS, the value of information for $\bar{\pi}_g$ is higher than the value of information for $\bar{\pi}_h$, so we can still conclude that $P_g \succsim P_h$.

5 Related Literature

Our work is mostly closely related to three papers. First, our relation draws natural parallels to the seminal informativeness relation provided by Blackwell (1953). In comparing experiments, Blackwell holds fixed π_g and π_h, and for each u, evaluates the expected utility provided by all P_g and P_h consistent with those experiments. On the other hand, in comparing decisions, we hold fixed P_g and P_h, and for each u, evaluate the expected utility provided by all π_g and π_h consistent with those decisions. This change in perspective produces an important technical difference that is illustrated in the preceding section. For Blackwell's relation, every u must be considered because every experiment π_g has a consistent P_g for every u. However, every u does not need to be considered for our relation because P_g may not have a consistent π_g for some u.

Second, like our paper, Lu (2016) also orders unknown experiments, but his characterization requires much more than just the joint distribution of actions and states P_g and P_h. He also requires that the econometrician observes richer "test functions" F_g and F_h, which are not naturally occurring. A test function F indicates how often the actions are chosen when the set of actions is paired with every possible mixture between the best and worst action.

Third, our paper builds on the results of Caplin and Martin (2015), and we provide three innovations relative that work. First, we provide a novel geometric representation for using NIAS to rule out utility functions. Second, we provide a new result showing when NIAS holds strictly and weakly, which enhances the analytical and computational tractability of NIAS. Third, and most importantly, we provide an entirely new application of NIAS by showing exactly when it can rule out enough utility functions to allow the value of information to be ranked between two unknown experiments.

References

Abeler, Johannes and Simon Jager (2015). "Complex tax incentives". In: *American Economic Journal: Economic Policy* 7.3, pp. 1–28.

Apesteguia, Jose and Miguel A Ballester (2018). "Monotone stochastic choice models: The case of risk and time preferences". In: *Journal of Political Economy* 126.1, pp. 74–106.

Bhargava, Saurabh, George Loewenstein, and Justin Sydnor (2017). "Choose to lose: Health plan choices from a menu with dominated option". In: *The Quarterly Journal of Economics* 132.3, pp. 1319–1372.

Blackwell, David (1953). "Equivalent comparisons of experiments". In: *Annals of Mathematical Statistics*, pp. 265–272.

Bursztyn, Leonardo, Aakaash Rao, Christopher Roth, and David Yanagizawa-Drott (2020). "Misinformation during a pandemic". In: *University of Chicago, Becker Friedman Institute for Economics Working Paper* 2020-44.

Caplin, Andrew and Daniel Martin (2015). "A testable theory of imperfect perception". In: *The Economic Journal* 125.582, pp. 184–202.

Carpenter, Jeffrey et al. (2019). "Choice Architecture to Improve Financial Decision Making". In: *Review of Economics and Statistics*, pp. 1–52.

Carrera, Mariana and Sofia Villas-Boas (2015). *Generic aversion and observational learning in the over-the-counter drug market*. Department of Agricultural and Resource Economics, UCB, UC Berkeley.

Choi, James J, David Laibson, and Brigitte C Madrian (2009). "Reducing the complexity costs of 401 (k) participation through quick enrollment". In: *Developments in the Economics of Aging*. University of Chicago Press, pp. 57–82.

Dranove, David and Ginger Zhe Jin (2010). "Quality disclosure and certification: Theory and practice". In: *Journal of Economic Literature* 48.4, pp. 935–63.

Ericson, Keith M Marzilli and Amanda Starc (2016). "How product standardization affects choice: Evidence from the Massachusetts Health Insurance Exchange". In: *Journal of Health Economics* 50, pp. 71–85.

Esponda, Ignacio and Emanuel Vespa (2016). *Contingent preferences and the sure-thing principle: Revisiting classic anomalies in the laboratory*. Tech. rep. working paper.

Farkas, Julius (1902). "Uber die theorie der einfachen ungeichungen". In: *J. Reine Angew. Math.* 124, pp. 1–24.

Frankel, Alexander and Emir Kamenica (2018). "Quantifying information and uncertainty". In: *American Economic Review*.

Handel, Benjamin (2013). "Adverse selection and inertia in health insurance markets: When nudging hurts". In: *American Economic Review* 103.7, pp. 2643–82.

Hastings, Justine S and Lydia Tejeda-Ashton (2008). *Financial literacy, information, and demand elasticity: Survey and experimental evidence from Mexico*. Tech. rep. National Bureau of Economic Research.

Jin, Ginger Zhe, Michael Luca, and Daniel Martin (2015). *Is no news (perceived as) bad news? An experimental investigation of information disclosure*. Tech. rep. National Bureau of Economic Research.

— (2018). *Complex disclosure*. Tech. rep. National Bureau of Economic Research.

Kamenica, Emir (2019). "Bayesian persuasion and information design". In: *Annual Review of Economics* 11, pp. 249–272.

Kamenica, Emir and Matthew Gentzkow (2011). "Bayesian persuasion". In: *American Economic Review* 101.6, pp. 2590–2615.

Lara, Michel De and Olivier Gossner (2020). "Payoffs-Beliefs Duality and the Value of Information". In: *SIAM Journal on Optimization* 30.1, pp. 464–489.

Lu, Jay (2016). "Random choice and private information". In: *Econometrica* 84.6, pp. 1983–2027.

Manzini, Paola and Marco Mariotti (2014). "Stochastic choice and consideration sets". In: *Econometrica* 82.3, pp. 1153–1176.

McFadden, Daniel (1973). "Conditional logit analysis of qualitative choice behavior". In: *Frontiers in Econometrics*, pp. 105–142.

Milgrom, Paul R (1981). "Good news and bad news: Representation theorems and applications". In: *The Bell Journal of Economics*, pp. 380–391.

6 Appendix

Lemma 1: For every u that satisfies NIAS,

$$\sum_{\omega\in\Omega} P_f(a,\omega)u(x(a,\omega)) = \sum_{\omega\in\Omega} P_f(a,\omega)u(x(b,\omega)) \tag{2}$$

for $a, b \in A$ if and only if there exists a collection of N triples with generic element (P_n, a_n, b_n) having $P_n \in \{P_g, P_h\}$, $a_n \in A$, $b_n \in A$, and $(P_n, a_n, b_n) \neq (P_f, a, b)$ and non-positive weights $w_1, ..., w_N$ such that for every $x \in X$

$$\begin{aligned} &\sum_{\omega\in\Omega} P_f(a,\omega)(\mathbf{1}_{\{x(a,\omega)=x\}} - \mathbf{1}_{\{x(b,\omega)=x\}}) \\ =&\sum_{n=1}^{N} w_n * \left(\sum_{\omega\in\Omega} P_n(a_n,\omega)(\mathbf{1}_{\{x(a_n,\omega)=x\}} - \mathbf{1}_{\{x(b_n,\omega)=x\}})\right) \end{aligned} \tag{3}$$

Proof of Lemma 1:

Proof. First, for any u that satisfies NIAS, by definition

$$\sum_{x\in X}\left(\sum_{\omega\in\Omega} P_f(a,\omega)(\mathbf{1}_{\{x(a,\omega)=x\}} - \mathbf{1}_{\{x(b,\omega)=x\}}\right) u(x) \geq 0 \tag{4}$$

for any triple (P_n, a_n, b_n) where $P_n \in \{P_f, P_g\}$, $a_n \in A$, and $b_n \in A$. Thus, for any u that satisfies NIAS,

$$-1 * \sum_{x\in X}\left(\sum_{n=1}^{N} w_n * \left(\sum_{\omega\in\Omega} P_n(a_n,\omega)(\mathbf{1}_{\{x(a_n,\omega)=x\}} - \mathbf{1}_{\{x(b_n,\omega)=x\}}\right)\right) u(x) \geq 0$$

for any collection of triples (P_n, a_n, b_n) where $P_1, ..., P_N \in \{P_f, P_g\}$, $a_1, ..., a_N \in A$, and $b_1, ..., b_N \in A$ with $(P_n, a_n, b_n) \neq (P_f, a, b)$ and non-positive weights $w_1, ..., w_N$. Assuming that equation (3) holds, this implies that for any u that satisfies NIAS,

$$-1 * \sum_{x\in X}\left(\sum_{\omega\in\Omega} P_f(a,\omega)(\mathbf{1}_{\{x(a,\omega)=x\}} - \mathbf{1}_{\{x(b,\omega)=x\}}\right) u(x) \geq 0$$

Because of equation (4), this must equal 0, so equation (2) must hold.

Second, if equation (2) holds, then for all u that satisfy NIAS, then it cannot be that

$$-1 * \sum_{x\in X}\left(\sum_{\omega\in\Omega} P_f(a,\omega)(\mathbf{1}_{\{x(a,\omega)=x\}} - \mathbf{1}_{\{x(b,\omega)=x\}}\right) u(x) < 0 \tag{5}$$

By Farkas lemma, equation (4) and equation (5) means there must exist non-positive weights on that collection of NIAS inequalities that give equation (3), completing the proof. $\square$

Theorem 1: $P_g \succsim_V P_h$ if and only if there exists a weighting function t_{gh} that satisfies DISI for P_g over P_h.

Proof of Theorem 1:

Proof. The NIAS inequality for $P_f \in \{P_g, P_h\}$ and actions $a, b \in A$ can be expressed as a 1 by M row vector $\vec{d}_f(a, b)$ where element m of this vector gives the difference in the probability of receiving outcome x_m from taking action a and from taking action b with the same probability. This is given by

$$\sum_{\omega \in \Omega} P_h(a, \omega)(\mathbf{1}_{\{x(a,\omega)=x_m\}} - \mathbf{1}_{\{x(b,\omega)=x_m\}})$$

where $\mathbf{1}_{\{x(a,\omega)=x_m\}}$ is an indicator function that takes a value of 1 when the outcome from taking action a in state ω yields outcome x_m.

Stacking the row vectors for all NIAS inequalities for $P_f \in \{P_g, P_h\}$ produces a J^2 by M matrix D_h where

$$D_h = \begin{bmatrix} \vec{d}_h(a_1, a_1) \\ \vec{d}_h(a_1, a_2) \\ ... \\ \vec{d}_h(a_J, a_{J-1}) \\ \vec{d}_h(a_J, a_J) \end{bmatrix}$$

and stacking the matrix of NIAS inequalities for both distributions of actions and states produces a $2 * J^2$ by M matrix D where

$$D = \begin{bmatrix} D_f \\ D_g \end{bmatrix}$$

Based on this matrix D, NIAS can be restated as the M by 1 column vector $u \in \mathbb{R}^M$ satisfying $Du \geq 0$ with $Du(m) > 0$ for some $m \in \{1, ..., M\}$.

In addition, the requirement for P_g to be revealed to have a higher value of information than P_h can be expressed as a 1 by M row vector $\vec{d}$ where element m gives the expected gain in outcome x_m from choosing with P_g instead of P_h, which is given by

$$\sum_{a \in A} \sum_{\omega \in \Omega} (P_f(a, \omega) - P_g(a, \omega)) \mathbf{1}_{\{x(a,\omega)=x_m\}}$$

$P_g \succsim_V P_h$ can be restated as $\vec{d}u \geq 0$ for all $u \in \mathbb{R}^M$ that satisfy NIAS. With this notation, both directions of the theorem follow from Farkas lemma.

1. Exists $t \in \mathbb{R}_+^{2*J^2}$ s.t. $D^T t = (\vec{d})^T \Rightarrow$ For all $u \in \mathbb{R}^M$ satisfying NIAS, $\vec{d}u \geq 0$. Assume not. Take $u \in \mathbb{R}^M$ such that NIAS is satisfied, so that $Du \geq 0$, but $\vec{d}u < 0$. By

Farkas lemma, there cannot exist a $t \in \mathbb{R}_{+}^{2*J^2}$ s.t. $D^T t = (\vec{d})^T$, which is a contradiction.

2. For all $u \in \mathbb{R}^M$ satisfying NIAS, $\vec{d}u \geq 0 \Rightarrow$ Exists $t \in \mathbb{R}_{+}^{2*J^2}$ s.t. $D^T t = (\vec{d})^T$.
   Assume there does not exist $t \in \mathbb{R}_{++}^{2*J^2}$ such that $D^T t = (\vec{d})^T$. By Farkas lemma, there must exist a $u \in \mathbb{R}^M$ satisfying NIAS and with $\vec{d}u < 0$, which is a contradiction.

$\square$