A Testable Theory of Imperfect Perception*

Andrew Caplin† and Daniel Martin‡

First version: June 15, 2011; Latest version: July 28, 2013

Abstract

We provide a new characterization of Bayesian expected utility maximization. The novel feature of the model is our separation of actions from prizes. The signature of the resulting theory is the impossibility of raising utility by switching wholesale from one action to another. We provide applications to robustness, to the recovery of utility from choice data, and to model classification.

Key Words: Signal Processing, Bounded Rationality, Incomplete Information, Bayes' Rule, Rational Inattention, Consideration Set

1 Introduction

There is a standard approach to modeling signal processing and choice. Decision makers start out with prior beliefs concerning an underlying state of the world that determines the payoffs to all actions. They receive additional signals concerning this state and update their priors in a Bayesian manner. Their final choice of action maximizes expected utility given these posterior beliefs.

From an applied point of view, the devil is in the details. In a typical application it is impossible to know what form private signals take, let alone how well they are understood. By way of

*We thank Dirk Bergemann, Colin Camerer, David Cesarini, Olivier Compte, Mark Dean, Sen Geng, Paul Glimcher, Phillipe Jehiel, Paola Manzini, Marco Mariotti, Stephen Morris, Antonio Rangel, Natalia Shestakova, and Jonathan Weinstein for valuable comments.

†Center for Experimental Social Science and Department of Economics, New York University and National Bureau of Economic Research.

‡Center for Experimental Social Science and Department of Economics, New York University.

example, consider jurors in a trial. Even if one tightly monitors the trial, one cannot know how the proceedings were translated into informative signals concerning the guilt or innocence of the defendant. Nor can one know how such signals were processed. Is it accurate to model jurors as having perceived all information perfectly? Should we instead model them as perceiving noisy signals? If so, what form should these signals take and from what distribution should they be drawn? When there is no clear answer to these questions, it seems sensible to ask what we can say when we make no assumptions about the exact form of information processing.

In this paper we identify the behavioral limits of the standard model of Bayesian expected utility maximization. These "No Improving Action Switches" (NIAS) inequalities assert that the standard model can be applied if and only if it is impossible to improve utility by making wholesale switches from one action to another. These inequalities apply regardless of the exact form of private information. Moreover, if they apply, there is some specification of private signals and utilities that rationalizes the data.

We present three applications of the NIAS inequalities. In the first application, we make predictions for behavior that are "robust" to the exact form of signal processing. This is analogous to the robust predictions for games of incomplete information found in Bergemann and Morris [2013] (see section 6.1). For example, in a jury voting example, we show that when the defendant is more likely to be innocent, the standard model robustly predicts low levels of Type I error (voting to convict a innocent person), but puts no restriction on the degree of Type II error (voting to acquit a guilty person).

In our second application we illustrate the bounds that the NIAS inequalities place on utility. Our utility recovery question is similar in spirit to that of Rubinstein and Salant [2011], who show how to determine ordinal preference rankings from boundly rational choice procedures. The key difference in our approach is that the NIAS inequalities establish bounds on the relative strength of preference for one prize over another even when ordinal rankings cannot be identified.

In our final application, we use the NIAS inequalities to identify whether or not prominent forms of boundedly rational behavior can be rationalized with a standard Bayesian model. This is a nontrivial question, since there are many different models of choice that produce the same behavior (see Richter [2013]).[1] We show that models based on "consideration sets" (such as those of Manzini

[1] To illustrate that this is not always obvious, note that a simple satisficing model produces the same behavior as

and Mariotti [2007], Masatlioglu, Nakajima, and Ozbay [2009], and Manzini and Mariotti [2013]) produce behaviors that violate the NIAS inequalities, hence cannot be so rationalized. The same applies to behaviors generated by procedural models of list order search proposed by Rubinstein and Salant [2006] and Salant and Rubinstein [2008]. We show also that the standard logit model of discrete choice, while sometimes motivated by cognitive limits (McKelvey and Palfrey [1995]), is inconsistent with the standard Bayesian model. However, a variant of the logit model based on the rational inattention approach is consistent (see Matejka and McKay [2013]).

The centrality of Bayesian expected utility maximization makes the NIAS inequalities of wide applicability. Caplin and Martin [2013] extend them to categorize observed framing effects as having resulted from changes in attention or changes in utility. In a laboratory experiment, they find that list order effects can be modeled as purely attentional, while explaining default effects requires distortions in utility. Caplin and Dean [2013a] consider a more general act environment, and find that in addition to the NIAS inequalities, a "No Improving Attention Cycles" (NIAC) condition characterizes a general form of rational inattention theory. In a laboratory experiment involving a perception task, they find that subjects conform to both conditions in many, but not all, circumstances.

Our key modeling innovation is that we do not allow decision makers (DMs) to directly pick prizes. Rather, we explicitly separate an "action set" A from the prize set X. We treat DMs as directly choosing actions, which could deliver different prizes. For example we allow the juror only to pick guilty or not guilty and the potential purchaser of a good to choose only to buy or to not buy.

Our motive in separating actions and prizes is that we treat the DM's information processing as unobservable to the model-builder. The model-builder can observe prior beliefs, the underlying state of the world (the actual mapping between actions to prize), and how choice among actions depends on this state. This "data theoretic" approach to Bayesian expected utility maximization deliberately mimics the revealed preference approach. In introducing this approach, Samuelson [1939] noted that utility is unobservable. We extend this approach to accommodate the fact that both utility *and* information processing are unobservable.

The separation between actions and prizes on which our approach rests is not standard in the

a standard model of maximization in which all "unsatisfactory" items are indifferent (see Tyson [2008]).

decision theoretic tradition. For Savage, an "act" is literally defined by its consequences in each state of the world, and therefore cannot be separated from these consequences. In expected utility theory and in the Anscombe-Aumann model, the decision maker chooses directly among prize lotteries. In contrast, the separation of actions from prizes is standard both in game theory and in psychophysics, which produces interesting analogies between our approach and work in these fields.

In game theory, the separation of actions from prizes is required to allow for payoff uncertainty associated with the actions of others. Hence it is no surprise that the two papers that are closest to ours come from the strategic literature. Bergemann and Morris [2013] determine all outcomes that can be supported as equilibria of incomplete information games when individuals may have access to private information. We show in section 6.1 that in some cases the NIAS inequalities identify all such equilibrium when the utility function itself may be unknown. Kamenica and Gentzkow [2011] characterize conditions under which a sender can benefit from sending a signal to a receiver who takes an action that impacts both parties. Again, the NIAS inequalities identify the set of receiver actions that are possible for some signal process and some utility function.

In psychophysics, actions are often separated from prizes. Experiments related to the drift-diffusion model provide prominent examples (see Fehr and Rangel [2011]). In this approach, the separation is required to capture manipulations in the location of a prize for a decision maker presented with evidence (typically dots moving to the left or right with a certain coherence) concerning the location of a prize (Shadlen and Newsome [2001], Ratcliff and McKoon [2008]).[2] Experimental implementation of our model and of its further developments have an analogous structure (e.g. Martin [2012] and Caplin and Dean [2013a,b]).

In section 2 we introduce our formal model and establish that the NIAS inequalities characterize expected utility maximization. In section 3 we demonstrate use of these bounds in restricting observed behavior in a jury trial example. In section 4 we show how to use choice data to bound utilities. In section 5 we use the NIAS inequalities to classify behaviors according to whether or not they are rationalizable in a Bayesian model. We discuss strategic analogs and additional applications in section 6.

[2] Krajbich, Lu, Camerer, and Rangel [2012] provide evidence of the drift-diffusion model in consumer choice using eye-tracking technology.

2 Bayesian Expected Utility Representations

We model decision makers (DMs) who may not perfectly perceive the payoffs to taking each action. This could be because the payoffs themselves are complex, and thus hard to determine, or because there is information about the payoffs that is complex.

2.1 Decision Problem

A DM has available a known set of actions A of cardinality $|A| = J$ with generic element $a \in A$. These actions may be of many different forms: voting guilty or not guilty in a trial; buying or not buying a good; choosing one of J positions in a list (Rubinstein and Salant [2006], Caplin, Dean, and Martin [2011], Caplin and Martin [2013]); choosing one of J possible prices (Martin [2012]), or making one of J available guesses concerning the number of blue balls in a display (Caplin and Dean [2013a,b]).

Separate from the actions, there is a set of possible prizes X of cardinality $|X| = N$. The DM knows the set of possible prizes, but may be unsure about the exact prize that will result from choosing each action.

We define the underlying state of the world as specifying precisely the connection between actions and prizes. Given the finiteness of X and A, there are finitely many such states, $\Omega = \{1, .., m, ..M\}$, with state m corresponding to function $\omega_m : A \to X$.

The DM's uncertainty about the state of the world is summarized by $\mu \in \Gamma = \Delta(\Omega)$, with μ_m the prior probability of state m. Given that μ is fixed from this point forward, we discard impossible states, so that $\mu_m > 0$ for all $m \in \Omega$. A **decision problem** is defined by such a triple (X, A, μ).

2.2 Data Set

Even if private signals are unobservable, information processing can place restrictions on the observable relationship between stochasticity in choice and the underlying state of the world. Given a decision problem (X, A, μ), a state dependent stochastic function q identifies the probability distribution over action choices as it depends on the state,

$$q : \Omega \longrightarrow \Delta(A).$$

Let q_m^a be the probability of action a when the state is m. We call the quadruple (X, A, μ, q) the **data set**.

The applications found in this paper are theoretical in nature and do not require collecting this data set. The feasibility of collecting this data is established in the experiments of Martin [2012], Caplin and Dean [2013a,b], and Caplin and Martin [2013]. There are other possible ways to collect this data. The prior μ may be estimated, induced, or elicited. With additional assumptions about stationarity over time or homogeneity across individuals, a data set q can be constructed not only from experimental data, but also from more standard observational data.

2.3 Unobservables

Our goal is not to specify the detailed process of signal extraction and processing, but rather to characterize the general properties that are associated with the standard theory of signal processing and decision making. To that end, we cut through the details of the signal extraction and processing technology and work directly with a mapping from the prior to each possible posterior belief $\gamma \in \Gamma$.[3]

In technical terms, the theories of choice that we characterize are based on a standard expected utility function $U : X \to \mathbb{R}_+$ and two less standard elements: a **perception function** π that produces posterior beliefs and a **choice function** C that maps posterior beliefs into action choices.

Given the utility function over prizes, the utility that results from taking action a in state m can be generated by taking the composition of the utility function and the map between actions and prizes in this state, so that $U_m^a = U(\omega_m(a))$.

The perception function maps each state $m \in \Omega$ into $\Delta(\Gamma)$, the probability distributions over Γ with finite support,

$$\pi : \Omega \longrightarrow \Delta(\Gamma).$$

[3] Kamenica and Gentzkow [2011] show that it is without loss of generality to work directly with posterior beliefs rather than with underlying signals.

This map reflects how the DM perceives and internalizes information about the underlying state of the world, and it allows for stochasticity in this process.

Letting $\pi_m(\gamma)$ be the probability of posterior γ in state m, we define $\Gamma(\pi)$ as the set of possible posteriors,

$$\Gamma(\pi) = \cup_{m=1}^{M}\{\gamma \in \Gamma | \pi_m(\gamma) > 0\}.$$

It is this finite set that serves as the domain of the choice function, which allows for mixed strategies,

$$C : \Gamma(\pi) \longrightarrow \Delta(A).$$

We let $C^a(\gamma)$ denote the probability of choosing action a with posterior $\gamma \in \Gamma(\pi)$.

2.4 BEU Representation

For π and C to provide a possible explanation of the data set requires that their composition generates the data set. To make them consistent with the standard model of signal processing, we insist that π satisfies Bayes' rule. To ensure consistency with the standard model of choice, we insist that C maximizes expected utility. A Bayesian expected utility maximizing (BEU) representation is defined by satisfaction of these three conditions.

Definition 1 (π, C, U) *is a **BEU representation** of* (X, A, μ, q) *if it satisfies:*

1. *Data Matching: For all* $m \in \Omega$ *and* $a \in A$,
$$q_m^a = \sum_{\gamma \in \Gamma(\pi)} \pi_m(\gamma) C^a(\gamma).$$
2. *Bayesian Updating: For all* $m \in \Omega$ *and* $\gamma \in \Gamma(\pi)$,
$$\gamma_m = \frac{\mu_m \pi_m(\gamma)}{\sum_{i=1}^{M} \mu_i \pi_i(\gamma)}.$$
3. *Maximization: For all* $\gamma \in \Gamma(\pi)$ *and* $a \in A$ *such that* $C^a(\gamma) > 0$,
$$\sum_{m=1}^{M} \gamma_m U_m^a \geq \sum_{m=1}^{M} \gamma_m U_m^b \text{ all } b \in A,$$
with the inequality strict for some $\gamma \in \Gamma(\pi)$ *and* $a, b \in A$.

We require strictness in the utility comparison of some pair of actions to prevent the conditions from being trivially satisfied by a utility function with all prizes indifferent.

2.5 The NIAS Theorem

In our data set, Bayesian expected utility maximization is characterized by the impossibility of raising utility by switching wholesale from one action to another. This condition is formalized in the "No Improving Action Switches" (NIAS) inequalities.

Definition 2 *Utility function $U : X \to \mathbb{R}$ satisfies the **NIAS inequalities** with respect to (X, A, μ, q) if,*

$$\sum_{m=1}^{M} \mu_m q_m^a U_m^a \geq \sum_{m=1}^{M} \mu_m q_m^a U_m^b,$$

for all $a, b \in A$, with at least one inequality strict.

Theorem 1 *(X, A, μ, q) has a BEU representation if and only if there exists $U : X \to \mathbb{R}$ satisfying the NIAS inequalities.*

A proof can be found in the Appendix. Necessity of the NIAS inequalities follows directly from the definition of a BEU representation. Sufficiency is established by constructing a BEU representation based on any utility function that satisfies the NIAS inequalities. The only intricacy involves cases in which the choice function C must involve mixed strategies.

For a specific (X, A, μ, q), each NIAS inequality imposes a linear constraint on prize utilities. Existence of a utility function that validates all such inequalities corresponds to establishing non-emptiness of the intersection of $(J-1)^2$ linear inequalities. This can be checked using standard linear programming methods.

Afriat [1967] similarly provided a set of data-defined linear inequalities such that a solution to the inequalities exists if and only if a non-satiated utility function exists that rationalizes the data. While not directly comparable because they are based on deterministic choices from budget sets, Afriat's constraints have a conceptual link with our constraints in that they both treat the utility function as unobservable.

3 Robust Prediction

Our first application of the NIAS inequalities is to produce predictions for behavior that are robust to assumptions about information processing. To illustrate, we use an example of a juror in a criminal trial, which is inspired by the lead example in Kamenica and Gentzkow [2011] and Bergemann and Morris [2013]. The juror has two actions: voting to acquit (action A) and voting to convict (action C), and there are two possible states: the defendant is innocent (state I) or the defendant is guilty (state G).

Three parameters define the data set. These are the prior probability of innocence $\mu_I \in (0,1)$ and the probabilities of voting to acquit when the defendant is either innocent or guilty, $\alpha_I \equiv q_I^A$, $\alpha_G \equiv q_G^A$. These parameters define the statistics that may be of outside interest:

1. The overall probability of voting to acquit is $\mu_I \alpha_I + (1-\mu_I)\alpha_G$.
2. The probability of voting to convict an innocent party, known as Type I error, is $1-\alpha_I$.
3. The probability of voting to acquit a guilty party, known as Type II error, is α_G.
4. The overall probability of a mistaken vote is $\mu_I(1-\alpha_I) + (1-\mu_I)\alpha_G$.

In this illustration, we use the NIAS inequalities to indicate when it is possible to robustly predict a low level of Type I error (voting to convict an innocent party) or Type II error (voting to acquit a guilty party).

3.1 NIAS Inequalities and Limits on Behavior

We assume the juror prefers voting correctly: to acquit if the defendant is innocent and to convict if the defendant is guilty. As a starting point, we follow Kamenica and Gentzkow [2011] and Bergemann and Morris [2013] in assuming that the juror only cares about voting correctly and in normalizing the utility of voting correctly to 1 and the utility of voting incorrectly to 0.

It is intuitively clear that the choice data produced by a Bayesian must satisfy certain conditions. For example, if the probability of innocence is high, there must be high likelihood of voting to acquit. The NIAS inequalities provide all such intuitive restrictions in a concise and precise manner. The

NIAS inequality for voting to acquit simplifies to,

$$\alpha_I \geq \left[\frac{1-\mu_I}{\mu_I}\right] \alpha_G.$$

The NIAS inequality for voting to convict simplifies to,

$$\alpha_I \geq (2 - \frac{1}{\mu_I}) + \left[\frac{1-\mu_I}{\mu_I}\right] \alpha_G.$$

As always, at least one inequality must be strict.[4]

With an even prior ($\mu_I = 0.5$), the NIAS inequalities assert that $\alpha_I > \alpha_G$, which means that the juror must be right strictly more than 50% of the time. It is clear that this is necessary for there to exist a BEU representation. That it is also sufficient shows that being very good at correctly identifying an innocent party can be consistent with being very bad at correctly identifying a guilty party without requiring non-Bayesian reasoning or caring about the type of error committed. If *correct* on innocence 99% of the time ($\alpha_I = .99$), such a juror can be *incorrect* on guilt 98% of the time ($\alpha_I = .98$).

Despite the reach of the standard model, precisely 50% of all conceivable data sets are ruled out by this condition, as illustrated by the shaded region beneath the main diagonal in figure 1. As the prior becomes more uneven, the implied restrictions on choice data become stronger. To illustrate, the larger shaded region in figure 1 indicates the choice probabilities that cannot be rationalized when the prior probability of innocence is $\mu_I = \frac{1}{3}$, in which case the relevant inequality identifying existence of a BEU is,

$$\alpha_I \geq 2\alpha_G.$$

Note that 75% of all conceivable data sets are ruled out by this condition.

3.2 Type I and Type II Error

Note that when $\mu_I = \frac{1}{3}$ there is no outright restriction on Type I error (voting to convict an innocent person), while the rate of Type II error (voting to acquit a guilty person) cannot be above 50%. As the probability of innocence falls, the upper bound on Type II error falls.

The bounds on errors are interdependent. If $\alpha_I = 0$ so that the juror always makes a Type I error, then must be no Type II error. If $\alpha_I = 1$ so that the juror never makes Type I error, then

[4] When $\mu_I < .5$ the second inequality is slack, so holds strictly, while when $\mu_i > .5$ the first inequality is slack, so hold strictly. When $\mu_i = .5$ both hold strictly.

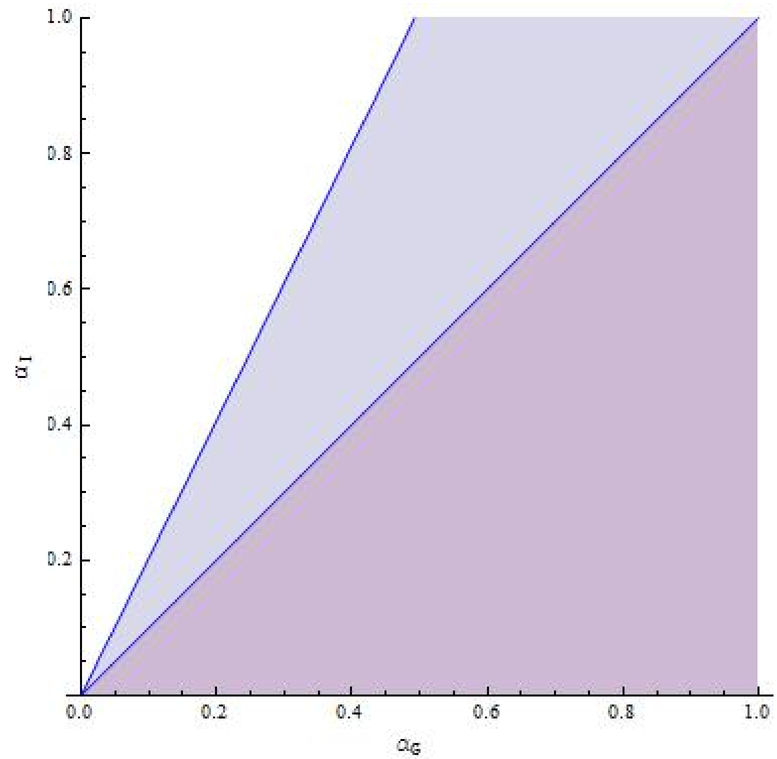

Figure 1: Robust predictions for $\mu_I = .5$ and $\mu_I = \frac{1}{3}$.

there can be up to a 50% rate of Type II error. Between these extremes, the relationship between the Type I error and the maximum Type II error is linear.

Finally, note that when the defendant is more likely to be innocent, say $\mu_I = \frac{2}{3}$, the relationship between Type I and Type II errors reverses. For example, there is no outright restriction on Type II errors, while the Type I error rate cannot be above 50%.

3.3 Caring about Errors

We now allow for the juror to dislike Type I errors (voting to convict a innocent person) differently from Type II errors. We partially continue the normalization from before, so that the utility of voting correctly is 1 and the utility of making a Type II error is 0. However, now the utility of making a Type I error is $\upsilon < 1$. For values of $\upsilon \in (0, 1)$, the juror dislikes Type II error (acquitting a guilty party) more than Type I error (convicting an innocent party), while the converse holds for $\upsilon < 0$.

The NIAS inequality for voting to acquit now simplifies to,

$$\alpha_I \geq \left[\frac{1-\mu_I}{(1-\upsilon)\,\mu_I}\right]\alpha_G;$$

while the NIAS inequality for voting to convict now simplifies to,

$$\alpha_I \geq \left(\frac{2-\upsilon}{1-\upsilon} - \frac{1}{(1-\upsilon)\,\mu_I}\right) + \left[\frac{1-\mu_I}{(1-\upsilon)\,\mu_I}\right]\alpha_G.$$

How behaviorally restrictive the NIAS inequalities are now depends both on the prior and on the value of υ. With $\mu_I = 0.5$ and $\upsilon = -1$ the inequality is,

$$\alpha_I \geq \frac{1}{2} + \frac{1}{2}\alpha_G.$$

This implies that the juror must correctly vote to acquit the innocent at least 50% of the time, with this bound becoming ever more restrictive the more they incorrectly acquit the guilty. Keeping this same prior and increasing the asymmetry in the utility function to the point where $k = -9$, the constraint becomes,

$$\alpha_I \geq \frac{9}{10} + \frac{1}{10}\alpha_G.$$

so that the juror must correctly acquit the innocent at least 90% of the time. Fully 95% of all conceivable data sets are ruled out by this condition, as illustrated by the larger shaded region in figure 2.

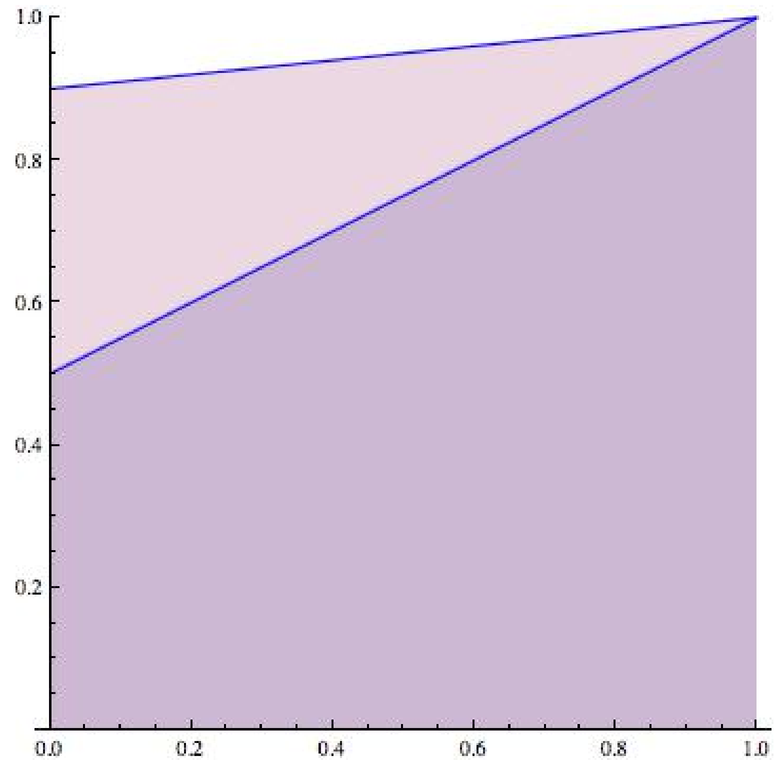

Figure 2: Robust predictions for $\upsilon = -1$ and $\upsilon = -9$ when $\mu_I = .5$.

With $\mu_I = 0.5$ and $\upsilon = 0$, there were no restrictions on Type I error, but with $\upsilon = -1$, we see a substantial restriction on the rate of Type I error. As would be expected, with $\upsilon = -9$ we see even tighter restrictions on the probability of a Type I error.

4 Bounds on Expected Utility

The NIAS inequalities allow us to use choice data to place bounds on unknown utilities even when the exact form of information processing is unspecified. This is similar in spirit to the recovery of preferences in the presence of bounded rationality by Rubinstein and Salant [2011]. They show how to reverse engineer ordinal preference rankings from several deterministic choice procedures. An important difference is that the NIAS inequalities establish bounds on the relative strength of preference for one prize over another even when ordinal rankings cannot be determined.

To illustrate, consider again the example of the previous section, but where the relative importance of Type I and Type II errors in the juror's utility function is unknown. Again this is determined by a single parameter,

$$\upsilon < 1,$$

the utility of Type I error.

In technical terms, we use the NIAS inequalities to constrain υ. The inequality that makes voting to acquit at least as beneficial to the juror as voting to convict simplifies to,

$$\upsilon \leq 1 - \left[\frac{1-\mu_I}{\mu_I}\right]\left(\frac{\alpha_G}{\alpha_I}\right).$$

The NIAS inequality that makes voting to convict at least as beneficial as voting to acquit simplifies to,

$$\upsilon \geq 1 - \left[\frac{1-\mu_I}{\mu_I}\right]\left(\frac{1-\alpha_G}{1-\alpha_I}\right).$$

Overall the requirement is,

$$1 - \left[\frac{1-\mu_I}{\mu_I}\right]\left(\frac{\alpha_G}{\alpha_I}\right) \geq \upsilon \geq 1 - \left[\frac{1-\mu_I}{\mu_I}\right]\left(\frac{1-\alpha_G}{1-\alpha_I}\right),$$

with at least one inequality strict.

To see how the NIAS inequalities constrain relative error costs, consider first a case with $\mu_I = 0.5$, $\alpha_I = \frac{2}{3}$, and $\alpha_G = \frac{1}{3}$. In this case the inequalities assert,

$$\frac{1}{2} \geq \upsilon \geq -1.$$

These constraints do not pin down whether Type I or Type II error is seen as worse because the utility of Type II error is zero. Rather the inequalities constrain the ratio of the losses associated Type I errors relative to Type II errors.

In fact, with $\mu_I = 0.5$, the bounds are on the opposite sides of zero for any choice data,

$$1 - \frac{\alpha_G}{\alpha_I} \geq \upsilon \geq 1 - \frac{1-\alpha_G}{1-\alpha_I}.$$

Hence one cannot know ordinal rankings of Type I and Type II error in this case. However, in the limit as α_I approaches α_G, the utility function is almost exactly pinned down. For $\mu_I = 0.5$, in the limit as α_I approaches α_G, the juror must be close to indifferent between Type I errors and Type II errors. The bounds tighten when the conditional choice probabilities get closer because there is less room for imperfect perception to explain variation in choice.

When the prior is uneven, it is possible to pin down which type of error is worse and also provide bounds on the extent of the difference. For example, with $\mu_I = \frac{1}{4}$, the inequality becomes,

$$1 - 3\left(\frac{\alpha_G}{\alpha_I}\right) \geq \upsilon \geq 1 - 3\left(\frac{1-\alpha_G}{1-\alpha_I}\right).$$

For example, when $\alpha_I = \frac{2}{3}$ and $\alpha_G = \frac{1}{3}$, the requirement is,

$$-\frac{1}{2} \geq \upsilon \geq -5.$$

In this case it is known for sure that Type I errors are regarded as worse than Type II errors, although the precise extent of this preference in not known. Once again, in the limit as α_I falls toward α_G, the utility function is almost pinned down. In this case, the utility bounds tighten around $\upsilon = -2$.

5 Model Classification

The NIAS inequalities can be used to classify models of choice as either consistent or inconsistent the Bayesian expected utility maximization directly from the choice data they produce. To match a typical application in the literature on bounded rationality, we now consider a consumer choosing between two goods x_1 and x_2. The consumer strictly prefers product x_1, and the corresponding utility function is normalized to $U(x_1) = 1$ and $U(x_2) = 0$. However the goods look somewhat similar and are put side-by-side on a shelf. Hence it may be hard for the consumer to determine which good is which.

The consumer can choose the good on the left (action L) or choose the good on the right (action R). The preferred good is either on the left (state l) and on the right (state r), with the prior $\mu_l \in (0,1)$ identifying the probability that it is on the left. The two other parameters that define the data set are $\lambda_l \equiv q_l^L$, $\lambda_r \equiv q_r^L$, the probabilities of picking the good on the left in either state.

5.1 Stochastic Consideration of Prizes

The bounded rationality literature offers many approaches to modeling imperfect perception. Manzini and Mariotti [2013] propose a form of stochastic consideration where a prize x_n is considered with probability $\eta_n \in (0,1)$ and the optimal option inside the consideration set is chosen.[5] The default choice for an empty consideration set is left unspecified. In this example, we assume that the default choice gives some third inferior good x_3 with certainty.

The data produced by this theory reflect the fact that x_1 is chosen if considered and x_2 is chosen only if it is the only prize considered,

$$\begin{aligned} \lambda_l &= \eta_1, \\ \lambda_r &= (1-\eta_1)\,\eta_2. \end{aligned}$$

To illustrate failure of the NIAS conditions, note that substituting this data into the NIAS inequality for action L produces,

$$\eta_1 \geq \frac{(1-\mu_l)\,\eta_2}{\mu_l + (1-\mu_l)\,\eta_2},$$

which is clearly violated when η_1 is close to 1 and μ_l is large.

Note that in this simple example it is hardly surprising that the stochastic consideration set model produces clear failures of updating. After all, it is trivial if one looks at *either* object to select the best prize: if the prize is x_1 it should be chosen, otherwise the other prize should be chosen. Refusal to pick outside a consideration set is hard to justify in a Bayesian model unless the option is sure to be dominated.

[5] Alternatively, Manzini and Mariotti [2007] and Masatlioglu, Nakajima, and Ozbay [2011] consider models of deterministic consideration of prizes.

5.2 Stochastic Consideration of Actions

Rubinstein and Salant [2006] and Salant and Rubinstein [2008] describe choice procedures that consist of searching a fixed number of positions in a list and then selecting the best searched option. In our language, such a procedure produces deterministic consideration of actions. To be more precise, let the action of choosing the first position in a list be a_1. Searching the first position in a list is equivalent to determining the prize associated with action a_1. Caplin and Dean [2011] consider analogous models of search allowing for stochasticity in the order of search. This gives rise to stochastic consideration of actions rather than of prizes as in Manzini and Mariotti [2013].

We amend the earlier analysis by interpreting η_1 as the probability that action L is considered. The data produced by this theory reflect the fact that action L is selected if and only if it is considered and gives the best prize of the considered actions,

$$\begin{aligned}\lambda_l &= \eta_1, \\ \lambda_r &= \eta_1 (1 - \eta_2).\end{aligned}$$

Substituting this data into the NIAS inequality for action L produces a constraint on the consideration of the other action,

$$\eta_2 \geq \frac{1 - 2\mu_l}{1 - \mu_l}.$$

When $\mu_l = .25$, this reduces to $\eta_2 \geq \frac{2}{3}$, which is readily violated.

5.3 Logit Demand

The most important model of discrete choice is the logit model. This approach is sometimes motivated as resulting from imperfect cognition (McKelvey and Palfrey [1995]). However we show the standard version of this model it is not consistent with the Bayesian expected utility maximization. To establish this we simply analyze whether or not the associated stochastic choice data satisfies the NIAS inequalities.

Logit demand, which arises when errors follow an extreme value distribution, produces the following data,

$$\begin{aligned}\lambda_l &= \frac{e^{\frac{1}{\kappa}}}{1 + e^{\frac{1}{\kappa}}}, \\ \lambda_r &= \frac{1}{1 + e^{\frac{1}{\kappa}}},\end{aligned}$$

where $\kappa > 0$ is a parameter of the distribution. By way of interpretation, when the good prize is on the left, it is seen as being on the left with a probability that is increasing in how much better it is than the prize on the right. Rewards in this sense shrink stochastic errors.

To illustrate failure of the NIAS conditions, note that substituting this data into the NIAS inequality for selecting the good on the left produces,

$$e^{\frac{1}{\kappa}} \geq \frac{(1-\mu_l)}{\mu_l},$$

which is violated whenever $\kappa > \frac{1}{\ln\left(\frac{(1-\mu_l)}{\mu_l}\right)}$. To understand why logit demand cannot be rationalized in a Bayesian manner, note that prior beliefs play no role in determining stochastic choice. With a sufficiently uneven prior, logit demand is inconsistent with the NIAS inequalities.

5.3.1 Rational Inattention

An increasingly popular approach to modeling imperfect perception involves rational inattention theory (Sims [2003], Matějka and McKay [2011], Martin [2012], Caplin and Dean [2013b]). In this approach, a perception function is chosen by the decision maker based on a cost that is proportional to the change in Shannon entropy. Matějka and McKay [2011] show that the data this theory produces are of a generalized logit form. In this example, rational inattention produces the following form of logit demand:

$$\begin{aligned}
\lambda_l &= \frac{e^{\frac{1}{\kappa}}}{e^{\frac{1}{\kappa}}-1} - \frac{e^{\frac{1}{\kappa}}}{\mu_l\left(e^{\frac{2}{\kappa}}-1\right)}; \\
\lambda_r &= \frac{(1-\mu_l)\, e^{\frac{1}{\kappa}}\left(\mu_l - \frac{e^{\frac{1}{\kappa}}-1}{e^{\frac{2}{\kappa}}-1}\right)}{\left(e^{\frac{2}{\kappa}}-1\right)}.
\end{aligned}$$

Substituting this data into the NIAS inequality for selecting the good on the left produces,

$$\frac{e^{\frac{2}{\kappa}}-e^{\frac{1}{\kappa}}}{e^{\frac{2}{\kappa}}-1} \geq \frac{1}{2}.$$

which is always satisfied because as κ approaches 0, the limit of the left-hand side is 1, and as κ approaches ∞, the limit of the left-hand side is $\frac{1}{2}$. Thus, rational inattention theory satisfies the standard assumptions.[6]

[6] Because of the symmetry in payoffs and attentional costs, we need consider only one of the constraints.

Figure 4 shows the choice probabilities produced by rational inattention as the linear cost of attention parameter κ varies when the prior is .7. Notice how, unlike the previous theories, they remain within the bounds provided by the NIAS inequalities.

6 Strategic Analogs and Additional Applications

Because action and prize separation is common in game theory, there are several strategic analogs to our model, two of which we discuss in this section.

6.1 Bayes Correlated Equilibrium

Bergemann and Morris [2013] study games of incomplete information in which players may or may not have access to additional private information about the state. For the case of one player, they provide a definition of Bayes correlated equilibrium (BCE) for an arbitrary game G and experiment S, where a game G is a triple (A, U, μ) and an experiment S is a set of signals T and an information structure π. They define BCE with an "obedience" condition on the player's decision rule $\sigma : T \times \Theta \to \Delta(A)$, where Θ is the set of payoff relevant states.

The authors show that σ is a BCE of (G, S) if and only if, for some "expansion" S^* of S, σ is a Bayes Nash equilibrium (BNE) of (G, S^*), where expansion places an ordering on the informativeness of experiments. In other words, they show that if the player is playing a BCE for the game G with some information given by S, it is as if they are playing a BNE for game G with additional information beyond S given by S^*.

Because we treat perception as entirely unobservable, we do not put any restrictions on the player's information. This is analogous to having S be completely uninformative, so that it contains only one signal. Such an S is called the "null" experiment and is denoted by $\underline{S}$. For the null experiment, the decision rule reduces to a function $\sigma : \Theta \to \Delta(A)$, which is the same observable content as our model, so the NIAS conditions can be applied to this function.

Not surprisingly, the conditions for the existence of a BNE in a one player game when information is entirely unobservable are identical to the conditions for a BEU representation. It is immediate from their obedience condition that σ is a BCE of $(G, \underline{S})$ for some non-trivial U if and

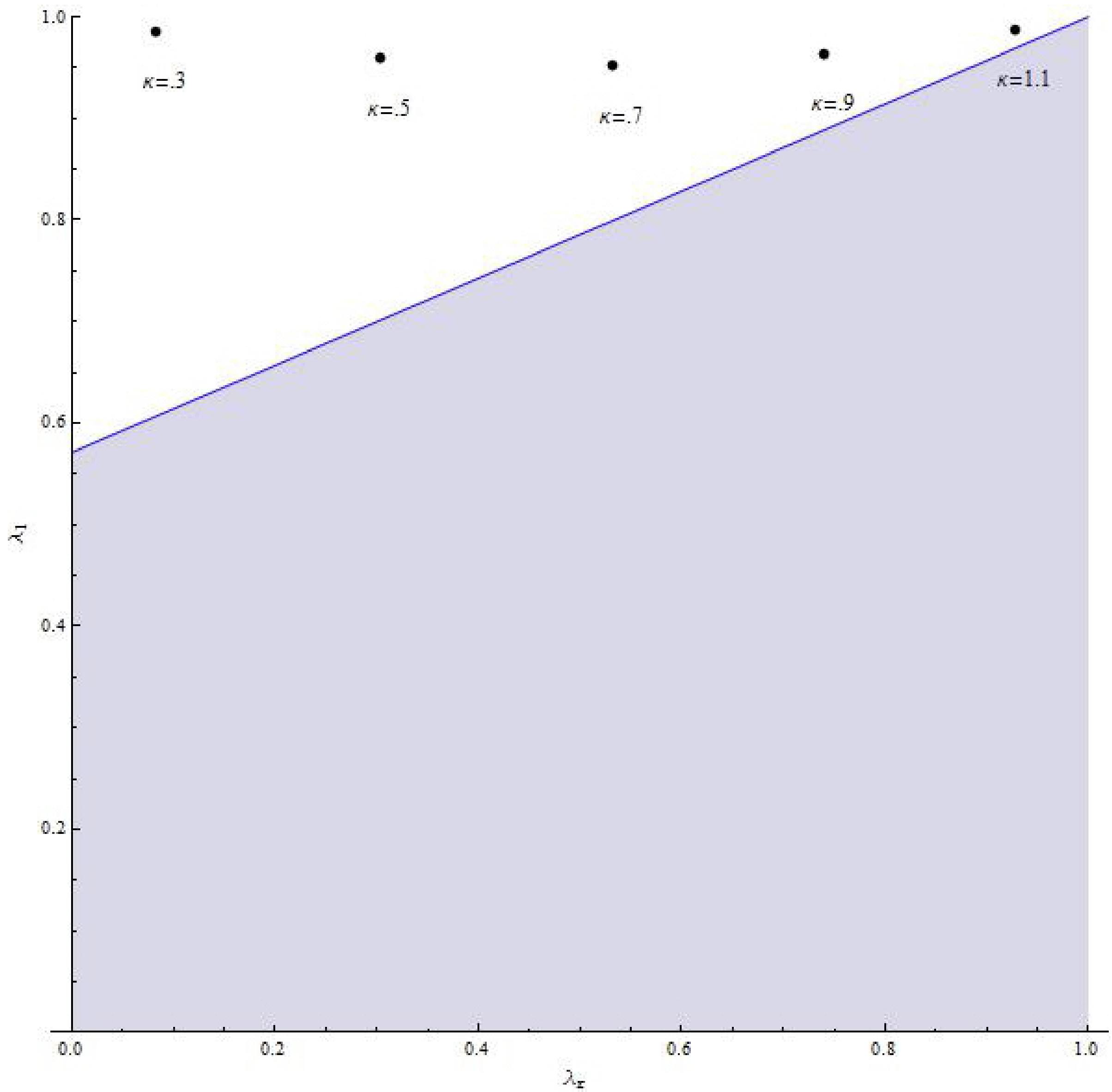

Figure 3: Choice probabilities for rational inattention theory when $\mu_l = .7$.

only if U satisfies the NIAS inequalities.

6.2 Strategic NIAS Inequalities

The NIAS inequalities can also be applied to decision rules in the multiplayer strategic setting of Bergemann and Morris [2013]. Let I be the set of players, finite and non-empty A^i the action set of player i, $\mathcal{A}$ the set of complete action profiles $\mathcal{A} \equiv \Pi_{i\in I} A^i$, and A^{-i} the act sets of other players $A^{-i} \equiv \Pi_{j\in I\setminus i} A^j$. A state dependent stochastic function is a mapping $Q : \Omega \to \Delta(\mathcal{A})$. It is now straightforward to define strategic NIAS inequalities based on utility functions $U^i : X^i \to \mathbb{R}$, with X^i the prize set of player i, where we let $U^i_m(a_i, a_{-i})$ be the state dependent payoff to player i as the result of actions $a_i \in A^i$ and $a_{-i} \in A^{-i}$.

Definition 3 *Utility functions $\{U^i\}_{i\in\{1,\ldots,I\}}$ satisfy the* ***strategic NIAS inequalities*** *with respect to $(X, \mathcal{A}, \mu, Q)$ if for every i,*

$$\sum_{m=1}^{M} \mu_m \sum_{a_{-i}\in A^{-i}} Q_m(a_i, a_{-i}) U^i_m(a_i, a_{-i}) \geq \sum_{m=1}^{M} \mu_m \sum_{a_{-i}\in A^{-i}} Q_m(a_i, a_{-i}) U^i_m(a'_i, a_{-i}),$$

for all $a_i, a'_i \in A^i$, with at least one inequality strict.

Similar to the one player setting, σ is a BCE of $(G, \underline{S})$ for some non-trivial $\{U_i\}_{i\in\{1,\ldots,I\}}$ if and only if $\{U_i\}_{i\in\{1,\ldots,I\}}$ satisfy the strategic NIAS inequalities. Because of this equivalence, our first application is analogous in this setting to the robust prediction found in Bergemann and Morris [2013] for the case of the null experiment.

The chief difference between the approach that we take and that of Bergemann and Morris is around the observably of utility. They treat utility functions as known and analyze possible equilibrium patterns of behavior. We treat the data as given and infer utilities when these data satisfy conditions consistent with equilibrium play. This enables us to recover bounds on utility functions in the strategic setting just as it does in the decision theoretic setting. The distinction is that these are joint restrictions on players' utility functions rather than restrictions that apply to each individual separately. Hence exploration of this revealed preference approach to strategic analysis may be of some interest.

Note also that there is a strategic analogy to our third application involving model classification. The strategic NIAS inequalities can be used to determine which theories of boundedly rational play in games are essentially inconsistent with standard equilibrium analysis.

6.3 Bayesian Persuasion

Kamenica and Gentzkow [2011] determine necessary and sufficient conditions characterizing when a sender can benefit from sending a signal to a receiver who takes a non-contractible action that impacts the utility of both parties. For the finite action case, the NIAS inequalities give the set of receiver actions that are possible under some signal choice and some receiver utility function. This statement reflects two differences between our setting and theirs. First, their action space is infinite rather than finite. Second, they treat the utility function as known.

6.4 Additional Applications

We are currently working to extend, apply, and test models that incorporate the NIAS inequalities. In the current model, the extent of information processing is taken as given. Hence no consideration is given to the cost of information processing. Caplin and Dean [2013a] consider a general model optimal choice of attention when there are such costs, while Caplin and Dean [2013b] consider the special case of Shannon costs. Caplin and Martin [2013a,b] apply the NIAS constraints to classify observed default and framing effects according to whether or not they are consistent with Bayesian expected utility maximizing behavior. Martin [2012] considers strategic applications. In all cases, the NIAS constraints play an essential role.

7 Bibliography

References

[1] Afriat, Sydney. 1967. "The Construction of a Utility Function from Demand Data." *International Economic Review*, 8: 67-77.

[2] Bergemann, Dirk, and Stephen Morris. February 2013. "Bayes Correlated Equilibrium and the Comparison of Information Structures." Mimeo.

[3] Caplin, Andrew, and Mark Dean. 2011. "Search, Choice, and Revealed Preference." *Theoretical Economics,* Volume 6, Issue 1, pages 19–48.

[4] Caplin, Andrew, and Mark Dean. 2013a. "Rational Inattention and State Dependent Stochastic Choice." Mimeo.

[5] Caplin, Andrew, and Mark Dean. 2013b. "Rational Inattention, Entropy, and Choice: The Posterior-Based Approach." Mimeo.

[6] Caplin, Andrew, Mark Dean, and Daniel Martin. 2011. "Search and Satisficing." *American Economic Review*, 101: 2899-2922.

[7] Caplin, Andrew, and Daniel Martin. 2013a. "Defaults and Attention: The Drop Out Effect". Mimeo.

[8] Caplin, Andrew, and Daniel Martin. 2013b. "Attentional Framing Effects." Mimeo.

[9] Fehr, Ernst, and Antonio Rangel. 2011. "Neuroeconomic foundations of economic choice - Recent advances." *Journal of Economic Perspectives*, 25(4): a3-30.

[10] Kamenica, Emir, and Matthew Gentzkow. 2011. "Bayesian Persuasion." *American Economic Review*, 101(6): 2590-2615.

[11] Krajbich, Ian, Dingchao Lu, Colin Camerer, and Antonio Rangel. 2012 "The attentional drift-diffusion model extends to simple purchasing decisions." *Frontiers in Cognitive Science*, 3: 193.

[12] Manzini, Paola, and Marco Mariotti. 2007. "Sequentially Rationalizable Choice." *American Economic Review*, 97(5): 1824-1840.

[13] Manzini, Paola, and Marco Mariotti. March 2013. "Stochastic Choice and Consideration Sets." Mimeo.

[14] Martin, Daniel. 2012. "Strategic Pricing and Rational Inattention to Quality." Mimeo.

[15] Masatlioglu, Yusufcan, Daisuke Nakajima and Erkut Ozbay. 2011. "Revealed Attention." *American Economic Review*.

[16] Matějka, Filip, and Alisdair McKay. November 2011. "Rational Inattention to Discrete Choices: A New Foundation for the Multinomial Logit Model." Mimeo.

[17] McKelvey, Richard, and Thomas Palfrey. 1995. "Quantal Response Equilibria for Normal Form Games." *Games and Economic Behavior*.

[18] Ratcliff, Roger, and Gail McKoon. 2008. "The diffusion decision model: theory and data for two-choice decision tasks." *Neural Computation*, 20: 873-922.

[19] Richter, Michael. 2013. "Choice Theory with Equivalence." Mimeo.

[20] Rubinstein, Ariel, and Yuval Salant. 2006. "A Model of Choice from Lists." *Theoretical Economics*, 1(1): 3-17.

[21] Rubinstein, Ariel, and Yuval Salant. 2011. "Eliciting Welfare Preferences from Behavioural Data Sets." *Review of Economic Studies*: 1-13.

[22] Salant, Yuval, and Ariel Rubinstein. 2008. "("A", "f"): Choice with Frames," *Review of Economic Studies*, 75(4): 1287-1296.

[23] Shadlen, Michael N., and William T. Newsome. 2001. "Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey." *Journal of Neurophysiology*, 86: 1916-1936.

[24] Sims, Christopher A. 2003. "Implications of Rational Inattention." *Journal of Monetary Economics*, 50 (3), 665–690.

[25] Tyson, Christopher J. 2008. "Cognitive constraints, contraction consistency, and the satisficing criterion." *Journal of Economic Theory*, 138(1): 51-70.

8 Appendix: Proof of Theorem 1

Proof. Necessity: Suppose that (π, C, U) define a BEU of (X, A, μ, q). We show directly that $U : X \to \mathbb{R}$ must satisfy the NIAS inequalities. Note first that from Maximization, given any $\gamma \in \Gamma(\pi)$ and $a \in A$,

$$C^a(\gamma)\left[\sum_{m=1}^{M} \gamma_m U_m^a\right] \geq C^a(\gamma)\left[\sum_{m=1}^{M} \gamma_m U_m^b\right] \text{ all } b \in A.$$

Adding up across $\gamma \in \Gamma(\pi)$, using the Bayesian Updating property to substitute for γ_m^a, changing order of addition, and cancelling common terms $\left[\sum_{i=1}^{M} \mu_i \pi_i(\gamma)\right]$ in all denominators, we derive,

$$\sum_{m=1}^{M} \mu_m \left[\sum_{\gamma \in \Gamma(\pi)} \pi_m(\gamma) C^a(\gamma)\right] U_m^a \geq \sum_{m=1}^{M} \mu_m \left[\sum_{\gamma \in \Gamma(\pi)} \pi_m(\gamma) C^a(\gamma)\right] U_m^b \text{ all } b \in A,$$

We now use Data Matching to substitute for the inner summations and derive,

$$\sum_{m=1}^{M} \mu_m q_m^a U_m^a \geq \sum_{m=1}^{M} \mu_m q_m^a U_m^b,$$

verifying the all NIAS inequalities hold at least weakly. To confirm that at least one is a strict inequality, pick $a \in A$ such that there exists $\gamma \in \Gamma(\pi)$ with $C^a(\gamma) > 0$, so that,

$$\sum_{m=1}^{M} \gamma_m U_m^a > \sum_{m=1}^{M} \gamma_m U_m^b \text{ some } b \in A.$$

With this we know that,

$$C^a(\gamma)\left[\sum_{m=1}^{M} \gamma_m U_m^a\right] > C^a(\gamma)\left[\sum_{m=1}^{M} \gamma_m U_m^b\right].$$

Repeating other steps from this point forward reveals that the corresponding NIAS inequality holds strictly,

$$\sum_{m=1}^{M} \mu_m q_m^a U_m^a > \sum_{m=1}^{M} \mu_m q_m U_m^b.$$

Sufficiency: Consider a function $\bar{U} : X \to \mathbb{R}$ that satisfies the NIAS inequalities with respect to (X, A, μ, q). We now identify perception and choice strategies such that $(\pi, C, \bar{U})$ identify a BEU of (X, A, μ, q). First, define $\bar{\gamma}_m^a$ for each action $a_j \in A$ and each $m \in \Omega$ by,

$$\bar{\gamma}_m^a \equiv \frac{\mu_m q_m^a}{\sum_{i=1}^{M} \mu_i q^a}$$

We partition the set of possible acts a into some $P < J$ sets A^p with identical posteriors γ^p within each such set and distinct posteriors in distinct such sets, so that $a \in A^p$ if and only if $\bar{\gamma}^a = \gamma^p$. The perception function is now defined by,

$$\bar{\pi}_m(\gamma^p) = \sum_{a \in A^p} q_m^a,$$

while the choice function is defined by,

$$\bar{C}^a(\gamma^p) = \frac{\sum_{i=1}^M \mu_i q^a}{\sum_{b \in A^p} \sum_{i=1}^M \mu_i q_i^b} > 0 \text{ if and only if } a \in A^p.$$

To confirm that this identifies a BEU, note first that given $m \in \Omega$ and $b \in A^p$,

$$\sum_{\gamma^q \in \Gamma(\pi)} \bar{\pi}_m(\gamma^q)\bar{C}^b(\gamma^q) = \pi_m(\gamma^p)\bar{C}^b(\gamma^p) = \sum_{a \in A^p} q_m^a \left[\frac{\sum_{l=1}^M \mu_l q_l^b}{\sum_{a \in A^p} \sum_{l=1}^M \mu_l q_l^j}\right]$$

Now note that, given $m \in \Omega$ and $a \in A^p$,

$$\sum_{l=1}^M \mu_l q_l^a = \mu_m q_m^a \left[\frac{\sum_{l=1}^M \mu_l q_l^b}{\mu_m q_m^b}\right],$$

by the shared membership $a, b \in A^p$. Hence,

$$\sum_{a \in A^p} \sum_{l=1}^M \mu_l q_l^a = \sum_{a \in A^p} \mu_m q_m^a \left[\frac{\sum_{l=1}^M \mu_l q_l^b}{\mu_m q_m^b}\right].$$

Hence,

$$\begin{aligned}
\bar{\pi}_m(\gamma^p)\bar{C}^b(\gamma^p) &= \sum_{a \in A^p} q_m^a \left[\frac{\mu_m q_m^b}{\sum_{a \in A^p} \mu_m q_m^a}\right] \\
&= q_m^b \left[\frac{\sum_{a \in A^p} \mu_m q_m^a}{\sum_{a \in A^p} \mu_m q_m^a}\right] = q_m^b,
\end{aligned}$$

in confirmation of Data Matching.

To confirm Bayesian Updating, note that for all $m \in \Omega$, $\gamma^p \in \Gamma(\pi)$, and $b \in A^p$,

$$\gamma_m^p = \frac{\mu_m q_m^b}{\sum_{l=1}^M \mu_l q_l^b} = \frac{\mu_m \bar{\pi}_m(\gamma^p)\bar{C}^b(\gamma^p)}{\sum_{l=1}^M \mu_l \pi_l(\gamma^p)\bar{C}^b(\gamma^p)} = \frac{\mu_m \bar{\pi}_m(\gamma^p)}{\sum_{l=1}^M \mu_l \bar{\pi}_l(\gamma^p)}.$$

Finally, note that for each $m \in \Omega$, $1 \leq p \leq P$ and $a \in A^p$, we know definitionally that,

$$\mu_m q_m^a = \gamma_m^p \sum_{l=1}^M \mu_l q_l^a.$$

Substitution in the NIAS inequalities and division by the constant $\sum_{l=1}^M \mu_l^a q_l^a > 0$ yields,

$$\sum_{m=1}^M \gamma_m^p \bar{U}_m^a \geq \sum_{m=1}^M \gamma_m^p \bar{U}_m^b,$$

for all $m \in \Omega$, $1 \leq p \leq P$ and $a \in A^p$, with the inequality strict for some $a, b \in A$. This establishes Maximization and completes the proof. ■