NBER WORKING PAPER SERIES

RATIONALIZABLE LEARNING

Andrew Caplin
Daniel J. Martin
Philip Marx

Working Paper 30873
http://www.nber.org/papers/w30873

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2023

We thank participants in the Sloan-NOMIS Summer School on Cognitive Foundation of Economic Behavior for helpful comments. Caplin thanks the Alfred P. Sloan Foundation and the NOMIS Foundation for their support for the broader research program on the cognitive foundations of economic behavior. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Rationalizable Learning
Andrew Caplin, Daniel J. Martin, and Philip Marx
NBER Working Paper No. 30873
January 2023
JEL No. D83,D91

ABSTRACT

The central question we address in this paper is: what can an analyst infer from choice data about what a decision maker has learned? The key constraint we impose, which is shared across models of Bayesian learning, is that any learning must be rationalizable. To implement this constraint, we introduce two conditions, one of which refines the mean preserving spread of Blackwell (1953) to take account for optimality, and the other of which generalizes the NIAC condition (Caplin and Dean 2015) and the NIAS condition (Caplin and Martin 2015) to allow for arbitrary learning. We apply our framework to show how identification of what was learned can be strengthened with additional assumptions on the form of Bayesian learning.

Andrew Caplin
Department of Economics
New York University
19 W. 4th Street, 6th Floor
New York, NY 10012
and NBER
andrew.caplin@nyu.edu

Daniel J. Martin
Northwestern University
Kellogg School of Management
2211 Campus Drive
Evanston, IL 60208
and University of California, Santa Barbara
d-martin@kellogg.northwestern.edu

Philip Marx
Department of Economics
Louisiana State University
501 South Quad Dr.
Baton Rouge, LA 70803
pmarx@lsu.edu

1 Introduction

Bayesian learning models form a bedrock of modern social science and are ubiquitous in economic, psychological, and neuroscientific analysis. In these models, decision makers acquire informative signals about the state, update their beliefs using Bayes' rule, and then choose an action that maximizes expected utility. This broad framework includes fixed information models, where learning is not impacted by the decision context, such as in market models of incomplete information, auctions, and observational learning; capacity constrained learning models, where one of a feasible set of learning methods is chosen to maximize payoffs, such as in fixed capacity rational inattention (Sims 2003) and optimal encoding (Woodford 2014); and costly learning models, where distinct methods of learning have different costs and the chosen method maximizes net payoffs, as in sequential search, bandit problems, and variable capacity rational inattention (Matejka and McKay 2015; Caplin, Dean, and Leahy 2017).

A key component of all Bayesian learning models is the learning itself, which is represented by the decision maker's information structure. However, in practice what a decision maker learns is rarely observed directly. Instead, imagine an analyst who would like to infer what the decision maker learned but only observes the following joint distribution over states $\{\omega_1, \omega_2\}$ and actions in respective choice sets $A^1 = \{a_1, a_2\}$ and $A^2 = \{a_1, a_2, a_3\}$:

$$P^1 = \begin{array}{c} \\ a_1 \\ a_2 \end{array}\!\!\begin{array}{c} \begin{array}{cc} \omega_1 & \omega_2 \end{array} \\ \begin{pmatrix} 0.4 & 0.1 \\ 0.1 & 0.4 \end{pmatrix} \end{array} \qquad \text{and} \qquad P^2 = \begin{array}{c} \\ a_1 \\ a_2 \\ a_3 \end{array}\!\!\begin{array}{c} \begin{array}{cc} \omega_1 & \omega_2 \end{array} \\ \begin{pmatrix} 0.25 & 0 \\ 0.05 & 0.2 \\ 0.2 & 0.3 \end{pmatrix} \end{array}$$

Such state-dependent stochastic choice data has long been analyzed in psychology, neuroscience, and economics, especially when estimating and evaluating models of Bayesian learning.[1] In psychometrics experiments, the state could be luminosity, weight, or the direction of movement; in an economic setting, it could be the fundamentals of a stock, the effectiveness of a vaccine, characteristics of a health plan, and so on.

The central question we address in this paper is: what can the analyst infer from such data about what a decision maker has learned? The key constraint we impose, which is shared across models of Bayesians learning, is that any learning must be consistent with rational choice. In other words, the decision maker's learning must be rationalizable.

We consider rationalizable learning on two levels of generality. First, we consider ratio-

[1]For instance, Caplin and Martin (2015), Caplin and Dean (2015), Chambers, Liu, and Rehbeck (2017), Caplin, Dean, and Leahy (2017), and Denti (2022) use state-dependent stochastic choice to characterize different forms of Bayesian learning, and Lipnowski and Ravid (2022) use knowledge of learning costs to predict state-dependent stochastic choice.

nalizable learning given choice data from a single decision problem. Here, our main result (Theorem 1) is that an information structure is rationalizable if and only if it constitutes a *mean and optimality preserving spread (MOPS)* of the least informative information structure consistent with the data, which itself is readily revealed by the data. As the name suggests, our MOPS operation refines the mean preserving spread of Blackwell (1953) to take account for optimality. That is, it restricts attention to mean preserving spreads of posterior beliefs that also preserve optimal choice. Its value in our context is to provide a constructive characterization of all information structures rationalizable within a given decision problem.

Second, we consider rationalizable learning given choice data from multiple decision problems. Such choice data has been shown to be crucial for testing and distinguishing among Bayesian models of learning (e.g., Caplin and Dean 2015, Denti 2022). In order to accommodate the additional constraints on rationalizable learning that arise *across* decision problems, we complement MOPS with an approach grounded in the value of information.

Here, our main result (Theorem 2) is that a set of information structures is rationalizable under costly learning if and only if (1) each is a MOPS of the least informative information structure consistent with the data from a given decision problem and (2) as a collection they satisfy a *Generalized No Improving Cycles* (G-NIC) condition, which generalizes the NIAC (Caplin and Dean 2015) and NIAS (Caplin and Martin 2015) conditions as a function of arbitrary information structures. The G-NIC condition is compactly summarized using the *indirect value difference function*, which is a matrix-valued function that builds on the classic revealed preference approach of Varian (1982). Analogous to MOPS in the single-problem case, the indirect value difference function encodes the requirements of optimality across decision problems. Thus, our framework specializes the classic three-way equivalence of Blackwell (1953) informativeness between mean preserving spreads, the value of information, and signal garbling (which relates naturally to rationalizability in our setting) to precisely target the optimality embedded in models of Bayesian learning.

We apply our framework to show that identification of what was learned can be strengthened with assumptions on the form of Bayesian learning because doing so adds requirements on optimality across decision problems. Specifically, we identify what could have been learned under the further assumptions that produce capacity constrained learning (Proposition 2) and fixed information (Proposition 3) by strengthening the indirect value difference function and MOPS requirements, respectively.

The paper proceeds as follows. In Section 1.1 we provide a motivating example which we use to set ideas throughout the paper. In Section 1.2 we discuss related literature and additional contributions. In Section 2 we formalize the decision problem and introduce the key objects of analysis. In Section 3 we introduce MOPS and consider rationalizable learning

within a decision problem. In Section 4 we introduce the indirect value difference function and consider rationalizable learning with choice data from multiple decision problems. In Section 5, we recover the costs of what was learned (Subsection 5.1), characterize rationalizable learning under nested models (Subsection 5.2), and recover all consistent utility functions (Subsection 5.3).

1.1 A Motivating Example

To illustrate the challenge of recovering information structures from choice data, consider again the example data sets introduced previously. Two choice sets are faced: $A^1 = \{a_1, a_2\}$ and $A^2 = \{a_1, a_2, a_3\}$. Choice data P^1 from the first choice set is summarized by:

$$P^1 = \begin{array}{c} \begin{matrix} \omega_1 & \omega_2 \end{matrix} \\ \begin{pmatrix} 0.4 & 0.1 \\ 0.1 & 0.4 \end{pmatrix} \end{array} \begin{matrix} a_1 \\ a_2 \end{matrix}$$

Clearly, this exhibits symmetric choice patterns. The choice data P^2 observed from the second choice set is somewhat asymmetric:

$$P^2 = \begin{array}{c} \begin{matrix} \omega_1 & \omega_2 \end{matrix} \\ \begin{pmatrix} 0.25 & 0 \\ 0.05 & 0.2 \\ 0.2 & 0.3 \end{pmatrix} \end{array} \begin{matrix} a_1 \\ a_2 \\ a_3 \end{matrix}$$

In the first state, the first and third actions are selected more often, and in the second state, the second and third actions are selected more often. Further, these actions can yield one of three prizes $\{z_G, z_M, z_B\}$, which are known to correspond to actions and states as follows:

Action	**State** ω_1	**State** ω_2
a_1	z_G	z_B
a_2	z_B	z_G
a_3	z_M	z_M

What does the choice data reveal about learning? First, as detailed in Caplin and Dean (2015), the *revealed information structures* for the data are always consistent with the data under costly learning. Revealed information structures are the distributions of posteriors for each choice set derived as if each action in the choice data had been chosen at a single posterior, as in action recommendation strategies. Each data set P^1 and P^2 corresponds to a revealed information structure $\bar{Q}^1$ and $\bar{Q}^2$, which is summarized by revealed posteriors $\bar{\gamma}$

of state ω_1 and probabilities of posteriors given by $\bar{Q}(\gamma)$:

$$\begin{array}{cc} \bar{\gamma} & \bar{Q}^1(\bar{\gamma}) \\ \begin{pmatrix} 4/5 \\ 1/5 \end{pmatrix} & \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix} \end{array} \qquad \begin{array}{cc} \bar{\gamma} & \bar{Q}^2(\bar{\gamma}) \\ \begin{pmatrix} 1 \\ 1/5 \\ 2/5 \end{pmatrix} & \begin{pmatrix} 1/4 \\ 1/4 \\ 1/2 \end{pmatrix} \end{array}$$

One possibility is that these revealed information structures are in fact the information structures of the decision maker. However, this is not necessary, nor is it consistent with the decision maker having fixed information.

Alternatively, consider the following three candidate information structures:

$$\begin{array}{cc} \gamma & Q(\gamma) \\ \begin{pmatrix} 1 \\ 1/2 \\ 0 \end{pmatrix} & \begin{pmatrix} 3/10 \\ 2/5 \\ 3/10 \end{pmatrix} \end{array} \qquad \begin{array}{cc} \gamma & Q(\gamma) \\ \begin{pmatrix} 1 \\ 1/2 \\ 1/8 \end{pmatrix} & \begin{pmatrix} 3/10 \\ 3/10 \\ 2/5 \end{pmatrix} \end{array} \qquad \begin{array}{cc} \gamma & Q(\gamma) \\ \begin{pmatrix} 1 \\ 3/5 \\ 1/5 \end{pmatrix} & \begin{pmatrix} 1/4 \\ 1/4 \\ 1/2 \end{pmatrix} \end{array}$$

Which of these information structures are consistent with the observed data? Our methods will clarify the following. The first information structure is rationalizable as a MOPS of $\bar{Q}^1$ but inconsistent with costly learning because it is too informative not to have been chosen in decision problem 2. The second information structure is rationalizable for decision problem 1 under costly learning but not under the more stringent model of capacity constrained learning. The third information structure is consistent not just with capacity constraints, but with fixed information. In fact, it is the least informative such structure consistent with the data under fixed information. In the remainder of the paper, we show exactly how to identify information structures, both in this example and in the general case.

1.2 Related Literature and Additional Contributions

In related work, Caplin and Dean (2015) identify a necessary condition for rationalizable learning: being as informative as the revealed information structure. Their condition is not sufficient because such an information structure can be so informative that utility maximization is not satisfied, either within or across decision problems, as demonstrated in our motivating example. We build on their contribution by establishing both necessary and sufficient conditions for information structures to be consistent with choice data. In addition, we characterize learning in two other Bayesian learning models: capacity constrained learning and fixed information.

Thus we also relate to Lu (2016), who identifies the set of possible information structures under fixed information using rich test functions that are generated over all possible payoffs.

Instead, we show what can be recovered about information structures under fixed information using choice data from just a finite set of possible payoffs, a much simpler data requirement. We also generalize this approach to cover capacity constrained and costly learning.

Our paper makes two additional contributions related to the testability of Bayesian learning models. First, we show that a simple condition on the indirect value difference function provides a test of capacity constrained learning, and to the best of our knowledge, this is the first general test of this model class. Second, we show that the indirect value difference function helps operationalize the test of costly learning established in Caplin and Dean (2015) because a simple condition on that function is equivalent to their test, and the output of the function can be efficiently computed using the simple and well-known polynomial-time algorithm of Floyd (1962) and Warshall (1962).

This paper also makes contributions related to the recovery of other key objects in Bayesian learning models: learning costs and utility. Caplin and Dean (2015) provide linear constraints that must be satisfied for learning costs to be consistent with the data, and in Subsection 5.1 we complement their result by showing that the indirect value difference function can also be used to recover all possible information costs that rationalize what was learned (Theorem 3). We further add to this by showing that the function also explicitly encodes a variety of extremal learning costs, which in turn define a single representative cost function of what was learned (Proposition 1).

In addition, we show how to recover consistent utility functions by generating utility cones in the space of prize lotteries (Subsection 5.3). This extends the geometric approach to recovering utility under Bayesian expected utility maximization introduced in Caplin and Martin (2021) to cover models of Bayesian learning that place restrictions of the form of optimal learning, such as costly learning and capacity constrained learning. Because our utility cones provide a ready means to recover all consistent utility functions, they enable welfare analysis across a wide set of Bayesian learning models.

2 Setup

There is a finite set of possible states of the world $\omega \in \Omega$ and a fixed prior $\mu \in \Delta(\Omega)$. There is a finite global set of actions $\mathcal{A}$. There is a finite prize set $Z = \{z_k\}_{k=1}^K$. In any given decision problem, a finite set of actions $A \subset \mathcal{A}$ with $|A| \geq 2$ is available. For each action, the prize that is realized depends on the state of the world according to a prize specification $z(a, \omega)$ that is known by the analyst. Because rewards depend on the state, the decision maker (DM) is motivated to learn about the state before making action choices.

2.1 Data

As in Caplin and Martin (2015) (CM15 hereafter), the data relevant to assessing what the DM learns before choosing in decision problem $A \subset \mathcal{A}$ is state-dependent stochastic choice (SDSC) data. This specifies the joint distribution of actions and states $P(a, \omega)$ for all $a \in A$ and $\omega \in \Omega$, with marginal distributions recovering the fixed prior and unconditional action probabilities:

$$\mu(\omega) = \sum_{a \in A} P(a, \omega),$$

$$P(a) \equiv \sum_{\omega \in \Omega} P(a, \omega);$$

Alternatively, the SDSC data is equivalently represented by a revealed information structure $\bar{Q}$, defined as the distribution over action-conditional posterior beliefs:

$$\bar{\gamma}^a(\omega) \equiv P(a, \omega) \big/ P(a). \tag{1}$$

for all chosen actions, $P(a) > 0$. When ambiguities exist, our convention throughout the paper is to distinguish revealed data objects from their theoretical counterparts with a bar.

2.2 Consistency with Bayesian Learning

Our main goal is to characterize which information structures are consistent with the observed data P under Bayesian learning. As in Kamenica and Gentzkow (2011), we specify an information structure as a Bayes consistent distribution Q of posteriors $\gamma \in \Delta(\Omega)$ with finite support $\Gamma(Q) \equiv \operatorname{supp} Q$, with their set given by:

$$\mathcal{Q} \equiv \{Q \in \Delta(\Delta(\Omega)) \text{ with } |\Gamma(Q)| < \infty \text{ and } \sum_{\gamma \in \Gamma(Q)} \gamma Q(\gamma) = \mu\}.$$

The DM has a mixed strategy over actions that is a function of the posterior, $q(a|\gamma) \in \Delta(A)$. To help determine consistency, we define $P_{(Q,q)}$ as the hypothetical SDSC that (Q, q) would generate,

$$P_{(Q,q)}(a, \omega) \equiv \sum_{\gamma \in \Gamma(Q)} q(a|\gamma) Q(\gamma) \gamma(\omega). \tag{2}$$

Thus, (Q, q) could have generated the data P if:

$$P_{(Q,q)} = P \tag{3}$$

We will say that (Q, q) *rationalizes* the data if it could have generated the data and could have arisen from optimal choice.

We model the DM's optimization problem in two stages, which we solve using backward induction. In the second stage, given an information structure Q and decision problem A, the DM chooses an action strategy to maximize expected utility. To this end, define the posterior expected utility of action a given a utility function $u : Z \to \mathbb{R}$ and a posterior γ as:

$$U(a|\gamma, u) \equiv \sum_{\omega \in \Omega} \gamma(\omega) u(z(a, \omega)) \tag{4}$$

and the gross expected utility of strategy (Q, q) given utility function u as:

$$g(Q, q|u) \equiv \sum_{\gamma \in \Gamma(Q)} \sum_{a \in \mathcal{A}} Q(\gamma) q(a|\gamma) U(a|\gamma, u). \tag{5}$$

As a function of information structure Q and choice set A the DM chooses an action strategy to solve:

$$\underset{q:\Gamma(Q)\to\Delta(A)}{\operatorname{argmax}} \; g(Q, q|u) \tag{6}$$

In what follows, it will also be useful to define the resulting gross value of learning an information structure Q in decision problem A given utility function u as:

$$G(Q|A, u) \equiv \max_{q:\Gamma(Q)\to\Delta(A)} g(Q, q|u) \tag{7}$$

In the first stage, the DM chooses an information structure to maximize this value of learning function minus a learning cost, which we summarize by a function $K : \mathcal{Q} \to \mathbb{R} \cup \{\infty\}$ as in CD15. That is, the DM chooses a learning strategy to solve:

$$\underset{Q \in \mathcal{Q}}{\operatorname{argmax}} \; G(Q|A, u) - K(Q) \tag{8}$$

While we model the information structure as chosen, this general structure nests all of the models considered in this paper. Capacity constrained learning is captured by the cost of an information structure being either 0 or ∞, and fixed information is captured by the cost of only one information structure being finite, so that the choice of information structure is trivial.

3 Rationalizing Within Decision Problem

We begin by considering what could have been learned in a single decision problem summarized by a choice set A and observed choice data P. Specifically, we are interested in characterizing the set of information structures Q for which there exists a mixed action strategy q satisfying expected utility maximization (6) such that (Q, q) generates the observed data (3). In this case we say that the strategy rationalizes the data within decision problem (according to expected utility maximization).

Throughout the following sections, we take as given a prize utility function. In Subsection 5.3 we consider the case where the utility function is unknown and must itself be recovered from the available data. In the case of our running example, the resulting characterization yields simple bounds, which we henceforth take as given:

$$u(z_B) = 0, \quad u(z_M) \in [0.6, 0.8], \quad u(z_G) = 1 \tag{9}$$

3.1 Mean and Optimality Preserving Spreads

The centerpiece of our characterization of what could have been learned under expected utility maximization in a single decision problem is what we term a mean *and optimality* preserving spread. To relate the standard concept of a mean preserving spread to optimality, we use the fact that every action a maps to a single revealed posterior $\bar{\gamma}^a$. Additionally, we define a shorthand for the set of posteriors $\gamma \in \Delta(\Omega)$ at which each action $a \in A$ is optimal:

$$\hat{\Gamma}(a|A,u) \equiv \{\gamma \in \Delta(\Omega) \,|\, U(a|\gamma,u) \geq U(b|\gamma,u) \text{ for all } b \in A\}$$

Then our definition of a mean and optimality preserving spread (MOPS) is as follows.

Definition 1. *Given decision problem A and utility function u, information structure Q is a* ***mean and optimality preserving spread*** *(MOPS) of revealed information structure $\bar{Q}$ if there exists a function $B : |\Gamma(\bar{Q})| \times |\Gamma(Q)| \to [0,1]$ satisfying the standard conditions of a mean preserving spread:*

$$\sum_{\gamma \in \Gamma(Q)} B(\bar{\gamma}, \gamma) = 1 \tag{10}$$

$$\sum_{\bar{\gamma} \in \Gamma(\bar{Q})} \bar{Q}(\bar{\gamma}) B(\bar{\gamma}, \gamma) = Q(\gamma) \tag{11}$$

$$\sum_{\gamma \in \Gamma(Q)} B(\bar{\gamma}, \gamma)\gamma = \bar{\gamma} \tag{12}$$

while additionally preserving optimality:

$$B(\bar{\gamma}^a, \gamma) > 0 \implies \gamma \in \hat{\Gamma}(a|A,u) \tag{13}$$

for all chosen actions $P(a) > 0$.

As is standard, the first condition (10) requires that for every posterior $\bar{\gamma} \in \Gamma(\bar{Q})$, $B(\bar{\gamma}, \cdot)$ is a probability distribution over $\Gamma(Q)$. The second condition (11) requires that for every posterior $\gamma \in \Gamma(Q)$, the probability $Q(\gamma)$ is obtained as the sum of spread mass from the posteriors $\bar{\gamma}$. The third condition (12) requires that the posteriors $\bar{\gamma} \in \Gamma(\bar{Q})$ are an average

of posteriors in $\Gamma(Q)$. Finally, condition (13) requires the spreading of revealed posteriors to preserve optimality of the corresponding actions.

A MOPS provides a simple way of generating information structures from the revealed information structure by spreading mass from revealed posteriors in a way that preserves optimality of their associated actions. The following result establishes its equivalence with the set of information structures that rationalize the observed data, and thus could have been learned.

Theorem 1. *Fix a decision problem with data (A, P) and revealed information structure $\bar{Q}$. Given utility function u, the following are equivalent for an information structure Q:*

1. *There exists an optimal action strategy q such that (Q, q) rationalizes the data P according to EU maximization.*
2. *The information structure Q is a mean and optimality preserving spread of $\bar{Q}$.*

We now illustrate the logic and value of the MOPS construction in each of the two decision problems (A^1, P^1) and (A^2, P^2) of our running example. Applying the result to characterize possible learning above and beyond the revealed information structure generally requires specifying the utility function u, which by the subsequent characterization (9) of Section 5.3 reduces to picking a scalar $u(z_M) \in [0.6, 0.8]$ with the normalization that $u(z_B) = 0$ and $u(z_G) = 1$. However, the characterization of possible learning in the first decision problem A^1 is independent of this value $u(z_M)$ because the corresponding prize z_M is never realized under actions a^1 and a^2. Given that only actions a^1, a^2 are available in A^1, the key sets of posteriors at which each action is optimal are given by:

$$\begin{aligned} \hat{\Gamma}(a_1|A^1) &= [0.5, 1]; \\ \hat{\Gamma}(a_2|A^1) &= [0, 0.5]. \end{aligned}$$

Theorem 1 states that an information structure can rationalize the data in decision problem 1 if and only if it can be obtained by spreading the revealed posteriors across posteriors that preserve optimality at each associated action. More specifically, the posteriors that a Markov matrix B permits from revealed posterior $\bar{\gamma}^{a_1} = 0.8$ must all preserve optimality of a_1, hence be in the range $[0.5, 1.0]$, while those permitted from revealed posterior $\bar{\gamma}^{a_2} = 0.2$ must all preserve optimality of a_2, hence be in the range $[0, 0.5]$.

Figure 1 illustrates such a construction for an information structure Q^1 defined as:

$$\begin{matrix} \gamma & Q^1(\gamma) \end{matrix} \\ \begin{pmatrix} 1 & 3/10 \\ 1/2 & 2/5 \\ 0 & 3/10 \end{pmatrix} \tag{14}$$

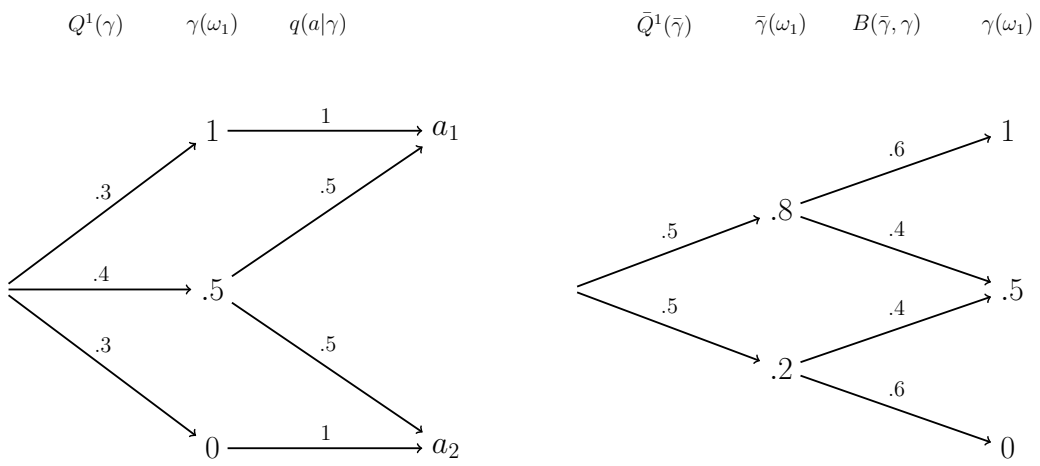

Figure 1: Rationalizing (left) and MOPS (right) constructions of an information structure Q^1 from the distribution of revealed posteriors $\bar{Q}^1$. Given (inferred) prize utilities $u(z_B) = 0$ and $u(z_G) = 1$, action a_1 is preferred to a_2 for posteriors $\gamma(\omega_1) \geq 0.5$, and action a_2 is preferred to a_1 for posteriors $\gamma(\omega_1) \leq 0.5$. The rationalizing action strategy $q(a|\gamma)$ is consistent with these preferences. Conversely, the corresponding MOPS function B spreads from revealed posteriors $\bar{\gamma}^{a_1} = 0.8$ and $\bar{\gamma}^{a_2} = 0.2$ to other posteriors where the optimality of the respective actions is preserved.

This information structure can be recovered as a MOPS of the revealed information structure $\bar{Q}^1$ by a mean and optimality preserving spread from the revealed posterior .8 of action a_1 to (optimality-preserving) posteriors .5 and 1, and from revealed posterior .2 of action a_2 to posteriors 0 and .5, with weights such that each revealed posterior is also preserved. Note that the spread involves mass from each revealed posterior on posterior .5, at which both actions a_1 and a_2 are optimal. Conversely, the MOPS construction guarantees an optimal action strategy such that data P^1 is rationalized by Q^1. In particular, the action strategy is mixed at revealed posterior .5, where again both actions are optimal.

The key distinction in decision problem A^2 is that feasible learning depends on the utility function, specifically the parameter $u(z_M)$, through the sets of posteriors inducing optimal actions:

$$\begin{aligned}
\hat{\Gamma}(a_1|A^2, u) &= [u(z_M), 1]; \\
\hat{\Gamma}(a_3|A^2, u) &= [1 - u(z_M), u(z_M)]; \\
\hat{\Gamma}(a_2|A^2, u) &= [0, 1 - u(z_M)].
\end{aligned}$$

Thus, an information structure may be feasible only for a subset of utility functions consistent with the choice model. This point is perhaps even more apparent upon characterizing what could have been learned in terms of its informational value.

3.2 The (Limited) Value of Information

We now briefly consider a simple alternative characterization of what could have been learned in terms of the information value. The value-based characterization will be especially useful when we consider what could have been learned with additional restrictions across multiple decision problems in Section 4. Following Blackwell (1951), we say for information structures $Q, \bar{Q} \in \mathcal{Q}$ that Q is *as (Blackwell) informative as* $\bar{Q}$, denoted $Q \succeq \bar{Q}$, if:

$$G(Q|A,u) \geq G(\bar{Q}|A,u) \quad \forall\, u : A \times \Omega \to \mathbb{R}.^{2} \tag{15}$$

For a given utility function u, define the *revealed gross utility* as:

$$\bar{G}(u) \equiv \sum_{a \in \mathcal{A}} \sum_{\omega \in \Omega} u(z(a,\omega)) P(a,\omega) \tag{16}$$

Analogously to Theorem 1, we then have the following lemma, which will underlie our approach to characterizing learning across decision problems in Section 4.

Lemma 1. *Fix a decision problem with data (A,P) and revealed information structure $\bar{Q}$. Given utility function u, the following are equivalent for an information structure Q:*

1. *There exists an optimal action strategy q such that (Q,q) rationalizes the data P according to EU maximization.*
2. *The information structure Q is as informative as the revealed information structure $\bar{Q}$ and yields maximal gross utility equal to what is revealed:*

$$G(Q|A,u) = \bar{G}(u) \tag{17}$$

Intuitively, condition (17) results from the combination of two binding inequalities. First is that the maximal gross utility at the revealed information structure is at least as high as the revealed gross utility:

$$G(\bar{Q}|A,u) \geq \bar{G}(u)$$

with equality if and only if the data satisfies the No Improving Action Switch (NIAS) condition of CM15. Second is that the maximal gross utility of what is learned is at least as high as what is revealed:

$$G(Q|A,u) \geq G(\bar{Q}|A,u)$$

by the ranking of Blackwell informativeness, with equality when the information structures are equally valuable in the decision problem A. Interesting subtleties in the value-based characterization will arise in the following Section 4 when we consider learning inferred from richer data encompassing more than one decision problem. In particular, what is learned within one decision problem may be relatively more valuable in another decision problem, which imposes significant additional constraints.

[2] Note this specification of utility u is consistent with our prize-based definition when $z(a,\omega) \equiv (a,\omega)$.

4 Rationalizing Across Decision Problems

We now consider what could have been learned across a finite set of $M > 1$ decision problems consisting of action sets $\mathbf{A} \equiv (A^1, \ldots, A^M)$ and generating corresponding SDSC data $\mathbf{P} \equiv (P^1, \ldots, P^M)$, which can be equivalently represented as corresponding revealed information structures $\bar{\mathbf{Q}} \equiv (\bar{Q}^1, \ldots, \bar{Q}^M)$. Such richer choice data allow us to test and estimate models of learning, which additionally impose structure on the acquisition of information. Furthermore, such models combined with richer data impose constraints on what could have been learned.

We begin by characterizing the set of information structures $\mathbf{Q} \equiv (Q^1, \ldots, Q^M)$ for which there exist respective mixed action strategies $\mathbf{q} \equiv (q^1, \ldots, q^M)$ satisfying expected utility maximization as in (6) and generating the observed data sets as in (3), and for which there exists a single cost function $K : \mathcal{Q} \to \mathbb{R} \cup \{\infty\}$ such that the choice of information is optimal according to (8). In this case we say that the set of action strategies $(\mathbf{Q}, \mathbf{q})$ rationalize the sets of observed data $\mathbf{P}$ according to costly learning, and that $\mathbf{Q}$ is a viable tuple of what could have been learned.

4.1 Costly Learning and the Indirect Value Difference Function

To motivate our approach to characterizing what could have been learned in multiple decision problems, consider again the information structure Q^1 defined previously in (14). As shown in Figure 1, this information structure is a MOPS of the revealed information structure $\bar{Q}^1$ and is thus consistent with expected utility maximization in decision problem 1. Yet, a model of information acquisition such as (8) imposes additional constraints on learning across decision problems. We now ask: could this information structure still have been learned in decision problem 1 under costly learning given the data (A^2, P^2) observed in decision problem 2?

The answer is no because Q^1 is too valuable to rationalize what was (not) learned in decision problem 2. Consider first the revealed value of what was learned in each decision problem. By Lemma 1, the definition (16), and the utility bounds (9), we can compute the gross utility for any learned information structures Q^1 and Q^2 in respective decision problems 1 and 2 as:

$$G(Q^1|A^1, u) = \bar{G}^1(u) = 0.8$$
$$G(Q^2|A^2, u) = \bar{G}^2(u) = 0.45 + 0.5u(z_M).$$

Given a specification (14) for information structure Q^1, we can also compute its gross value in choice set A^2:

$$G(Q^1|A^2, u) = 0.6 + 0.4u(z_M)$$

Finally, the gross value of learning Q^2 in decision problem A^1 is bounded below by the value of the revealed information structure $\bar{Q}^2$ by the Blackwell informativeness order (Lemma 1):

$$G(Q^2|A^1, u) \geq G(\bar{Q}^2|A^1, u) = 0.75$$

Next, consider the sum of revealed and counterfactual gross utilities when learning across the decision problems is switched. The sum of gross utilities revealed within decision problem is:

$$G(Q^1|A^1, u) + G(Q^2|A^2, u) = 1.25 + 0.5u(z_M)$$

Analogously, the sum of counterfactual gross utilities from switching learning across decision problems is:

$$G(Q^2|A^1, u) + G(Q^1|A^2, u) \geq 1.35 + 0.4u(z_M)$$

It follows that the sum of counterfactual gross utilities from switching learning across decision problems exceeds the sum of revealed gross utilities for all possible $u(z_M) \in [0.6, 0.8]$.[3]

This violates the costly learning model by a similar logic to CD15: a cycle of learning across decision problems is always feasible and furthermore holds fixed the sum of learning costs across decision problems. Thus, information Q^1 could not have been learned in decision problem 1 in conjunction with the data observed in decision problem 2. Unlike CD15, however, this is not a statement about model testability, but rather about what could have been learned. In particular, the observed data remains consistent with a costly learning representation.

We now introduce the *indirect value difference function* for summarizing these rich restrictions on learning in a simple, computationally tractable matrix form. To this end, we first define the *(direct) value difference* between the value of learning Q and the revealed gross utility in decision problem m:

$$D_0^m(Q|u) \equiv G(Q|A^m, u) - \bar{G}^m(u), \tag{18}$$

Note that we now index decision-problem-specific objects by their decision problem m. By Proposition 1, a necessary condition for an information structure Q^m to be rationalizable within decision problem m is that $D_0^m(Q^m|u) = 0$.

The value difference construction becomes eminently useful in summarizing the rich counterfactual attention strategies that must be considered across decision problems. To this end, we formalize the notion of an attention path and cycle. An *attention path* $\vec{h}$ of edge length $1 \leq J(\vec{h}) \leq M$ is a vector of decision problem indices $\vec{h} = (h^1, h^2, \ldots, h^{J(\vec{h})+1})$, with $1 \leq h^j \leq M$ and with the first $J(\vec{h})$ entries unique. An *attention cycle* is an attention path

[3] Specifically, as long as $u(z_M) < 1$.

where the first and last entries coincide: $h^{J(\vec{h})+1} = h^1$. For arbitrary decision problem indices $1 \leq m, n \leq M$, let:

$$H(m,n) \equiv \{\vec{h} \in H | h^1 = m, h^{J(\vec{h})+1} = n\}$$

denote the subset of attention paths that start at m and end at n. Maximizing the sum of direct value differences across each set of attention paths $H(m,n)$ for $1 \leq m, n \leq M$ then yields the central object of our method of information recovery, an $M \times M$ matrix for each set of information structures that we call the *indirect value difference function*. Evaluated at a point, we also refer to this object as the indirect value difference matrix.

Definition 2. *Given utility function u, the* ***indirect value difference function*** $D(\mathbf{Q}|u)$ *is, for each set of information structures $\mathbf{Q} \in \mathcal{Q}^M$, an $M \times M$ matrix defined element-wise as:*

$$D^{mn}(\mathbf{Q}|u) \equiv \max_{\vec{h} \in H(m,n)} \sum_{j=1}^{J(\vec{h})} D_0^{h^j}(Q^{h^{j+1}} \,|\, u) \tag{19}$$

Intuitively, the indirect value difference function sums up the changes in maximized expected utility along a path in which each decision problem A^{h^j} is shifted to attention strategy $Q^{h^{j+1}}$ corresponding to one higher index, with subsequent re-optimization of actions.

The indirect value difference function reflects a strong analogy with revealed preference theory — specifically, with the problem of computing the transitive closure of a revealed preference relation in order to test whether a finite set of price and choice data are consistent with utility maximization (Afriat 1967, Varian 1982). Two key differences stem from the richness and observability of choices in our informational setting. First, our counterfactual comparisons involve switching attention choices between action sets and re-optimizing actions. Second, the choice of attention is only partially identified by the action choice data, which requires generalizing the value difference matrices as functions of imperfectly observed learning. Nevertheless, as in Varian (1982), we can employ the Floyd-Warshall algorithm (Floyd 1962, Warshall 1962) to compute the indirect value difference matrix for a given set of information structures in polynomial time; further details are provided in Appendix B.

For a set of information structures to have been learned across decision problems, it is necessary that any such sum of value differences across an attention cycle be non-positive. The maximal sum of such changes is also non-negative, since the identity attention cycle mapping each information structure to its original decision problem is feasible. This yields a necessary condition, captured by the diagonal entries of the indirect value difference matrix, that we refer to as Generalized No Improving Cycles (G-NIC):

$$\text{diag}(D(\mathbf{Q}|u)) = 0 \tag{G-NIC}$$

where $\text{diag}(\cdot)$ denotes the operator recovering the main diagonal entries of a square matrix. Fundamentally, (G-NIC) resembles the No Improving Attention Cycle (NIAC) condition of

CD15, but generalized in two ways. First, it is a function of an arbitrary set of information structures $\mathbf{Q}$. Second, by including edges to and from the same node and computing value differences relative to the revealed gross utility $\bar{G}^m(u)$ rather than the gross utility at revealed information $G(\bar{Q}^m|A^m, u)$, the condition (combined with Blackwell informativeness) also subsumes the within-decision-problem value constraints required by Lemma 1, including the No Improving Action Switches (NIAS) condition of CM15 and the within-problem MOPS property of Theorem 1.

Since all information structures not chosen in some decision problem could simply be infeasible without further assumptions, this condition becomes jointly sufficient for characterizing what could have been learned once combined with the informativeness condition required for an expected utility rationalization within decision problem. The following result summarizes the characterization of what could have been learned across decision problems.

Theorem 2. *Fix a set of decision problems with data* $(\mathbf{A}, \mathbf{P})$ *and corresponding revealed information structures* $\bar{\mathbf{Q}}$*. For a given utility function* u*, the following are equivalent:*

1. *There exists a learning cost function* K *and a set of action strategies* $\mathbf{q}$ *such that* $(\mathbf{Q}, \mathbf{q})$ *rationalizes the data sets* $\mathbf{P}$ *according to costly learning.*

2. *The information structures* $\mathbf{Q}$ *satisfy* (G-NIC) *and are each as Blackwell informative as their revealed counterparts* $\bar{\mathbf{Q}}$*.*

An appealing feature of the value difference characterization of Theorem 2 is that it circumvents the need to specify corresponding optimal action strategies $\mathbf{q}$. Still, the proof of rationalization requires such a construction, which follows by Lemma 1. In addition to its analytical and computational tractability, we show in the following Section 5 how the indirect value difference function also encodes valuable information about the possible learning costs and the nature of learning.

5 Rationalizing Learning

We now consider the question of *why* a set of candidate information structures $\mathbf{Q}$ could have been learned. We explore this question in three complementary ways. First, in Subsection 5.1, we derive the identified set of all information cost functions that rationalize such learning. Second, in Subsection 5.2, we further elucidate the nature of learning by considering consistency with two important nested classes of the costly learning model, namely capacity constrained learning and fixed information. In each case, we emphasize the importance of our value difference and MOPS constructions. Finally, in Subsection 5.3 we recover the set of possible utility functions for applications in which this is not known.

5.1 Costs of Learning

We begin with recovery of learning costs as a function of a utility function u and a set of learned information structures $\mathbf{Q}$. For each such combination, we derive the identified set of all cost functions $K : \mathcal{Q} \to \mathbb{R} \cup \{\infty\}$ that rationalize the data. This includes costing both what was and was not learned.

Theorem 3. *Given a utility function u and a rationalizable tuple of what was learned $\mathbf{Q} = (Q^1, \ldots, Q^M)$, the sharp identified set of learning cost functions consists of cost functions that rationalize what was learned:*

$$K(Q^n) - K(Q^m) \geq D^{mn}(\mathbf{Q}|u) \tag{20}$$

and rationalize what was not learned:

$$K(Q) \geq \max_{1 \leq m \leq M} [D_0^m(Q|u) + K(Q^m)]. \tag{21}$$

for all $1 \leq m, n, \leq M$ and $Q \in \mathcal{Q}$.

The first condition (20) characterizes the learning cost function evaluated at what was learned. In particular, this condition is necessary and sufficient for what was learned in each decision problem to have been optimal, relative to what was learned in other decision problems. The second condition (21) places lower bounds on the cost of what was not learned to ensure suboptimality relative to what was learned.

Theorem 3 extends the related cost recovery of CD15 in several ways. First, it recovers the sharp identified set of costs conditional on any combination of information structures $\mathbf{Q}$ that could have been learned across decision problems, rather than only the sharp set corresponding to the revealed information structures $\bar{\mathbf{Q}}$. Note, however, that the bounds corresponding to the revealed information structures $\bar{\mathbf{Q}}$ nest those of any information structures $\mathbf{Q}$ that could have been learned because the bounding indirect value difference function is increasing in the (element-wise) informativeness of its information structures (Lemma 4):

$$D(\mathbf{Q}|u) \geq D(\bar{\mathbf{Q}}|u).$$

Thus, the bounds for the revealed information structures remain valid for candidate learned information structures but may cease to be sharp conditioning on what was learned. In the following Subsection 5.2, for example, the unconditional bounds would typically be insufficient for determining which (if any) information structures could have been learned in nested variants of the model. Second, Theorem 3 also characterizes the possible cost functions on the remainder of their domain.

Perhaps most importantly, however, the characterization of Theorem 3 highlights the centrality of our indirect value difference function $D(\mathbf{Q}|u)$. This object operationalizes the

literature through the computational simplicity of the underlying matrix calculations using the Floyd-Warshall algorithm, as described in Appendix B. Additionally, the indirect value difference function itself encodes a variety of cost functions, including a representative cost function, on the domain of what was learned. In order to succinctly state these results, let $\mathcal{K}^M(\mathbf{Q}, u) \subseteq \mathbb{R}^M$ denote the set of rationalizing cost functions restricted to the domain of what was learned and expressed as a vector superscripted by decision problem m.

Proposition 1 (Indirect Value Difference Matrix as Costs of What Was Learned)**.** *Suppose the information structures* $\mathbf{Q}$ *are rationalizable under costly learning. Then:*

1. *For each* $1 \leq m \leq M$*, the row* $D^{m*}(\mathbf{Q}|u)$ *and sign-inverted column* $-D^{*m}(\mathbf{Q}|u)$ *are respectively the minimum and maximum elements (and thus extreme points) of the subset of* m*-normalized costs of what was learned:*

$$\{\tilde{K} \in \mathcal{K}^M(\mathbf{Q}, u) : \tilde{K}^m = 0\} \tag{22}$$

2. *The average of their midpoints across* $1 \leq m \leq M$ *is a "representative" cost function of what was learned, obtained as the average of row means and sign-inverted column means:*

$$\frac{1}{2M}\left[\sum_{m=1}^{M} D^{m*}(\mathbf{Q}|u) - \sum_{m=1}^{M} D^{*m}(\mathbf{Q}|u)\right] \tag{23}$$

The first part of Proposition 1 shows that each column and row have an interpretation as a (negative) cost function of what was learned, which are furthermore extremal among a subset of normalized costs. The second part of Proposition 1 uses this fact and the convexity of $\mathcal{K}^M(\mathbf{Q}, u)$ to derive a single, representative cost function on the domain of what was learned from the indirect value difference function. Of course, ours is not the only possible method of defining a specific feasible cost function as representative, and the literature on costly learning contains alternative definitions. Notably, Denti (2022) (following Rockafellar 1973) defines and (partially) identifies the minimal monotone cost function as the solution to a linear program.

We now recover learning costs in the running example. For simplicity, we focus on relative learning costs of the revealed information structures $\bar{Q}^1$ and $\bar{Q}^2$ for the normalized utility function (9), as derived in Subsection 5.3. Following Theorem 3, we begin by constructing the indirect value difference matrix $D(\bar{\mathbf{Q}}|u)$. This depends on the direct value differences (18), which in turn depend on realized and counterfactual gross utilities $\bar{G}^m(u)$ and $G(\bar{Q}^n|A^m, u)$ for all decision problems $1 \leq m, n \leq M$. The gross utilities within decision problem can be computed as:

$$\begin{aligned} G(\bar{Q}^1|A^1, u) &= \bar{G}^1(u) = 0.8 \\ G(\bar{Q}^2|A^2, u) &= \bar{G}^2(u) = 0.45 + 0.5u(z_M) \end{aligned} \tag{24}$$

which is consistent with condition (17) of Lemma 1 for rationalizability within decision problem. Additionally, the gross utilities from switching revealed information structures across decision problems can be computed as:

$$\begin{aligned} G(\bar{Q}^2|A^1, u) &= 0.75 \\ G(\bar{Q}^1|A^2, u) &= 0.8 \end{aligned} \tag{25}$$

From these computed gross utilities, we can then obtain the indirect value difference matrix at revealed information as:

$$D(\bar{\mathbf{Q}}|u) = \begin{pmatrix} 0 & -0.05 \\ 0.35 - 0.5u(z_M) & 0 \end{pmatrix} \tag{26}$$

In our simple example, the indirect value difference matrix follows readily upon observing its elementwise equality with the direct value differences:

$$D^{mn}(\bar{\mathbf{Q}}|u) = D_0^m(\bar{Q}^n) \equiv G(\bar{Q}^n|A^m, u) - \bar{G}(u)$$

for $1 \leq m, n \leq 2$. In turn, this equality between direct and indirect value differences follows from two facts. First, the off-diagonal entries must be equal because the only feasible path involves a single attention (and action) switch across the decision problems. Second, the diagonal entries are zero because the summed value of the attention cycle $0.35 - 0.5u(z_M) - 0.05$ is less than zero when $u(z_M) \geq 0.6$.

Condition (20) of Theorem 3 then yields the bounds:

$$K(\bar{Q}^1) - K(\bar{Q}^2) \in [0.35 - 0.5u(z_M), 0.05] \tag{27}$$

In our simple example where the direct and indirect value differences coincide, these bounds also correspond exactly to those from the pairwise incentive compatibility constraints on learning across decision problem:

$$\begin{aligned} 0.8 - K(\bar{Q}^1) &\geq 0.75 - K(\bar{Q}^2); \\ 0.45 + 0.5u(z_M) - K(\bar{Q}^2) &\geq 0.8 - K(\bar{Q}^1). \end{aligned}$$

The set of cost differences (27) is non-empty for all $u(z_M) \in [0.6, 0.8]$. When $u(z_M) = 0.8$, there is a wide range of rationalizing cost differences,

$$K(\bar{Q}^1) - K(\bar{Q}^2) \in [-0.05, 0.05].$$

When $u(z_M) = 0.6$ the only rationalizing cost difference is $K(\bar{Q}^1) - K(\bar{Q}^2) = 0.05$..

In addition to tractably summarizing bounds on cost functions, the indirect value difference function encodes viable learning costs in its (average) rows and columns by Proposition 1. Namely, as a function of the utility parameter, each row:

$$D^{1*}(\bar{\mathbf{Q}}|u) = (0, -0.05), \quad D^{2*}(\bar{\mathbf{Q}}|u) = (0.35 - 0.5u(z_M), 0)$$

and sign-inverted column:

$$-D^{1*}(\bar{\mathbf{Q}}|u) = (0, -0.35 + 0.5u(z_M)), \quad -D^{2*}(\bar{\mathbf{Q}}|u) = (0.05, 0)$$

is a constrained extremal cost function restricted to revealed learning, and their average (23) yields a representative such cost:

$$(0.1 - 0.125u(z_M), 0.125u(z_M) - 0.1)$$

Notably, the representative cost places higher cost on $\bar{Q}^1$ than $\bar{Q}^2$ for almost all feasible $u(z_M) \in [0.6, 0.8)$, even though this need not be true for a cost function to rationalize the data.

5.2 Nested Models of Learning

Complementing cost recovery, we may also be interested in what this implies about the nature of learning. For example, can the observed data be rationalized under alternative models of learning and, if so, what could have been learned? In this section, we use our machinery to address these questions for two additional important classes of learning model, which are nested based on additional structure of the cost function K. First, a *capacity constrained* model entails cost-free acquisition of information from an exogenous feasible set. This model is a special case of costly learning where the cost function indicates feasibility:

$$K : \mathcal{Q} \to \{0, \infty\}. \tag{28}$$

Second, in a *fixed information* model, information is exogenous to the decision problem, and the only choice is how to use that information across different problems. This model is a special case of capacity constraints (28) with a single-valued feasible set:

$$|\{Q \in \mathcal{Q} : K(Q) < \infty\}| = 1. \tag{29}$$

We now consider these models in turn.

The characterization of what could have been learned in a capacity constrained model adopts our value difference approach. By a simple revealed preference argument, any (attention and/or action) switch cannot raise expected utility relative to what was revealed. This is captured through the direct value difference construction as a non-positivity requirement:

$$D_0^m(Q^n|u) \leq 0$$

for all $1 \leq m, n \leq M$. This condition is also expressible in terms of our *in*direct value difference function, which we refer to as Generalized No Improving (Attention or Action) Switches (G-NIS):

$$D(\mathbf{Q}|u) \leq 0 \tag{G-NIS}$$

This condition is sufficient for possible learning once combined with the informativeness restriction required for generating the observed data.

Proposition 2. *Fix a set of decision problems with data* $(\mathbf{A}, \mathbf{P})$ *and corresponding revealed information structures* $\bar{\mathbf{Q}}$*. For a given utility function* u*, the following are equivalent:*

1. *There exists a feasibility indicator function* $K : \mathcal{Q} \to \{0, \infty\}$ *and a set of action strategies* $\mathbf{q}$ *such that* $(\mathbf{Q}, \mathbf{q})$ *rationalizes the data sets* $\mathbf{P}$ *according to capacity constrained learning.*

2. *The information structures* $\mathbf{Q}$ *satisfy* (G-NIS) *and are each as informative as their revealed counterparts* $\bar{\mathbf{Q}}$*.*

An advantage of expressing the condition in terms of the indirect value difference matrix is that Proposition 2 follows immediately as a corollary of Theorem 3. By (20), only in this case could the set of information structures have been learned under identically zero costs:

$$K(Q^1) = \cdots = K(Q^M) = 0,$$

and the remaining information structures in $\mathcal{Q}$ can always be taken to be infeasible. Additionally, the indirect value difference expression of (G-NIS) clearly nests (G-NIC) because the main diagonal of $D(\mathbf{Q}|u)$ is by construction non-negative, since the identity cycle is always feasible. Note, however, that (G-NIS) holds only if the direct and indirect value differences coincide:

$$D^{mn}(\mathbf{Q}|u) = D_0^m(Q^n|u)$$

for all $1 \leq m, n \leq M$. Conversely, equality of the direct and indirect value differences does not imply (G-NIS), as evidenced by the indirect value difference matrix (26) constructed in the preceding section.

We now show in the running example how Proposition 2 restricts both what could have been learned and the set of possible prize utilities. Consider the information structure:

$$\begin{array}{cc} \gamma & Q(\gamma) \\ \begin{pmatrix} 1 \\ 1/2 \\ 1/8 \end{pmatrix} & \begin{pmatrix} 3/10 \\ 3/10 \\ 2/5 \end{pmatrix} \end{array}$$

for which we can compute the gross utilities across decision problems as:

$$\begin{aligned} G(Q|A^1, u) &= 0.8 \\ G(Q|A^2, u) &= 0.65 + 0.3u(z_M). \end{aligned}$$

Also, observe that Q is as informative as the revealed information $\bar{Q}^1$ in decision problem 1. For the candidate set of information structures $\mathbf{Q} = (Q, \bar{Q}^2)$, we can combine these gross utilities with the revealed gross utilities and the gross value of revealed information $\bar{Q}^2$ from (24) and (25) to construct the indirect value difference matrix as:

$$D(\mathbf{Q}|u) = \begin{pmatrix} \max\{0.15 - 0.2u(z_M), 0\} & -0.05 \\ 0.2 - 0.2u(z_M) & \max\{0.15 - 0.2u(z_M), 0\} \end{pmatrix}$$

By Theorem 2 and the diagonal restrictions, this set of information structures $\mathbf{Q}$ is consistent with costly learning for $u(z_M) \geq 0.75$; yet, by Proposition 2 and the off-diagonal restrictions, it is not consistent with capacity constrained learning for any $u(z_M) < 1$, which includes all possible prize utility values $u(z_M) \in [0.6, 0.8]$ from (9). In summary, the information is inconsistent with capacity constrained learning and consistent with costly learning only for a subset of possible utility functions. Similarly, the costly information model is only consistent with the subset of utility functions $u(z_M) \geq 0.7$ for which the revealed indirect value difference matrix (26) is non-positive. This follows from Lemma 4, which implies that the indirect value difference matrix is smallest at the revealed information structures. We return more systematically to this point and the implications for utility recovery in the following Subsection 5.3.

Next, recall that the fixed information model assumes the existence of a single feasible information structure, independently of incentives or decision problem. Fixed information can be trivially incorporated into the value difference result for capacity constrained learning (Proposition 2) by imposing the additional constraint that all information structures are equal:

$$Q^1 = \cdots = Q^M.$$

Yet, in this case, the indirect value difference construction is largely redundant because there are no optimality restrictions on the choice of information across decision problem. For the same reason, the MOPS construction becomes operable even in the presence of choice data from multiple decision problems. To simplify notation, let $\bar{\mathcal{Q}}^m(u)$ denote the set of MOPS of revealed information $\bar{Q}^m$ in decision problem A^m under utility function u. Applying our previous MOPS characterization (Theorem 1) across decision problems immediately yields the following result.

Proposition 3. *Fix a set of decision problems with data* $(\mathbf{A}, \mathbf{P})$ *and corresponding revealed information structures* $\bar{\mathbf{Q}}$*. For a given utility function* u*, the following are equivalent:*

1. *There exists a single information structure* Q *and a set of action strategies* $\mathbf{q}$ *that rationalize the data sets* $\mathbf{P}$ *according to fixed information.*

2. The information structure Q is a MOPS of each revealed information structure:

$$Q \in \bigcap_{m=1}^{M} \bar{\mathcal{Q}}^m(u)$$

We now illustrate in the running example how this MOPS characterization operationalizes the construction of rationalizing information sets.[4] We begin by arguing that any fixed information structure Q in the running example must place exactly probability 0.25 on the certain posterior $\gamma = \gamma(\omega_1) = 1$ that the state is ω_1:

$$Q(1) = 0.25$$

This follows from two observations in decision problem 2. First, $Q(1) \geq 0.25$ since the revealed information structure places this probability $\bar{Q}^2(1) = 0.25$, which cannot be recovered as a mixture of other posteriors. Second, $Q(1) \leq 0.25$ since it is strictly optimal to choose a_1 at this posterior, and the probability of choosing action a_1 in decision problem 2 is $P^2(a_1) = 0.25$.

This leaves a probability mass 0.25 of uncommitted posterior probabilities that can be spread from revealed posterior $\bar{\gamma}_1^{a_1}$ in (subscripted) decision problem 1. We now argue that this remaining mass must all stem from posteriors in $[0.5, 0.8]$, and the average posterior in this range must be 0.6:

$$\sum_{\gamma \in [0.5,0.8]} Q(\gamma) = 0.25, \qquad \sum_{\gamma \in [0.5,0.8]} \gamma Q(\gamma) = 0.25 \times 0.6$$

First, to preserve optimality of a^1 in choice set A^1, this mass can be spread on $[0.5, 1]$. However, this mass cannot be spread to posteriors in the range $(0.8, 1]$ because, as before, optimality in choice set A^2 would imply a strictly higher probability of choosing action a_1 than observed in data set P^2. Second, the average revealed posterior $\bar{\gamma}_1^{a_1}$ in decision problem 1 is 0.8, and action a^1 is chosen in this decision problem with probability $P^1(a^1) = 0.5$. Since we already deduced that fixed information must satisfy $Q(1) = 0.25$, this implies that the spread posteriors must average to 0.6 for the remaining probability mass 0.25.

Next, we argue that a fixed information structure must put mass 0.5 on posterior 0.2:

$$Q(0.2) = 0.5.$$

The choice of action a_2 in both choice sets A^1 and A^2 has a common revealed posterior $\bar{\gamma}_1^{a_2}(\omega_1) = \bar{\gamma}_2^{a_2}(\omega_1) = 0.2$, but a probability $P^1(a_2) = 0.5$ in choice set A^1 rather than a probability $P^2(a_2) = 0.25$ in choice set A^2. Furthermore, optimality implies that the set of posteriors at which a_2 is chosen in A^2 is a subset of those where a_2 is chosen in A^1, since

[4] By a similar token, this characterization could also be used to rule out fixed information representations.

the upper posterior cutoff $1 - u(z_M) \leq 0.4$ for action a_2 in choice set A^2 is lower than the upper posterior cutoff 0.5 in choice set A^1. Consider then the set of posteriors at which a_2 is chosen in A^1 but not in A^2. By optimality, their support has a lower bound of 0.2, the lowest posterior at which it could be optimal not to choose a_2 in A^2. If, however, there is positive probability on any posteriors above 0.2, then removing mass from these posteriors in choice set A^1 relative to A^2 would strictly decrease the revealed posterior of a_2 in A^1, which it does not. Therefore a_3 must be chosen in A^2 at posterior 0.2. In this case, the posteriors at which a_2 is chosen in A^1 and A^2 are also bounded above by 0.2, which is only possible when 0.2 is the only posterior at which a_2 is chosen. Therefore the only possibility is that $Q(0.2) = 0.5$, as desired.

Finally, the average posterior other than 0.2 at which a_3 is chosen in data set P^2 must be 0.6 to rationalize the revealed posterior $\bar{\gamma}_2^{a_3} = 0.4$ in decision problem 2. However, this was already implied previously and therefore adds no further restrictions. We conclude that there are no more restrictions. Figure 2 illustrates a fixed information rationalization with $Q(0.6) = 0.25$. By the preceding arguments, we can generalize this example in only one respect. We can spread the mass on posterior 0.6 to any set of posteriors on the support $[0.5, 0.8]$ in a mean-preserving way and then set the corresponding strategy of deterministically choosing a_1 at all such posteriors in A^1 and a_3 at all such posteriors in A^2. This also implies that the information structure in Figure 2 is the least informative fixed information structure that can rationalize the observed data.

We conclude by noting how the fixed information model can impose additional restrictions on prize utilities, even beyond those of the capacity constrained model. In particular, the preceding derivation showed that a_3 must be chosen in A^2 with positive probability at posterior 0.2 in any fixed information rationalization; in turn, such a choice is only optimal for the normalized utility function where $u(z_M) = 0.8$. Thus, the utility function is effectively point identified in the example under a model of fixed information, and the requisite identification arguments arise naturally in the derivation of a fixed information structure using MOPS. Next, we propose a generalized method of recovering prize utilities for the costly learning model and its capacity constrained variant.

5.3 Prize Utilities

So far, our leading characterizations of what could have been learned took the utility function as given. We now consider recovery of the prize utility function in cases where it is not a known primitive of the empirical application.

A starting point for utility recovery is the question of utility *consistency*: namely, when is a utility function consistent with some costly learning (or capacity constrained) represen-

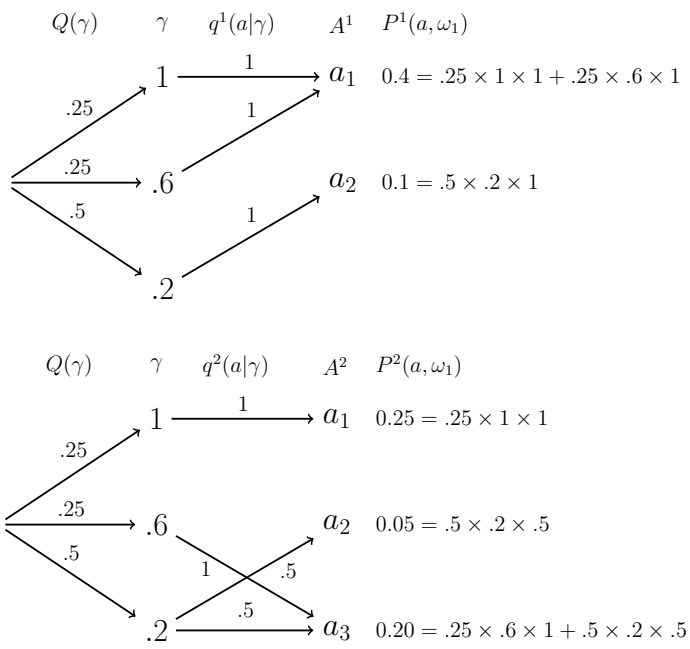

Figure 2: Existence of a common distribution of posteriors and mixed action strategies q^1, q^2 that rationalize the data sets P^1, P^2.

tation? Our methods provide a simple answer to this preliminary question in terms of the indirect value difference function evaluated at the revealed information structures.

Proposition 4 (Consistency of a Known Utility Function)**.** *Given data set* $(\mathbf{A}, \mathbf{P})$*, a utility function* $u : Z \to \mathbb{R}$ *is consistent with costly learning if and only if there exist No Improving Cycles at the revealed set of information:*

$$\text{diag}(D(\bar{\mathbf{Q}}|u)) = 0 \tag{30}$$

The utility function is consistent with capacity constrained learning if and only if there exist No Improving Switches at the revealed set of information:

$$D(\bar{\mathbf{Q}}|u) \leq 0 \tag{31}$$

These results follow respectively (and immediately) from Theorem 2 and Proposition 2, combined with the revealed lower bound on $D(\cdot|u)$ from Lemma 4.

The first condition (30) distills and operationalizes the consistency characterization (Theorem 1) of CD15 in terms of the indirect value difference matrix at revealed information. Namely, it can be verified by definition that (30) is equivalent to the combination of the No Improving Action Switch (NIAS) and No Improving Attention Cycle (NIAC) conditions of Caplin and Martin (2015) and CD15, respectively. The second statement provides the first general representation theorem for capacity constrained learning, which underlies vast

literatures in psychology, cognition, and neuroscience. The idea embodied in condition (31) is that there can exist No Improving Switches in revealed information across (or within) decision problem upon re-optimizing actions.

The main contribution of this subsection is to show how we can use the implicit and existential results of Proposition 4 to obtain explicit expressions for the sets of utility functions consistent with the data and the respective models. To do so, we extend a geometric approach, introduced by Caplin and Martin (2021), that summarizes the consistency restrictions in the space of utilities and (changes in) prize lotteries. More specifically, differences between realized and feasible prize lotteries generate a cone, whose dual cone equals the set of feasible utility functions. Intuitively, this approach recovers feasible utility functions because the inequalities defining the dual cone rule out improvements from counterfactual attention and action strategies. As in the characterization of CD15 and the above Proposition 4, for the sake of utility recovery it suffices to proceed as if revealed information structures coincide with those that are learned.

In that case, we define an *attention and action switch of revealed data* as a morphism $s : A^n \to A^m$ from a source choice set A^n to target action set A^m. Let S^{mn} denote the set of such switches, and let $S \equiv \cup S^{mn}$ denote the set of all such switches across $1 \leq m, n \leq M$. Technically, our only reason for defining an attention and action switch as a morphism rather than as a function is to preserve information about a switch's domain and codomain, so that we can without ambiguity refer to m^s and n^s as the target and source decision problems associated with switch s, which simplifies subsequent notation. Intuitively, s switches attention in decision problem m^s with attention in decision problem n^s and deterministically switches actions across the decision problem choice sets. An *attention and action cycle of revealed data* c comprises an attention cycle $\vec{h}$ combined with a set of corresponding attention and action switches $s^j \in S^{h^j h^{j+1}}$ for $1 \leq j \leq J(\vec{h})$, where (recall) $J(\vec{h})$ denotes the length of the cycle. Let C denote the set of such attention and action cycles.

The characterizations of utility functions that admit a costly information or capacity constrained representation are based on ruling out utility improvements from attention and action switches and cycles, respectively. For each switch $s \in S$ and with a slight abuse of previous notation, define the counterfactual switched dataset as:

$$P^s(a, \omega) \equiv \sum_{b \in s^{-1}(a)} P^{n^s}(b, \omega) \tag{32}$$

where $s^{-1}(a) = \{b \in A^{n^s} : s(b) = a\}$ denotes the set of actions in A^{n^s} that are switched by s to a. We can then define a revealed and a counterfactual lottery associated with each switch as follows. Define the lottery revealed in decision problem m, $\bar{L}^m = (\bar{L}^m_1, \ldots, \bar{L}^m_K) \in \Delta(Z)$,

on a prize-by-prize basis $1 \leq k \leq K$ by:

$$\bar{L}_k^m \equiv \sum_{a \in \mathcal{A}} \sum_{\omega \in \Omega} P^m(a, \omega) I\{z(a, \omega) = z_k\} \tag{33}$$

where $I\{z(a, \omega) = z_k\}$ is an indicator function that the prize in action and state (a, ω) is z_k. Analogously, define the lottery over prizes associated with the switched dataset by:

$$L_k^s \equiv \sum_{a \in \mathcal{A}} \sum_{\omega \in \Omega} P^s(a, \omega) I\{z(a, \omega) = z_k\}. \tag{34}$$

Despite more elaborate notation, we can compute similar lotteries for each attention and action cycle $c \in C$. The main difference from the lotteries associated with switches is that we now average over all (respectively revealed and switched) lotteries in the attention cycle. Again proceeding prize-by-prize and with a slight abuse of previous notation, the revealed and counterfactual lotteries associated with the cycle c are defined respectively as:

$$\bar{L}_k^c \equiv \frac{1}{J(\vec{h}^c)} \sum_{j=1}^{J(\vec{h}^c)} \bar{L}_k^{h^{cj}} \qquad \text{and} \qquad L_k^c \equiv \frac{1}{J(\vec{h}^c)} \sum_{j=1}^{J(\vec{h}^c)} L_k^{s^{cj}} \tag{35}$$

where $1 \leq h^{cj} \leq M$ is the jth node in the attention cycle of c, and s^{cj} is the associated attention and action switch.

In each case, the difference between the revealed lottery and counterfactual lottery gives the change in prize distributions that could have been obtained under an alternative attention and action strategy, relative to what was chosen. Allowing for positively weighted combinations of such differences defines a convex cone, which in turn defines a dual (and also convex) cone. The *lottery cone of attention and action switches* is the convex cone $\mathcal{S} \subseteq \mathbb{R}^K$ formed by all changes in prize lotteries associated with attention and action switches:

$$\mathcal{S} \equiv \Big\{ \sum_{s \in S} \alpha^s [\bar{L}^{m^s} - L^s] \,\Big|\, \alpha \in \mathbb{R}_+^{|S|} \Big\},$$

with its dual cone $\mathcal{S}^* \subseteq \mathbb{R}^K$ defined as:

$$\mathcal{S}^* \equiv \{u \in \mathbb{R}^K | x \cdot u \geq 0 \quad \forall x \in \mathcal{S}\}$$

Analogously, the *lottery cone of attention and action cycles* is the convex cone $\mathcal{C} \subseteq R^K$ formed by all summed changes in prize lotteries associated with attention and action cycles:

$$\mathcal{C} \equiv \Big\{ \sum_{c \in C} \alpha^c [\bar{L}^c - L^c] \,\Big|\, \alpha \in \mathbb{R}_+^{|C|} \Big\}.$$

with its dual cone $\mathcal{C}^* \subseteq \mathbb{R}^K$ defined as:

$$\mathcal{C}^* \equiv \{u \in \mathbb{R}^K | x \cdot u \geq 0 \quad \forall x \in \mathcal{C}\}$$

Our key result of utility recovery is that the sets of utility functions consistent with costly information and capacity constrained learning are given by the dual cones of attention and action cycles and switches, respectively.

Proposition 5 (Recovering Prize Utilities)**.** *The set $\mathcal{U}^{CI}$ of utility functions consistent with costly information is equal to the dual cone of attention and action cycles:*

$$\mathcal{U}^{CI} = \mathcal{C}^*$$

The set $\mathcal{U}^{CC}$ of utility functions consistent with capacity constrained learning is equal to the dual cone of attention and action switches:

$$\mathcal{U}^{CC} = \mathcal{S}^*$$

Intuitively, the change in prize distributions for a feasible strategy cannot strictly increase expected utility relative to what was chosen. For capacity constrained learning, for example, any attention and action switch is feasible because the requisite attention strategy is feasible, given that it was chosen in some decision problem. Thus, a necessary condition for a utility function u to admit a capacity constrained learning representation is that:

$$[\bar{L}^{m^s} - L^s] \cdot u \geq 0 \quad \forall s \in S. \tag{36}$$

For costly learning, any attention and action cycle holds fixed the aggregate cost of learning across its included decision problems, and therefore the cycle cannot increase gross utility relative to what was actually chosen. Given that each cone $\mathcal{C}$ and $\mathcal{S}$ comprises only positively weighted combinations of the corresponding lottery differences, the respective inequalities à la (36) extend to the entire cones. By definition, therefore, the sets of consistent utility functions are exactly given by the *dual* cones.

In essence, Proposition 5 is just an alternative way of geometrically expressing Proposition 4 along the lines of Caplin and Martin (2021), where Proposition 4 (in the case of costly information) in turn summarizes the previous consistency characterization of CD15. Nevertheless, from the standpoint of recovery, this is a conceptually useful distinction. Furthermore, a subtlety arising in the present context is that the dual logic proceeded in terms of direct action switches and cycles of the *revealed* data. It is less apparent how (or whether) such a dual characterization could be attained for, say, a fixed information representation or for learned information structures distinct from those that were directly revealed.

We now illustrate the utility cones in our running example, introduced in Subsection 1.1. The revealed lotteries over the three prizes (z_G, z_M, z_B) associated with the decision problems $m = 1, 2$ are respectively given by:

$$\bar{L}^1 = (0.8, 0, 0.2) \tag{37}$$

$$\bar{L}^2 = (0.45, 0.5, 0.05) \tag{38}$$

For the sake of exposition, we begin by ruling out improving action switches, holding attention fixed. This corresponds to the No Improving Action Switch condition of Caplin and

Martin (2015) and is required for any Bayesian expected utility maximization representation, including costly information and capacity constrained representations. Within choice set A^1, the possible action switches are choosing a_2 in place of a_1 and vice versa. In each case, the counterfactual prize lotteries attained yield $(0.2, 0, 0.8)$. The difference between actual and counterfactual lotteries is therefore $(0.6, 0, -0.6)$. Given that choices are optimal, we conclude that prize utilities $(u(z_G), u(z_M), u(z_B))$ must have a positive dot product with this difference vector, so that $u(z_G) \geq u(z_B)$.

In choice set A^2, consider first the action switches from a_1. When chosen, a_1 yields pure prize z_G and utility $u(z_G)$. Switching from a_1 to a_2 is not strictly improving since it yields z_B, and already $u(z_B) \leq u(z_G)$. Switching from a_1 to a_3 would yield z_M, which implies that $u(z_G) \geq u(z_M)$. Next, consider action switches from a_2. Switching from a_2 to a_1 cannot be improving since it only lowers the probability of receiving z_G rather than z_B. For a switch from a_2 to a_3 to be non-improving requires that the lottery $(0.8, 0, 0.2)$ resulting from a_2 be at least as good as the pure prize z_M achievable by switching to a_3,

$$0.8u(z_G) + 0.2u(z_B) \geq u(z_M).$$

Finally, consider switches from a_3, which produces pure prize z_M. For the best alternative — to switch to a_2 and get the lottery $(0.6, 0, 0.4)$ — to be non-improving requires that:

$$u(z_M) \geq 0.6u(z_G) + 0.4u(z_B).$$

Combining the above inequalities we confirm that the only non-trivial rationalization involves $u(z_G) > u(z_M) > u(z_B)$, as foreshadowed by the good, medium, and bad prize subscripts. Normalizing $u(z_G) = 1$ and $u(z_B) = 0$, the preceding implications are summarized by the condition that $u(z_M) \in [0.6, 0.8]$.

The remaining conditions required for capacity constrained learning involve also switching attention across decision problems. First, consider switching attention from decision problem 2 to 1. Given the preceding conclusions, at each posterior it is best to choose the action in A^1 that is more likely to yield prize z_G rather than z_B. The corresponding counterfactual lottery is $(0.75, 0, 0.25)$, while the chosen lottery was $\bar{L}^1 = (0.8, 0, 0.2)$. Hence there is a difference of $(0.05, 0, -0.05)$ between the chosen and counterfactual lotteries. For the counterfactual lottery to be non-improving, it must be that $u(z_G) \geq u(z_B)$, which was already established previously from considering only action switches.

The implications of switching attention from decision problem 1 to 2 are equally simple: given the bounds $u(z_M) \leq 0.8$ already established, the maximum utility derives from picking the action that yields lottery $(0.8, 0, 0.2)$. The chosen lottery was $\bar{L}^2 = (0.45, 0.5, 0.05)$. Hence there is a difference of $(-0.35, 0.5, -0.15)$ between the chosen and counterfactual lotteries. For the counterfactual lottery to be non-improving, it must be that:

$$u(z_M) \geq 0.7.$$

Combining with the preceding conditions, the conditions for capacity constrained learning with normalization $u(z_G) = 1$ and $u(z_B) = 0$ are summarized by $u(z_M) \in [0.7, 0.8]$.

Finally, consider the remaining conditions required for costly learning. The attention and action cycle constraints prevent utility rising only when both attention switches occur simultaneously, because in this case learning costs cancel out of the equation. Consider the arithmetic average of the change in lottery when attention is switched across the two decision problems, with subsequent optimization in action. Based on the separate implications of each such attention switch discussed previously, the averaged change in lotteries is:

$$\frac{(-0.35, 0.5, -0.15) + (0.05, 0, -0.05)}{2} = (-0.15, 0.25, -0.1).$$

Given what is already known, the corresponding inequality on utilities for which this yields positive utility is:

$$u(z_M) \geq 0.6,$$

which adds no new constraints over those already derived. In summary, the conditions for costly learning with normalization $u(z_G) = 1$ and $u(z_B) = 0$ are summarized by $u(z_M) \in [0.6, 0.8]$. These are precisely the conditions (9) which were taken as given throughout the preceding analysis.

References

Afriat, Sydney N (1967). "The construction of utility functions from expenditure data". In: *International economic review* 8.1, pp. 67–77.

Blackwell, David (1951). "Comparison of experiments". In: *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, pp. 93–102.

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

Caplin, Andrew and Mark Dean (2015). "Revealed preference, rational inattention, and costly information acquisition". In: *American Economic Review* 105.7, pp. 2183–2203.

Caplin, Andrew, Mark Dean, and John Leahy (2017). *Rationally inattentive behavior: Characterizing and generalizing Shannon entropy*. Tech. rep. National Bureau of Economic Research.

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

— (2021). "Comparison of decisions under unknown experiments". In: *Journal of Political Economy* 129.11, pp. 3185–3205.

Chambers, Christopher P, Ce Liu, and John Rehbeck (2017). "Nonseparable Costly Information Acquisition and Revealed Preference". In.

Denti, Tommaso (2022). *Posterior-separable cost of information*. Tech. rep. working paper.

Floyd, Robert W (1962). "Algorithm 97: shortest path". In: *Communications of the ACM* 5.6, p. 345.

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

Koopmans, Tjalling C and Martin Beckmann (1957). "Assignment problems and the location of economic activities". In: *Econometrica* 25.1, pp. 53–76.

Lipnowski, Elliot and Doron Ravid (2022). "Predicting Choice from Information Costs". In: *arXiv preprint arXiv:2205.10434*.

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

Matejka, Filip and Alisdair McKay (2015). "Rational inattention to discrete choices: A new foundation for the multinomial logit model". In: *American Economic Review* 105.1, pp. 272–98.

Perez-Richet, Eduardo (2017). "A Proof of Blackwell's Theorem". In.

Rochet, Jean-Charles (1987). "A Necessary and sufficient condition for rationalizability in a quasi-linear context". In: *Journal of Mathematical Economics* 16, pp. 191–200.

Rockafellar, R Tyrrell (1970). *Convex analysis*. Princeton University Press, Princeton, NJ, USA.

— (1973). "The multiplier method of Hestenes and Powell applied to convex programming". In: *Journal of Optimization Theory and Applications* 12.6, pp. 555–562.

Sims, Christopher A (2003). "Implications of Rational Inattention". In: *Journal of Monetary Economics* 50.3, pp. 665–690.

Varian, Hal R (1982). "The nonparametric approach to demand analysis". In: *Econometrica: Journal of the Econometric Society*, pp. 945–973.

Warshall, Stephen (1962). "A theorem on boolean matrices". In: *Journal of the ACM (JACM)* 9.1, pp. 11–12.

Woodford, Michael (2014). "Stochastic choice: An optimizing neuroeconomic model". In: *American Economic Review* 104.5, pp. 495–500.

A Proofs

A.1 Theorem 1 and Lemma 1

We prove Theorem 1 and Lemma 1 simultaneously to clarify the three-way Blackwellian equivalence between signal garbling, mean preserving spreads, and the value of information, specialized in our setting to optimal choice. As the proof makes clear, a rationalizing action strategy q is effectively a garbling of the revealed information structure, which is furthermore optimal in the sense that it only puts weight on actions that are optimal given posteriors.

As a preliminary, we use Lemma 2 to collect two observations on the gross value of learning function $G(Q|A,u)$ for use in what follows.

Lemma 2. *The gross value of learning is equivalently defined in terms of pure actions:*

$$G(Q|A,u) = \sum_{\gamma\in\Gamma(Q)} Q(\gamma)\left[\max_{a\in A} U(a|\gamma,u)\right] \tag{39}$$

The gross value of learning revealed information weakly exceeds revealed gross utility:

$$G(\bar{Q}|A,u) \geq \bar{G}(u) \tag{40}$$

Proof of Lemma 2. By definitions (5) and (7),

$$G(Q|A,u) \equiv \max_{q:\Gamma(Q)\to\Delta(A)} \sum_{\gamma\in\Gamma(Q)}\sum_{a\in\mathcal{A}} Q(\gamma)q(a|\gamma)U(a|\gamma,u).$$

This optimization problem is a linear program, consequently with optimal solutions at its extreme points. Therefore it is without loss of generality for defining the value function to restrict to extreme points, which yields exactly (39). For (40),

$$\begin{aligned}
\bar{G}(u) &= \sum_{a\in\mathcal{A}}\sum_{\omega\in\Omega} P(a,\omega)u(z(a,\omega)) && \text{by (16)}\\
&= \sum_{a\in\mathcal{A}}\sum_{\omega\in\Omega} P(a)\bar{\gamma}^a(\omega)u(z(a,\omega)) && \text{by (1)}\\
&= \sum_{a\in\mathcal{A}} P(a)\sum_{\omega\in\Omega} \bar{\gamma}^a(\omega)u(z(a,\omega)) && \text{by rearrangement}\\
&= \sum_{a\in\mathcal{A}} P(a)U(a|\bar{\gamma}^a,u) && \text{by (4)}\\
&\leq \sum_{a\in\mathcal{A}} P(a)\left[\max_{b\in A} U(b|\bar{\gamma}^a,u)\right] && \text{by optimization}\\
&= \sum_{\bar{\gamma}\in\Gamma(\bar{Q})} \bar{Q}(\bar{\gamma})\left[\max_{b\in A} U(b|\bar{\gamma},u)\right] && \text{collecting terms}\\
&= G(\bar{Q}|A,u). && \text{by (39)}
\end{aligned}$$

□

Proof of Theorem 1 and Lemma 1. In combination, the two results establish the following three-way equivalence. Fix a decision problem with data (A, P) and revealed information structure $\bar{Q}$. Given utility function u, the following are equivalent for an information structure Q:

1. There exists an optimal action strategy q such that (Q, q) rationalizes the data P according to EU maximization.
2. The information structure Q is a mean and optimality preserving spread of $\bar{Q}$.
3. The information structure Q is as informative as the revealed information structure $\bar{Q}$ and yields maximal gross utility equal to what is revealed.

The proof proceeds by showing that (3) $\implies$ (2) $\implies$ (1) $\implies$ (3).

(3 $\implies$ 2) Suppose that Q is as informative as $\bar{Q}$ and yields maximal gross utility equal to what is revealed (17). Since Q is assumed at least as valuable as $\bar{Q}$, Blackwell's Theorem (Blackwell 1953, Theorem 2) implies the existence of a function $B : |\Gamma(\bar{Q})| \times |\Gamma(Q)| \to [0, 1]$ defining Q as a mean preserving spread of $\bar{Q}$. For the sake of contradiction, suppose that this function B does not preserve the optimality (13) additionally required for a MOPS (Definition 1). Then there exists a chosen action $P(a) > 0$ and a posterior $\gamma \in \Gamma(Q)$ such that a is not optimal at spread posterior γ:

$$P(a) > 0, \quad B(\bar{\gamma}^a, \gamma) > 0, \quad \gamma \notin \hat{\Gamma}(a|A, u) \tag{41}$$

This contradicts the assumption that Q yields maximal gross utility equal to what is revealed

(17) because:

$$\begin{aligned}
\bar{G}(u) &= \sum_{a\in A}\sum_{\omega\in\Omega} u(z(a,\omega))P(a,\omega) && \text{by (16)}\\
&= \sum_{a\in A}\sum_{\omega\in\Omega} u(z(a,\omega))P(a)\bar{\gamma}^a(\omega) && \text{by (1)}\\
&= \sum_{a\in A}\sum_{\omega\in\Omega} u(z(a,\omega))P(a)\sum_{\gamma\in\Gamma(Q)} B(\bar{\gamma}^a,\gamma)\gamma(\omega) && \text{by (12)}\\
&= \sum_{a\in A}\sum_{\gamma\in\Gamma(Q)} P(a)B(\bar{\gamma}^a,\gamma)\sum_{\omega\in\Omega}\gamma(\omega)u(z(a,\omega)) && \text{by rearranging}\\
&= \sum_{a\in A}\sum_{\gamma\in\Gamma(Q)} P(a)B(\bar{\gamma}^a,\gamma)U(a|\gamma,u) && \text{by (4)}\\
&< \sum_{a\in A}\sum_{\gamma\in\Gamma(Q)} P(a)B(\bar{\gamma}^a,\gamma)\left[\max_{b\in A} U(b|\gamma,u)\right] && \text{by (41)}\\
&= \sum_{\gamma\in\Gamma(Q)}\left[\max_{b\in A} U(b|\gamma,u)\right]\sum_{a\in A} P(a)B(\bar{\gamma}^a,\gamma) && \text{by rearranging}\\
&= \sum_{\gamma\in\Gamma(Q)}\left[\max_{b\in A} U(b|\gamma,u)\right]\sum_{\bar{\gamma}\in\Gamma(\bar{Q})}\bar{Q}(\bar{\gamma})B(\bar{\gamma},\gamma) && \text{by collecting terms}\\
&= \sum_{\gamma\in\Gamma(Q)}\left[\max_{b\in A} U(b|\gamma,u)\right]Q(\gamma) && \text{by (11)}\\
&= G(Q|A,u) && \text{by (7)}
\end{aligned}$$

(2 $\implies$ 1) Suppose there exists a function $B : |\Gamma(\bar{Q})| \times |\Gamma(Q)| \to [0,1]$ defining Q as a mean and optimality preserving spread (MOPS) of $\bar{Q}$. We use the function B to construct a mixed strategy $q : \Gamma(Q) \to \Delta(A)$ such that (Q,q) generates the data (3) in an optimal way (6). For this, it suffices to restrict to chosen actions $P(a) > 0$, since for other actions rationalization is achieved by setting $q(a|\gamma) = 0$ for any posterior γ.

We now establish that the following mixed strategy $q : \Gamma(Q) \to \Delta(A)$ has the property that it combines with Q to generate the data and does so while focused only on optimal choices:

$$q(a|\gamma) = \frac{P(a)B(\bar{\gamma}^a,\gamma)}{Q(\gamma)} \tag{42}$$

Note first that q as defined in (42) are mixed strategies by construction since they are nonnegative and summing the numerators across actions yields the denominator:

$$\sum_{a\in A} P(a)B(\bar{\gamma}^a,\gamma) = \sum_{a\in A} q(a|\gamma)Q(\gamma) = Q(\gamma).$$

To confirm that (Q,q) rationalizes the data, note given any chosen action $P(a) > 0$ and

state ω:

$$\begin{aligned}
P_{(Q,q)}(a,\omega) &= \sum_{\gamma\in\Gamma(Q)} q(a|\gamma)Q(\gamma)\gamma(\omega) && \text{by (2)}\\
&= \sum_{\gamma\in\Gamma(Q)} P(a)B(\bar{\gamma}^a,\gamma)\gamma(\omega) && \text{by (42)}\\
&= P(a)\sum_{\gamma\in\Gamma(Q)} B(\bar{\gamma}^a,\gamma)\gamma(\omega) && \text{by rearranging}\\
&= P(a)\bar{\gamma}^a(\omega) && \text{by (12)}\\
&= P(a,\omega). && \text{by (1)}
\end{aligned}$$

Finally, we show that the mixed strategy q identified above only chooses actions at posteriors where they are optimal, as in (6). By construction (42), $q(a|\gamma) > 0$ implies $B(\bar{\gamma}^a,\gamma) > 0$. By the defining property (13) of a mean *and optimality* preserving spread,

$$q(a|\gamma) > 0 \implies \gamma \in \hat{\Gamma}(a|A,u). \tag{43}$$

In that case, we have:

$$\begin{aligned}
g(q,Q) &= \sum_{\gamma\in\Gamma(Q)}\sum_{a\in\mathcal{A}} Q(\gamma)q(a|\gamma)U(a|\gamma,u) && \text{by (5)}\\
&= \sum_{\gamma\in\Gamma(Q)}\sum_{a\in\mathcal{A}} Q(\gamma)q(a|\gamma)\left[\max_{b\in A} U(b|\gamma,u)\right] && \text{by (43)}\\
&= \sum_{\gamma\in\Gamma(Q)} Q(\gamma)\left[\max_{b\in A} U(b|\gamma,u)\right]\sum_{a\in\mathcal{A}} q(a|\gamma) && \text{by rearrangement}\\
&= \sum_{\gamma\in\Gamma(Q)} Q(\gamma)\left[\max_{b\in A} U(b|\gamma,u)\right] && \text{because } q(\cdot|\gamma)\in\Delta(A)\\
&= G(Q|A,u) && \text{by (39)}
\end{aligned}$$

which implies by definition of the gross value of learning function (7) that q is optimal for choice set A given u.

(2 $\implies$ 1) Suppose there exists an action strategy q such that (Q,q) rationalizes the data according to EU maximization (6). First, observe from rationalization that for all $\bar{\gamma}\in\Gamma(\bar{Q})$

and $\omega \in \Omega$ we have:

$$\begin{aligned}
\bar{Q}(\bar{\gamma})\bar{\gamma}(\omega) &= \sum_{a:\bar{\gamma}^a=\bar{\gamma}} P(a)\bar{\gamma}(\omega) && \text{collecting terms} \\
&= \sum_{a:\bar{\gamma}^a=\bar{\gamma}} P(a,\omega) && \text{by (1)} \\
&= \sum_{a:\bar{\gamma}^a=\bar{\gamma}} P_{(Q,q)}(a,\omega) && \text{by (3)} \\
&= \sum_{a:\bar{\gamma}^a=\bar{\gamma}} \sum_{\gamma\in\Gamma(Q)} q(a|\gamma)Q(\gamma)\gamma(\omega) && \text{by (2)} \\
&= \sum_{\gamma\in\Gamma(Q)} \left[\sum_{a:\bar{\gamma}^a=\bar{\gamma}} q(a|\gamma)\right] Q(\gamma)\gamma(\omega) && \text{by rearranging}
\end{aligned}$$

The outer equality relates the joint distributions over posteriors and states (i.e. Blackwell experiments with posteriors as signal realizations) $\bar{Q}(\bar{\gamma})\bar{\gamma}(\omega)$ and $Q(\gamma)\gamma(\omega)$ via the garbling function:

$$f(\bar{\gamma}|\gamma) \equiv \sum_{a:\bar{\gamma}^a=\bar{\gamma}} q(a|\gamma).$$

By Blackwell's Theorem (Blackwell 1953, Theorem 5),[5] this implies that information structure Q is at least as valuable as $\bar{Q}$. It remains to show that Q yields maximal gross utility equal to what is revealed (17):

$$\begin{aligned}
G(Q|A,u) &\geq G(\bar{Q}|A,u) && \text{by informativeness} \\
&\geq \bar{G}(u) && \text{by (40)} \\
&= \sum_{a\in A}\sum_{\omega\in\Omega} u(z(a,\omega))P(a,\omega) && \text{by (16)} \\
&= \sum_{a\in A}\sum_{\omega\in\Omega} u(z(a,\omega))P_{(Q,q)}(a,\omega) && \text{by (3)} \\
&= \sum_{a\in A}\sum_{\omega\in\Omega} u(z(a,\omega)) \sum_{\gamma\in\Gamma(Q)} q(a|\gamma)Q(\gamma)\gamma(\omega) && \text{by (2)} \\
&= \sum_{\gamma\in\Gamma(Q)} Q(\gamma) \sum_{a\in A} q(a|\gamma) \sum_{\omega\in\Omega} \gamma(\omega)u(z(a,\omega)) && \text{by rearranging} \\
&= \sum_{\gamma\in\Gamma(Q)} Q(\gamma) \sum_{a\in A} q(a|\gamma)U(a|\gamma,u) && \text{by (4)} \\
&= g(q,Q|u) && \text{by (5)} \\
&= G(Q|A,u)
\end{aligned}$$

where the last step follows from the starting assumption that q was optimal (6) for choice set A given utility u. □

[5] For a statement and proof of the result in our notation, see also Perez-Richet (2017).

A.2 Theorems 2 and 3

We begin by isolating the cyclically monotone aspect of our proof argument for Theorem 2 in a separate Lemma 3, which is essentially attributable to Rochet (1987) in the context of implementation in a quasi-linear context; see also Koopmans and Beckmann (1957) in the context of optimal assignment problems, Rockafellar (1970) in the context of subdifferentials of convex functions, and Caplin and Dean (2015) for a preceding implementation of this logic for testing the costly information acquisition model.

Lemma 3 (Rochet 1987, Theorem 1)**.** *For a tuple of information structures* $\mathbf{Q} \equiv (Q^1, \ldots, Q^M)$*, there exists a* $\tilde{K} \in \mathbb{R}^M$ *satisfying:*

$$G(Q^m|A^m, u) - \tilde{K}^m \geq G(Q^n|A^m, u) - \tilde{K}^n \qquad \forall\, 1 \leq m, n \leq M \tag{44}$$

if and only if:

$$\max_{\vec{h} \in H(m,m)} \sum_{j=1}^{J(\vec{h})} G(Q^{h^{j+1}}|A^{h^j}, u) - G(Q^{h^j}|A^{h^j}, u) = 0 \quad \forall\, 1 \leq m \leq M \tag{45}$$

Proof of Lemma 3. For completeness, we repeat the short and constructive proof of Rochet (1987) (itself adapted from arguments in the proof of Theorem 24.8 in Rockafellar (1973)) using our notation and indexing. First, suppose there exists a $\tilde{K} \in \mathbb{R}^M$ satisfying (44). Rearrangement yields:

$$G(Q^n|A^m, u) - G(Q^m|A^m, u) \leq \tilde{K}^n - \tilde{K}^m \qquad \forall\, 1 \leq m, n \leq M$$

Summing over any cycle $\vec{h}$ of indices,

$$\sum_{j=1}^{J(\vec{h})} [G(Q^{h^{j+1}}|A^{h^j}, u) - G(Q^{h^j}|A^{h^j}, u)] \leq \sum_{j=1}^{J(\vec{h})} [\tilde{K}^{h^{j+1}} - \tilde{K}^{h^j}] = 0$$

Since a length-1 cycle $h(1) = h(2)$ achieves the zero bound, this implies (45).

Conversely, assume (45) and define as a function of the index $1 \leq m \leq M$,

$$V(m) = \max_{\vec{h} \in H(1,m)} \sum_{j=1}^{J(\vec{h})} G(Q^{h^{j+1}}|A^{h^j}, u) - G(Q^{h^j}|A^{h^j}, u)$$

Note that the maximum exists in our case since there are only finitely many cycles. For any $1 \leq m, n \leq M$, the construction implies:

$$V(m) \geq V(n) + G(Q^m|A^n, u) - G(Q^n|A^n, u)$$

Defining $\tilde{K}^m \equiv G(Q^m|A^m, u) - V(m)$ and substituting yields (44). □

Proof of Theorem 2. Suppose there exists a learning cost function K and a set of action strategies $\mathbf{q}$ such that $(\mathbf{Q}, \mathbf{q})$ rationalizes the data sets $\mathbf{P}$ according to costly learning. By Lemma 1, rationalizability within decision problem implies that Q^m is as informative as $\bar{Q}^m$ and yields revealed utility (17):

$$G(Q^m|A^m, u) = \bar{G}^m(u)$$

for all $1 \leq m \leq M$. By definition, rationalizability across decision problem requires the existence of a learning cost function $K : \mathcal{Q} \to \mathbb{R} \cup \{\infty\}$ satisfying:

$$G(Q^m|A^m, u) - K(Q^m) \geq G(Q^n|A^m, u) - K(Q^n) \quad \forall\, 1 \leq m, n \leq M$$

which by Lemma 3 implies cyclic monotonicity (45). Plugging revealed utility from (17) into this condition and substituting for direct and direct value difference functions (18) and (19) yields the (G-NIC) condition because:

$$\begin{aligned}
0 &= \max_{\vec{h} \in H(m,m)} \sum_{j=1}^{J(\vec{h})} G(Q^{h^{j+1}}|A^{h^j}, u) - G(Q^{h^j}|A^{h^j}, u) \\
&= \max_{\vec{h} \in H(m,m)} \sum_{j=1}^{J(\vec{h})} G(Q^{h^{j+1}}|A^{h^j}, u) - \bar{G}^{h^j}(u) \\
&= \max_{\vec{h} \in H(m,m)} \sum_{j=1}^{J(\vec{h})} D_0^{h^j}(Q^{h^{j+1}}|u) \\
&= D^{mm}(\bar{Q}|u)
\end{aligned}$$

Conversely, suppose the information structures $\mathbf{Q}$ satisfy (G-NIC) and are each as informative as their revealed counterparts $\bar{\mathbf{Q}}$. For any $1 \leq m \leq M$ and length-1 cycle $h(1) = h(2) = m$, we have:

$$\sum_{j=1}^{J(\vec{h})} D_0^{h^j}(Q^{h^{j+1}}|u) = D_0^m(Q^m|u)$$

Since each set of cycles $H(m, m)$ includes such a cycle, (G-NIC) implies:

$$D_0^m(Q^m|u) \equiv G(Q^m|A^m, u) - \bar{G}^m(u) \leq 0$$

By informativeness $Q^m \succeq \bar{Q}^m$ and condition (40) of Lemma 2, we also have:

$$G(Q^m|A^m, u) \geq G(\bar{Q}^m|A^m, u) \geq \bar{G}^m(u)$$

Combining the preceding two equalities yields:

$$G(Q^m|A^m, u) = \bar{G}^m(u)$$

which combined with informativeness implies rationalizability within decision problem according to expected utility maximization (6), by Lemma 1. Additionally, plugging this equality into (G-NIC) yields cyclic monotonicity (45). In turn, by Lemma 3 this implies the existence of a $\tilde{K} \in \mathbb{R}^M$ satisfying (44). For rationalizability across decision problems, it then suffices to define a cost function $K : \mathcal{Q} \to \mathbb{R} \cup \{\infty\}$ to equal $\tilde{K}^m$ when evaluated at Q^m, $1 \leq m \leq M$, and to equal infinity otherwise. □

Proof of Theorem 3. Suppose the data $\mathbf{P}$ is rationalized according to costly learning by information structures $\mathbf{Q} = (Q^1, \ldots, Q^M)$ in combination with a cost function $K : \mathcal{Q} \to \mathbb{R} \cup \{\infty\}$. Rationalizability across decision problems implies:

$$G(Q^m|A^m, u) - K(Q^m) \geq G(Q|A^m, u) - K(Q) \quad \forall\, 1 \leq m \leq M, Q \in \mathcal{Q} \tag{46}$$

Upon substituting for revealed gross utility by rationalizability within decision problem (Lemma 1), rearranging, and substituting again for direct value difference (18), we obtain:

$$K(Q) - K(Q^m) \geq D_0^m(Q|u) \quad \forall\, 1 \leq m \leq M, Q \in \mathcal{Q}$$

Chaining such inequalities for any attention path $\vec{h} \in H(m, n)$ among chosen information structures implies:

$$K(Q^n) - K(Q^m) = \sum_{j=1}^{J(\vec{h})} [K(Q^{h^{j+1}}) - K(Q^{h^j})] \geq \sum_{j=1}^{J(\vec{h})} D_0(Q^{h^{j+1}}|u)$$

which by definition of the indirect value difference function implies (20). For any other information structure $Q \in \mathcal{Q}$ that could have been learned, the inequalities imply:

$$K(Q) \geq D_0^m(Q|u) + K(Q^m) \quad \forall\, 1 \leq m \leq M$$

which yields (21).

Conversely, suppose that K satisfies (20) and (21). Evaluating (20) at $m = n$ implies $D^{mm}(\mathbf{Q}|u) \leq 0$, so that $\mathbf{Q}$ is rationalizable within each decision problem. That the learning cost function K rationalizes learning across decision problems follows from reversing the preceding arguments to conclude (46). □

A.3 Remaining Propositions

Proof of Proposition 1. For the sake of this result, it suffices to restrict the domain of cost functions to what was learned. In particular, any vector $\tilde{K} \in \mathbb{R}^M$ consistent with the constraints on the costs of what was learned (20) can be extended to a cost function on $\mathcal{Q}$ that also satisfies constraints (21) on the costs of what was not learned.

For the first part, it suffices to show the result for an arbitrary row and (negative) column of $D(\mathbf{Q}|u)$, say those indexed by $m = 1$. By Theorem 3, any rationalizing cost function (on the domain of what was learned) must satisfy:

$$\tilde{K}^m \geq \tilde{K}^1 + D^{1m}(\mathbf{Q}|u)$$
$$\tilde{K}^m \leq \tilde{K}^1 - D^{m1}(\mathbf{Q}|u)$$

Among such vectors $\tilde{K}$ satisfying $\tilde{K}^1 = 0$, these inequalities become:

$$D^{1m}(\mathbf{Q}|u) \leq \tilde{K}^m \leq -D^{m1}(\mathbf{Q}|u)$$

or expressed in vector notation,

$$D^{1*}(\mathbf{Q}|u) \leq \tilde{K} \leq -D^{*1}(\mathbf{Q}|u)$$

Thus, $D^{1*}(\mathbf{Q}|u)$ and $-D^{*1}(\mathbf{Q}|u)$ are lower and upper bounds on the set:

$$\{\tilde{K} \in \mathcal{K}^M(\mathbf{Q}|u) : \tilde{K}^1 = 0\}$$

In order for them to be its minimum and maximum elements (and thus extreme points), it remains to confirm that they are indeed elements of this set. We confirm this only for the lower bound $D^{1*}(\mathbf{Q}|u)$, since the arguments for the upper bound are analogous. By condition (20) of Theorem 3, it suffices to verify that the vector satisfies the cost bounds on what was learned:

$$D^{1n}(\mathbf{Q}|u) - D^{1m}(\mathbf{Q}|u) \geq D^{mn}(\mathbf{Q}|u) \quad \forall\, 1 \leq m, n \leq M$$

or, rearranging,

$$D^{1n}(\mathbf{Q}|u) \geq D^{1m}(\mathbf{Q}|u) + D^{mn}(\mathbf{Q}|u) \quad \forall\, 1 \leq m, n \leq M$$

For any $1 \leq m, n \leq M$, consider a pair of paths $(1, \ldots, m)$ and $(m, \ldots, n)$ that, by definition (19) of the indirect value difference function, attain the optimal values of the maximization problems defining the right-hand terms. If the combined path $(1, \ldots, m, \ldots n)$ contains no cycle, then it is a feasible attention path for the maximization problem defining the left-hand term, implying the lower bound. If the combined path does contain a cycle $(1, \ldots, r, \ldots m, \ldots, r, \ldots m)$, then it is a lower bound for any attention path $(1, \ldots, r \ldots, n)$ obtained by cutting out the cycle(s), since by rationalizability a length-1 cycle (r, r) attains

the zero upper bound $D^{rr}(\mathbf{Q}|u) = 0$ among all cycles in $H(r, r)$. In turn, the attention path obtained by eliminating cycles is a feasible attention path for the maximization problem defining the left-hand term, implying the desired lower bound. This proves the first part of the result.

For the second part, observe that the set of rationalizing costs $\mathcal{K}^M(\mathbf{Q}, u)$ is a convex polyhedron. By part 1, each term $D^{m*}(\mathbf{Q}|u)$ and $-D^{*m}(\mathbf{Q}|u)$ is an element of this set. By convexity of the set, their average (23) is also an element of the set. □

Proof of Proposition 2. To begin, observe that (G-NIS) implies (G-NIC) because $\text{diag}(D(\mathbf{Q}|u)) \geq 0$ by construction, and so $D(\mathbf{Q}|u) \leq 0$ implies $\text{diag}(D(\mathbf{Q}|u)) = 0$. Suppose the information structures $\mathbf{Q}$ satisfy (G-NIS) and are each as informative as their revealed counterparts; by Theorem 2 they are rationalizable under costly learning. Furthermore, by the sharp cost bounds of Theorem 3, they are rationalizable by a cost function satisfying $K(Q^m) = 0$ for all $1 \leq m \leq M$ and $K(Q) = \infty$ otherwise. Thus, they are rationalizable in a capacity constrained model.

Conversely, if the information structures are not as informative as their revealed counterparts, then they are not rationalizable within decision problem; if they violate (G-NIS), by Theorem 3 there exists some $1 \leq m, n \leq M$ for which:

$$K(Q^n) - K(Q^m) \geq D^{mn}(\mathbf{Q}|u) > 0$$

which implies they are not rationalized by any cost function satisfying $K(Q^m) = 0$ for all $1 \leq m \leq M$. Thus, they are not rationalizable under capacity constraints. □

Proof of Proposition 3. This result is immediate by construction upon applying the within-problem characterization (Theorem 1) across decision problem. Distinctly from the costly learning and capacity constrained models, the across-problem constraints on learning in the fixed information model arise through common rationalizability, rather than through incentive compatibility constraints on learning across decision problem. Nevertheless, a fixed information model is a special case of capacity constraints with a singleton feasible set, so that the models are nested. □

In order to simplify the proof of Proposition 4 and because it is of inherent interest, we isolate a monotonicity property of the indirect value difference function in a separate Lemma 4.

Lemma 4. *The indirect value difference function is increasing in the Blackwell order. For* $\mathbf{Q}$ *as element-wise informative as* $\bar{\mathbf{Q}}$*,*

$$D(\mathbf{Q}|u) \geq D(\bar{\mathbf{Q}}|u)$$

Thus, among feasible sets of information structures characterized in Theorem 2, the function is minimized at the set of revealed information structures.

Proof of Lemma 4. Fix two tuples of information structures $\mathbf{Q}, \bar{\mathbf{Q}}$ ranked element-wise in the Blackwell order, $Q^m \succeq \bar{Q}^m$ for all $1 \leq m \leq M$. By definition (18) of the direct value difference function and definition (15) of the Blackwell order, we have:

$$D_0^m(Q^n|u) = G(Q^n|A^m, u) - \bar{G}^m(u) \geq G(\bar{Q}^n|A^m, u) - \bar{G}^m(u) = D_0^m(\bar{Q}^n|u)$$

for all $1 \leq m, n \leq M$. The desired inequality then follows elementwise by definition (19) of the indirect value difference function:

$$\begin{aligned} D^{mn}(\mathbf{Q}|u) &= \max_{\vec{h} \in H(m,n)} \sum_{j=1}^{J(\vec{h})} D_0^{h^j}(Q^{h^{j+1}} \,|\, u) \\ &\geq \max_{\vec{h} \in H(m,n)} \sum_{j=1}^{J(\vec{h})} D_0^{h^j}(\bar{Q}^{h^{j+1}} \,|\, u) \\ &= D^{mn}(\bar{\mathbf{Q}}|u) \end{aligned}$$

Finally, the fact that $D(\cdot|u)$ is minimized among all rationalizable tuples at the revealed tuple of information structures $\bar{\mathbf{Q}}$ follows immediately from the fact that this is the element-wise least informative tuple in the rationalizable set, by Theorem 2. □

Proof of Proposition 4. We consider the cases of costly and capacity constrained learning concurrently, since the logic is essentially the same. Suppose $u : Z \to \mathbb{R}$ is consistent with costly or capacity constrained learning, in the sense that there exists some tuple of information structures $\mathbf{Q}$ that rationalizes the observed data under the respective models. By Theorem 2 and Proposition 2, $\mathbf{Q}$ is then element-wise as informative as the revealed information structures $\bar{\mathbf{Q}}$ and satisfies, respectively, (G-NIC) or (G-NIS). By monotonicity of the indirect value difference function (Lemma 4), we have in each case:

$$\begin{aligned} \operatorname{diag}(D(\bar{\mathbf{Q}}|u)) \leq \operatorname{diag}(D(\mathbf{Q}|u)) = 0 \\ D(\bar{\mathbf{Q}}|u) \leq D(\mathbf{Q}|u) \leq 0 \end{aligned}$$

In the first case, we also have by construction that $\operatorname{diag}(D(\bar{\mathbf{Q}}|u)) \geq 0$, which joint with the preceding inequality implies:

$$\operatorname{diag}(D(\bar{\mathbf{Q}}|u)) = 0$$

Since $\bar{\mathbf{Q}}$ is trivially as informative as itself, it follows by Theorem 2 and Proposition 2 that conditions (30) and (31) hold, and thus that the revealed information structures also rationalize the data in each case.

Conversely, suppose a utility function is inconsistent with the costly or capacity constrained model. Since revealed information $\bar{\mathbf{Q}}$ is again trivially as informative as itself, the characterizations of Theorem 2 and Proposition 2 imply that the conditions (30) and (31) must be violated, respectively. □

As a preliminary to the final Proposition 5, it is helpful to isolate two observations. First, the characterization of capacity constrained learning is possible in terms of both the direct or indirect value difference functions. While we have used the indirect version for consistency in the remainder of the paper, the counterfactual switches of lottery for utility recovery will be more easily expressible in terms of the direct function.

Lemma 5. *Condition* (G-NIS) *holds if and only if there is no benefit from any direct switch of information across decision problem:*

$$D_0^m(Q^n|u) \leq 0 \quad \forall\, 1 \leq m, n \leq M \tag{47}$$

In this case, the indirect and direct value difference functions also coincide:

$$D^{mn}(\mathbf{Q}|u) = D_0^m(Q^n|u) \quad \forall\, 1 \leq m, n \leq M \tag{48}$$

Proof of Lemma 5. If (G-NIS) holds, the definition (19) of $D(\mathbf{Q}|u)$ and the fact that the simple path (m, n) is a candidate attention path in $H(m, n)$ imply (47). Conversely, if (47) holds, then this simple path (m, n) is also optimal in $H(m, n)$ for the computation of $D^{mn}(\mathbf{Q}|u)$, which implies the equality (48) and thus also (G-NIS). Note that (48) may hold even when (G-NIS) does not; such an example (26) is provided in Subsection 5.1. □

Additionally, Propositions 4 and 5 are logically equivalent but differ in their perspective. The following Lemma 6 clarifies the relations between their respective objects, gross expected utility and prize lotteries.

Lemma 6. *Gross expected utilities and (switched) lotteries are related by:*

$$\bar{G}^m(u) = \bar{L}^m \cdot u \tag{49}$$

$$G(\bar{Q}^n|A^m, u) = \max_{s \in S^{mn}} L^s \cdot u \tag{50}$$

where $u = (u(z_1), \ldots, u(z_K))$ *is expressed in vector notation.*

Proof of Lemma 6. Beginning with equation (49),

$$\begin{aligned}
\bar{G}^m(u) &= \sum_{a\in\mathcal{A}}\sum_{\omega\in\Omega} P^m(a,\omega)u(z(a,\omega)) && \text{by (16)}\\
&= \sum_{k=1}^{K}\sum_{a\in\mathcal{A}}\sum_{\omega\in\Omega} P^m(a,\omega)I\{z(a,\omega)=z_k\}u(z_k) && \text{by expansion}\\
&= \sum_{k=1}^{K}\bar{L}^m_k u(z_k) = \bar{L}^m\cdot u && \text{by (33)}
\end{aligned}$$

For equation (50),

$$\begin{aligned}
G(\bar{Q}^n|A^m,u) &= \sum_{\bar{\gamma}\in\Gamma(\bar{Q}^n)}\bar{Q}^n(\bar{\gamma})\left[\max_{b\in A^m} U(b|\bar{\gamma},u)\right] && \text{by (39)}\\
&= \sum_{a\in\mathcal{A}} P^n(a)\left[\max_{b\in A^m} U(b|\bar{\gamma}^a_n,u)\right] && \text{decomposing terms}\\
&= \max_{s\in S^{mn}}\sum_{a\in\mathcal{A}} P^n(a)U(s(a)|\bar{\gamma}^a_n,u) && \text{by definition of } s, S^{mn}
\end{aligned}$$

Focusing on the maximand of the last line for a given switch $s\in S^{mn}$,

$$\begin{aligned}
\sum_{a\in\mathcal{A}} P^n(a)U(s(a)|\bar{\gamma}^a_n,u) &= \sum_{a\in\mathcal{A}}\sum_{\omega\in\Omega} P^n(a)\bar{\gamma}^a_n(\omega)u(z(s(a),\omega)) && \text{by (4)}\\
&= \sum_{a\in\mathcal{A}}\sum_{\omega\in\Omega} P^n(a,\omega)u(z(s(a),\omega)) && \text{by (1)}\\
&= \sum_{a\in\mathcal{A}}\sum_{\omega\in\Omega} P^s(a,\omega)u(z(a,\omega)) && \text{by (32)}\\
&= \sum_{k=1}^{K}\sum_{a\in\mathcal{A}}\sum_{\omega\in\Omega} P^s(a,\omega)I\{z(a,\omega)=z_k\}u(z_k) && \text{by expansion}\\
&= \sum_{k=1}^{K} L^s_k u(z_k) = L^s\cdot u && \text{by (34)}
\end{aligned}$$

Plugging back into the preceding block of derivations yields the desired equation (50). □

Proof of Proposition 5. We prove Proposition 5 by establishing its equivalence with Proposition 4. Specifically, $u\in\mathcal{S}^*$ iff u satisfies (G-NIS), and $u\in\mathcal{C}^*$ iff u satisfies (G-NIC).

Beginning with the first equivalence, suppose $u\in\mathcal{S}^*$. By definition of $\mathcal{S}^*$ and since $-[L^s-\bar{L}^{m^s}]\in\mathcal{S}$,

$$[L^s-\bar{L}^{m^s}]\cdot u\le 0 \quad \forall\, s\in S \tag{51}$$

Optimizing over $s\in S^{mn}$,

$$\left[\max_{s\in S^{mn}} L^s\cdot u\right]-\bar{L}^m\cdot u\le 0 \quad \forall\, 1\le m,n\le M$$

Plugging in from (49) and (50) in Lemma 6 and the direct value difference definition (18),

$$D_0^m(\bar{Q}^n|u) \equiv G(\bar{Q}^n|A^m, u) - \bar{G}^m(u) \leq 0 \quad \forall\, 1 \leq m, n \leq M$$

By Lemma 5, this implies (G-NIS). Conversely, suppose (G-NIS). Reversing the preceding logic implies (51), which extends to the entire cone $\mathcal{S}$. Therefore $u \in \mathcal{S}^*$.

For the second equivalence, suppose $u \in \mathcal{C}^*$. By definition of $\mathcal{C}^*$ and since $-[L^c - \bar{L}^c] \in \mathcal{C}$,

$$[L^c - \bar{L}^c] \cdot u \leq 0 \quad \forall\, c \in C \tag{52}$$

Expanding the definitions (35) and multiplying by $J(\vec{h}^c) > 0$,

$$\sum_{j=1}^{J(\vec{h}^c)} \left[L^{s^{cj}} - \bar{L}^{h^{cj}}\right] \cdot u \quad \forall\, c \in C$$

Alternatively enumerating cycles and optimizing,

$$\max_{\vec{h} \in H(m,m)} \sum_{j=1}^{J(\vec{h})} \left[\left[\max_{s \in S^{h^j h^{j+1}}} L^s \cdot u\right] - \bar{L}^{h^j} \cdot u\right] \leq 0 \quad \forall\, 1 \leq m \leq M$$

Plugging in from (49) and (50) in Lemma 6 and the direct and indirect value difference definition (18) and (19),

$$D^{mm}(\bar{\mathbf{Q}}|u) = \max_{\vec{h} \in H(m,m)} \sum_{j=1}^{J(\vec{h})} D_0^{h^j}(\bar{Q}^{h^{j+1}}|u) \leq 0 \quad \forall\, 1 \leq m \leq M$$

which is (G-NIC). Again, reversing the preceding logic implies (52), which extends to the entire cone $\mathcal{C}^*$. Therefore $u \in \mathcal{C}^*$. □

B Floyd–Warshall Algorithm

The Floyd–Warshall algorithm takes as an input a directed graph with weight $W(i,j)$ on the vertex from node i to node j and cycles through these weights for all $1 \leq i, j, k \leq M$, identifying when $W(i,j) > W(i,k) + W(k,j)$ and correspondingly reducing it to equality by setting $W'(i,j) := W(i,k) + W(k,j)$. The key step in using the Floyd–Warshall algorithm for our purposes is to construct a complete weighted directed graph with M nodes, with the weight $W(m,n) = -D_0^m(Q^n|u)$ on the directed edge from node m to node n. By definition,

$$-D^{mn}(\mathbf{Q}|u) \equiv \min_{\{\vec{h} \in H(m,n)\}} \sum_{j=1}^{J(\vec{h})} -D_0^{h^j}(Q^{h^{j+1}}, u)$$

In graph-theoretic terms, $H(m, n)$ identifies the set of all non-repeating directed paths from node m to node n in the graph. For any such path, the sum on the RHS is precisely the sum of these weights. Hence, $D^{mn}(\mathbf{Q}|u)$ defines the minimal sum of weights on all directed paths from m to n, and the Floyd–Warshall algorithm efficiently identifies all such paths.

An important property of the Floyd-Warshall algorithm is that it only recovers the true weighting matrix (in our case, the indirect value difference matrix) when no cycles exist (G-NIC is satisfied). Nevertheless, this is readily verifiable from the diagonal of the algorithm output matrix, which is identically zero if and only if no cycles exist. Thus, while the Floyd-Warshall algorithm may not always recover the matrix of interest, it still suffices for the joint purposes of verifying consistency and, in the case where consistency is satisfied, recovering the true indirect value difference matrix.