Economic Theory
https://doi.org/10.1007/s00199-024-01627-z

RESEARCH ARTICLE

Rationalizable learning

Andrew Caplin[1] · **Daniel Martin**[2] · **Philip Marx**[3]

Received: 19 April 2024 / Accepted: 15 November 2024

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. We use our framework to show how identification can be strengthened as one imposes the assumptions behind more restrictive forms of Bayesian learning.

Keywords Rational inattention · Revealed preference · Stochastic choice · Information acquisition · Learning · Identification

JEL classification: D11 · D81 · D83

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 both fixed information models, where learning is not impacted by the decision context, as in many market models of incomplete information, and costly learning models, where distinct methods of learning have different costs and the chosen method maximizes net payoffs, as in sequential search (Caplin et al. 2011) and variable-capacity rational inattention (Matejka and McKay 2015; Caplin et al. 2022).

We thank Leshan Xu for excellent research assistance, and Mark Dean, Haoran Wang, and participants in the Sloan-NOMIS Summer School on Cognitive Foundation of Economic Behavior for helpful comments. We also thank the Alfred P. Sloan Foundation and the NOMIS Foundation for their support for the broader research program on the cognitive foundations of economic behavior.

✉ Daniel Martin
daniel@martinonline.org

[1] Department of Economics, New York University, New York, USA

[2] Department of Economics, University of California, Santa Barbara, USA

[3] Department of Economics, Louisiana State University, Baton Rouge, USA

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 $A^2 = \{a_1, a_2, a_3\}$:

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

Such state-dependent stochastic choice data has long been analyzed in psychology, neuroscience, and economics, especially when studying models of Bayesian learning.[1] In recent theoretical advances, (Caplin and Martin 2015), Caplin and Dean (2015), Chambers et al. (2020), Caplin et al. (2022), Denti (2022), Mensch and Malik (2023) 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.

The central question we address in this paper is: what can the analyst infer from such 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 consistent with rational choice. In other words, the decision maker's learning must be rationalizable.

Caplin and Dean (2015) identify a necessary condition for an information structure to be rationalizable under Bayesian learning, which is that it is a mean preserving spread of the least informative information structure consistent with the data. Their condition is not sufficient because such an information structure can be so informative that utility maximization is not satisfied, which we demonstrate in a simple example introduced in Sect. 1.1. As a result, their condition cannot be used to identify all information structures that are rationalizable under Bayesian learning.

We build on their contribution by providing the conditions that are both necessary and sufficient for Bayesian learning to be rationalizable. We first establish these conditions for 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

---

[1] In psychometric lab experiments, the state could be luminosity, weight, the direction of movement, or other perceptual distinctions — for instance, the proportion of colored balls (Dean and Neligh 2023), the sum of a numeric string (Almog and Daniel 2024), or the shape of a geometric figure (Caplin et al. 2020). Examples from the field include whether a pitch is in or out of the strike zone (Bhattacharya and Howard 2022) or whether a tennis ball lands inor out (Almog et al. 2024). For standard economic settings, the state could be the fundamentals of a stock or the characteristics of a health plan (Brown and Jeon 2024). Rambachan (2024) provides econometric methods for dealing with incomplete state-dependent stochastic choice data and applies them to judicial decisions.

spreads of posterior beliefs that also preserve optimal choice. Thus, it specializes the standard informativeness order to precisely target the optimality embedded in models of Bayesian learning. The value of this specialization in our context is to provide a constructive characterization of all information structures rationalizable within a given decision problem.

For data from multiple decision problems, the necessary and sufficient conditions for rationalizable learning depend on the particular model of Bayesian learning that is imposed, and we consider two nested possibilities: fixed information and costly learning. For costly learning, the more general form of Bayesian learning, we accommodate the additional constraints on rationalizable learning that arise across decision problems by complementing 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 (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).

We also identify what could have been learned under the much more restrictive model of fixed information (Proposition 2) by strengthening the MOPS requirement to generate necessary and sufficient conditions for rationalizable learning. Our result shows how identification of learning is strengthened with assumptions on the form of Bayesian learning because doing so adds requirements on optimality across decision problems.[2]

A related work is by Lu (2016), who identifies the set of possible information structures for fixed information, the more restrictive form of Bayesian learning we consider. We expand on his contribution by also showing how to identify rationalizing information structures under a common generalization of fixed information: costly learning. This expansion is needed when learning is impacted by the decision-making environment, such as when an increase in the payoff to being correct increases how much the decision maker chooses to learn (as suggested by experiments 1.2 and 2.2 of Dean and Neligh (2023). In addition, we complement his contribution by showing what can be inferred about rationalizable learning from finite sets of state-dependent stochastic choice, a data set increasingly used in the theoretical literature on costly learning.[3]

---

[2] It would also be possible identify what could have been learned under capacity-constrained learning, an intermediate case between fixed information and costly learning, by generalizing the test of capacity-constrained introduced in Caplin et al. (2024).

[3] Instead, Lu (2016) shows what can be recovered about information structures under fixed information using "test functions" that are generated over all possible payoffs. Duraj and Lin (2022) perform a similar recovery exercise with more limited observability of agent behavior, and also in the case where the agent can choose whether to acquire a given information structure at a fixed cost.

Our paper also makes a contribution related to the testability of Bayesian learning models. Namely, 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 polynomial-time algorithm of Floyd (1962) and Warshall (1962). This simple and well-known algorithm is often used in revealed preference testing to calculate the transitive closure of a revealed preference relation. Caplin et al. (2022) use our method for calculating the indirect value difference function to test whether the predictions of a machine learning algorithm are consistent with costly learning.

In addition, our paper makes contributions related to the recovery of another key object in learning models: learning costs. Caplin and Dean (2015) provide linear constraints that must be satisfied for learning costs to be consistent with the data, and in Sect. 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).[4] This representative cost is both "central" and easy to calculate, two properties that could be helpful for empirical analysis.[5]

The paper proceeds as follows. In Sect. 1.1 we provide a motivating example which we use to set ideas throughout the paper. In Sect. 2 we formalize the decision problem and introduce the key objects of analysis. In Sect. 3 we introduce MOPS and consider rationalizable learning within a decision problem. In Sect. 4 we introduce the indirect value difference function and consider rationalizable learning with choice data from multiple decision problems. In Sect. 5, we recover the costs of what was learned (Sect. 5.1) and characterize rationalizable learning under the nested model of fixed information (Sect. 5.2).

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:

$$P^1 = \begin{array}{cc} \omega_1 & \omega_2 & \\ \left(\begin{matrix} 0.4 & 0.1 \\ 0.1 & 0.4 \end{matrix}\right. & \left.\right) & \begin{matrix} a_1 \\ a_2 \end{matrix} \end{array}$$

---

[4] Caplin et al. (2022) also use our method for calculating a representative cost function to recover the machine learning algorithm's "pseudo-costs."

[5] This complements the canonical cost function considered by Denti (2022), which is the minimal cost consistent with the choice data under posterior-separability, a specialization of costly learning.

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

$$P^2 = \begin{array}{c} \\ \\ \\ \\ \end{array}\begin{array}{cc} \omega_1 & \omega_2 \\ \end{array}$$

$$\begin{array}{cccc} & \omega_1 & \omega_2 & \\ P^2 = & \left(\begin{array}{c} 0.25 \\ 0.05 \\ 0.2 \end{array}\right. & \left.\begin{array}{c} 0 \\ 0.2 \\ 0.3 \end{array}\right) & \begin{array}{c} a_1 \\ a_2 \\ a_3 \end{array} \end{array}$$

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 $\{x_G, x_M, x_B\}$, which are known to correspond to actions and states as follows:

Action	**State** ω_1	**State** ω_2
a_1	x_G	x_B
a_2	x_B	x_G
a_3	x_M	x_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 was 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 $\bar{Q}^2$ summarized by revealed posteriors $\bar{\gamma}$ of state ω_1 and probabilities of posteriors given by $\bar{Q}(\gamma)$:

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

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:

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

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 a more stringent model of fixed information. The third information structure is consistent even with the model of 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.

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 $X = \{x_k\}_{k=1}^K$. In any given decision problem, a set of actions $A \subset \mathcal{A}$ with $|A| \geq 2$ is available. For each action, the realized prize depends on the state of the world according to a prize specification $x(a, \omega)$ known by the analyst. Because rewards depend on the state, the decision maker (DM) is motivated to learn about the state before choosing actions.

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: $\mu(\omega) = \sum_{a \in A} P(a, \omega)$ and $P(a) \equiv \sum_{\omega \in \Omega} P(a, \omega)$. The SDSC data is equivalently represented by a revealed information structure $\bar{Q}$, defined as the distribution over action-conditional posterior beliefs:

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

$$\bar{Q}(\bar{\gamma}) \equiv \sum_{a:\bar{\gamma}^a = \bar{\gamma}} P(a) \tag{2}$$

for all chosen actions, $P(a) > 0$. When ambiguities exist, our convention in 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 $\Gamma(Q) \equiv \text{supp } Q$, 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 as a function of the posterior, $q(a|\gamma) \in \Delta(A)$. Define $P_{(Q,q)}$ as the hypothetical SDSC that (Q, q) would generate,

$$P_{(Q,q)}(a, \omega) \equiv \sum_{\gamma \in \Gamma(Q)} q(a|\gamma) Q(\gamma) \gamma(\omega). \tag{3}$$

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

$$P_{(Q,q)} = P \tag{4}$$

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 $u : X \to \mathbb{R}$ and a posterior γ as:

$$U(a|\gamma, u) \equiv \sum_{\omega \in \Omega} \gamma(\omega) u(x(a, \omega)) \tag{5}$$

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

$$g(Q, q|u) \equiv \sum_{\gamma \in \Gamma(Q)} \sum_{a \in \mathcal{A}} Q(\gamma) q(a|\gamma) U(a|\gamma, u). \tag{6}$$

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

$$\underset{q:\Gamma(Q) \to \Delta(A)}{\text{argmax}} \quad g(Q, q|u) \tag{7}$$

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:

$$G(Q|A, u) \equiv \max_{q:\Gamma(Q) \to \Delta(A)} g(Q, q|u) \tag{8}$$

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:

$$\underset{Q \in \mathcal{Q}}{\text{argmax}} \; G(Q|A, u) - K(Q) \tag{9}$$

This formulation with chosen information structures also captures fixed information by making the learning cost finite for only one information structure, 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 (7) such that (Q, q) generates the observed data (4). 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 $u : X \to \mathbb{R}$. However, in principle the utility can also be partially recovered from the available data using the (G-NIC) condition.[6] In the case of our running example, this would yield simple bounds, which we henceforth take as given:

$$u(x_B) = 0, \quad u(x_M) \in [0.6, 0.8], \quad u(x_G) = 1 \tag{10}$$

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 transition matrix $B : \Gamma(Q) \times \Gamma(\bar{Q}) \to [0, 1]$, denoted $B(\gamma|\bar{\gamma})$ for target posterior $\gamma \in \Gamma(Q)$ and source posterior $\bar{\gamma} \in \Gamma(\bar{Q})$, satisfying the standard conditions of a mean preserving spread:

$$\sum_{\gamma \in \Gamma(Q)} B(\gamma|\bar{\gamma}) = 1 \tag{11}$$

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

$$\sum_{\gamma \in \Gamma(Q)} \gamma B(\gamma|\bar{\gamma}) = \bar{\gamma} \tag{13}$$

---

[6] In practice, this can be accomplished by extending the geometric approach to recovering utility introduced in Caplin and Martin (2021).

while additionally preserving optimality:

$$B(\gamma|\bar{\gamma}^a) > 0 \implies \gamma \in \hat{\Gamma}(a|A, u) \tag{14}$$

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

As is standard, the first condition (11) requires that for every posterior $\bar{\gamma} \in \Gamma(\bar{Q})$, $B(\cdot|\bar{\gamma})$ is a probability distribution over $\Gamma(Q)$. The second condition (12) 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 (13) requires that the posteriors $\bar{\gamma} \in \Gamma(\bar{Q})$ are an average of posteriors in $\Gamma(Q)$. Finally, condition (14) 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 beyond the revealed information structure generally requires specifying the utility function u, which reduces to picking a scalar $u(x_M) \in [0.6, 0.8]$ with the normalization that $u(x_B) = 0$ and $u(x_G) = 1$. However, the characterization of possible learning in the first decision problem A^1 is independent of this value $u(x_M)$ because the corresponding prize x_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 sets of posteriors at which each action is optimal are $\hat{\Gamma}(a_1|A^1) = [0.5, 1]$ and $\hat{\Gamma}(a_2|A^1) = [0, 0.5]$. 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 transition matrix B permits from revealed posterior $\bar{\gamma}^{a_1} = 0.8$ must preserve optimality of a_1, hence be in the range [0.5, 1]; those permitted from revealed posterior $\bar{\gamma}^{a_2} = 0.2$ must preserve optimality of a_2, hence be in the range [0, 0.5].

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

$$\begin{array}{cc} \gamma & Q^1(\gamma) \\ \left(\begin{array}{c} 1 \\ 1/2 \\ 0 \end{array}\right. & \left.\begin{array}{c} 3/10 \\ 2/5 \\ 3/10 \end{array}\right) \end{array} \tag{15}$$

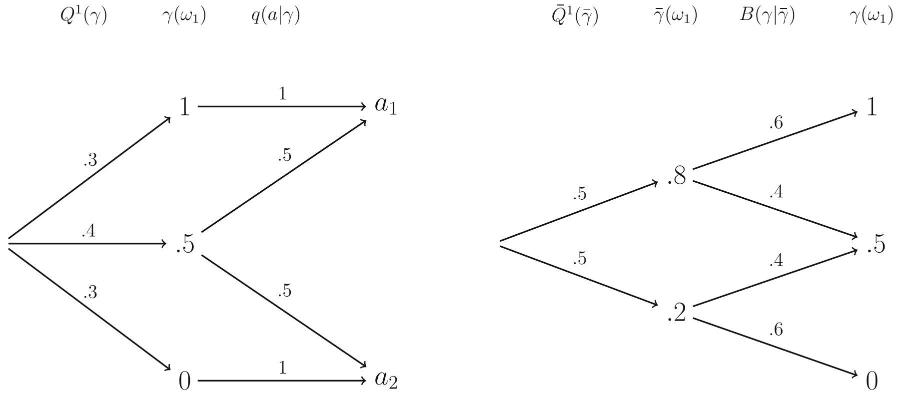

Fig. 1 Rationalizing (left) and MOPS (right) constructions of an information structure Q^1 from the distribution of revealed posteriors $\bar{Q}^1$. Given (inferred) prize utilities $u(x_B) = 0$ and $u(x_G) = 1$, action a_1 is preferred to a_2 for posteriors $\gamma(\omega_1) \geq 0.5$, and action a_2 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 matrix 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 $u(x_M)$, through the sets of posteriors inducing optimal actions:

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

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 Sect. 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:

$$G(Q|A,u) \geq G(\bar{Q}|A,u) \quad \forall u : A \times \Omega \to \mathbb{R}.^{7} \tag{16}$$

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

$$\bar{G}(u) \equiv \sum_{a \in \mathcal{A}} \sum_{\omega \in \Omega} u(x(a,\omega))P(a,\omega) \tag{17}$$

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

$$G(Q|A,u) = \bar{G}(u) \tag{18}$$

Intuitively, condition (18) results from the combination of two binding inequalities. First is that the maximal gross utility at the revealed information structure $G(\bar{Q}|A,u)$ is at least as high as the revealed gross utility $\bar{G}(u)$ 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 $G(Q|A,u)$ is at least as high as what is revealed $G(\bar{Q}|A,u)$ 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 Sect. 4 when we consider learning inferred from richer data with 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 additional constraints.

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 $\bar{\boldsymbol{Q}} \equiv (\bar{Q}^1, \ldots, \bar{Q}^M)$. 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 further constraints on what could have been learned.

We begin by characterizing the 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 (7) and generating the observed data sets as in (4), 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 (9). In this case we say that the action strategies $(\mathbf{Q}, \mathbf{q})$ rationalize the observed data $\mathbf{P}$ according to costly learning, and that $\mathbf{Q}$ is a viable tuple of what could have been learned.

4.1 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 (15). As shown in Fig. 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 (9) 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 (17), and the utility bounds (10), we can compute the gross utility for any learned information structures Q^1 and Q^2 in respective decision problems 1 and 2 as:

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

Given a specification (15) for information structure Q^1, we can also compute its gross value in choice set A^2 as $G(Q^1|A^2, u) = 0.6 + 0.4u(x_M)$. 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 $G(Q^1|A^1, u) + G(Q^2|A^2, u) = 1.25 + 0.5u(x_M)$. Analogously, the sum of counterfactual gross utilities from switching learning across decision problems is $G(Q^2|A^1, u) + G(Q^1|A^2, u) \geq 1.35 + 0.4u(x_M)$. 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 $u(x_M) \in [0.6, 0.8]$. [8]

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

---

[8] Specifically, as long as $u(x_M) < 1$.

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:

$$D_0^m(Q|u) \equiv G(Q|A^m, u) - \bar{G}^m(u), \tag{19}$$

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 is 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 where the first and last entries coincide: $h^{J(\vec{h})+1} = h^1$. For 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 $H(m,n)$ of attention paths for $1 \leq m, n \leq M$ yields our main object for 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:

$$D^{mn}(\mathbf{Q}|u) \equiv \max_{\vec{h} \in H(m,n)} \sum_{j=1}^{J(\vec{h})} D_0^{h^j}(Q^{h^{j+1}} \,|\, u) \tag{20}$$

Intuitively, the indirect value difference function sums the changes in maximized expected utility along a path in which each decision problem A^{h^j} is shifted to attention strategy $Q^{h^{j+1}}$ of 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 tuple of information structures to have been learned in the 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 on the diagonal entries of the indirect value difference matrix that we refer to as Generalized No Improving Cycles (G-NIC):

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

where 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, the (G-NIC) 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* $(\mathbf{A}, \mathbf{P})$ *generating revealed information structures* $\bar{\mathbf{Q}}$. *Given utility function u, the following are equivalent:*

1. *There exists a learning cost function* K *and 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* $\bar{\mathbf{Q}}$.

A feature of the value difference characterization of Theorem 2 is that it circumvents the need to specify a rationalizing action strategy $\mathbf{q}$. Still, the proof of rationalization requires such a construction, which follows by Lemma 1. We show in the following Sect. 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 two complementary ways. First, in

Sect. 5.1, we derive the identified set of all information cost functions that rationalize such learning. Second, in Sect. 5.2, we further elucidate the nature of learning by considering consistency with an important nested class of the costly learning model, namely fixed information. In each case, we emphasize the importance of our value difference and MOPS constructions.

5.1 Costs of learning

We begin with recovery of learning costs as a function of a utility function u and 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 identified set of rationalizing learning cost functions consists of cost functions that rationalize what was learned:*

$$K(Q^n) - K(Q^m) \geq D^{mn}(\mathbf{Q}|u) \tag{21}$$

and rationalize what was not learned:

$$K(Q) \geq \max_{1 \leq m \leq M} [D_0^m(Q|u) + K(Q^m)]. \tag{22}$$

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

The first condition (21) characterizes the learning cost function evaluated at what was learned. In particular, this condition is necessary and sufficient for optimal learning in each decision problem, relative to learning in other decision problems. The second condition (22) 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 $\bar{\boldsymbol{Q}}$. Note, however, that the bounds corresponding to the revealed information structures $\bar{\boldsymbol{Q}}$ 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): $D(\mathbf{Q}|u) \geq D(\bar{\boldsymbol{Q}}|u)$. 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 Sect. 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 on the domain of what was learned. 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 *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:*

$$\{\tilde{K} \in \mathcal{K}^M(\mathbf{Q}, u) : \tilde{K}^m = 0\} \tag{23}$$

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

$$\frac{1}{2M}\left[\sum_{m=1}^{M} D^{m*}(\mathbf{Q}|u) - \sum_{m=1}^{M} D^{*m}(\mathbf{Q}|u)\right] \tag{24}$$

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 define a single cost function on the domain of what was learned from the indirect value difference function; by its nature, we call this cost function "representative" in our context. Of course, ours is not the only possible definition of a representative or otherwise canonical cost function, and the literature on costly learning contains alternative suggestions. Notably, Denti (2022) (following Rockafellar 1973) considers the minimal monotone cost function that rationalizes the data and (partially) identifies it 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. Following Theorem 3, we begin by constructing the indirect value difference matrix $D(\bar{\mathbf{Q}}|u)$. This depends on the direct value differences (19), 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:

$$\begin{aligned} G(\bar{Q}^1|A^1, u) &= \bar{G}^1(u) = 0.8 \\ G(\bar{Q}^2|A^2, u) &= \bar{G}^2(u) = 0.45 + 0.5u(x_M) \end{aligned} \tag{25}$$

which is consistent with condition (18) of Lemma 1 for rationalizability within decision problem. Additionally, the gross utilities from switching revealed information structures across decision problems can be computed as:

$$\begin{aligned} G(\bar{Q}^2|A^1, u) &= 0.75 \\ G(\bar{Q}^1|A^2, u) &= 0.8 \end{aligned} \tag{26}$$

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

$$D(\bar{\boldsymbol{Q}}|u) = \begin{pmatrix} 0 & -0.05 \\ 0.35 - 0.5u(x_M) & 0 \end{pmatrix} \tag{27}$$

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

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

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 $0.35 - 0.5u(x_M) - 0.05$ is less than zero when $u(x_M) \geq 0.6$.

Condition (21) of Theorem 3 then yields the bounds:

$$K(\bar{Q}^1) - K(\bar{Q}^2) \in [0.35 - 0.5u(x_M), 0.05] \tag{28}$$

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:

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

The set of cost differences (28) is non-empty for all $u(x_M) \in [0.6, 0.8]$. When $u(x_M) = 0.8$, there is a wide range of rationalizing cost differences, $K(\bar{Q}^1) - K(\bar{Q}^2) \in [-0.05, 0.05]$. When $u(x_M) = 0.6$ 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:

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

and sign-inverted column:

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

is a constrained extremal cost function restricted to revealed learning, and their average (24) yields a representative such cost of $(0.1 - 0.125u(x_M), 0.125u(x_M) - 0.1)$. Notably, the representative cost places higher cost on $\bar{Q}^1$ than $\bar{Q}^2$ for almost all feasible $u(x_M) \in [0.6, 0.8)$, even though this need not be true for a cost function to rationalize the data.

5.2 Fixed information

Complementing cost recovery, we may also be interested in what is implied 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 an additional important class of learning model, in which learning is fixed and exogenous to the decision problem.

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

1. *There exist action strategies* $\mathbf{q}$ *that rationalize data* $\mathbf{P}$ *with fixed information Q.*
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.[9] 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, so that $Q(1) = 0.25$. This follows from two observations in decision problem 2. First, $Q(1) \geq 0.25$ since

---

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

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 the upper posterior cutoff $1 - u(x_M) \leq 0.4$ 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

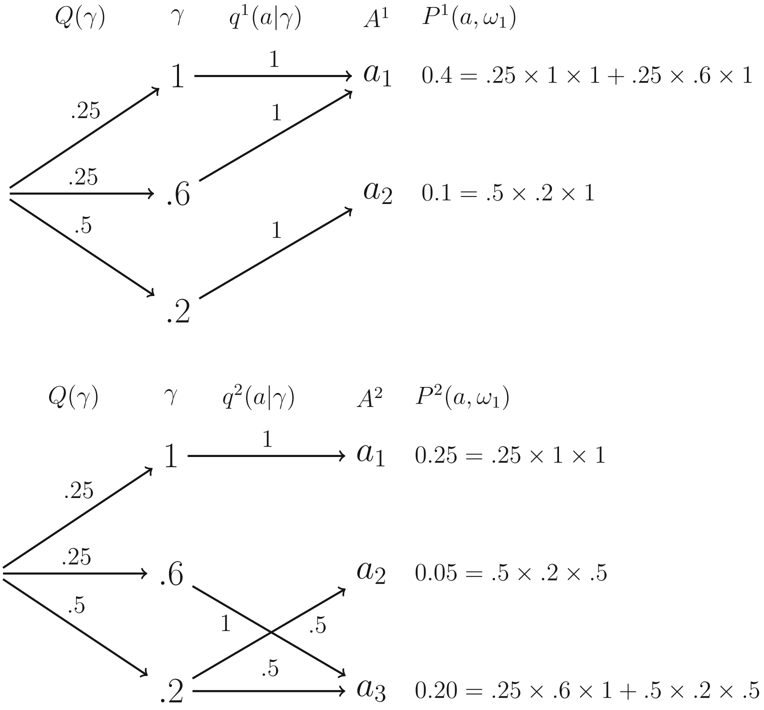

Fig. 2 Existence of a common distribution of posteriors and mixed action strategies q^1, q^2 that rationalize the data sets P^1, P^2

in Fig. 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 costly learning 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 $u(x_M) = 0.8$. 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.

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:*

$$G(Q|A, u) = \sum_{\gamma \in \Gamma(Q)} Q(\gamma) \left[\max_{a \in A} U(a|\gamma, u) \right] \tag{29}$$

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

$$G(\bar{Q}|A, u) \geq \bar{G}(u) \tag{30}$$

Proof *(Proof of Lemma 2)* By definitions (6) and (8),

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

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 (29). For (30),

$$\begin{aligned}
\bar{G}(u) &= \sum_{a \in \mathcal{A}} \sum_{\omega \in \Omega} P(a, \omega) u(x(a, \omega)) && \text{by (17)} \\
&= \sum_{a \in \mathcal{A}} \sum_{\omega \in \Omega} P(a) \bar{\gamma}^a(\omega) u(x(a, \omega)) && \text{by (1)} \\
&= \sum_{a \in \mathcal{A}} P(a) \sum_{\omega \in \Omega} \bar{\gamma}^a(\omega) u(x(a, \omega)) && \text{by rearrangement} \\
&= \sum_{a \in \mathcal{A}} P(a) U(a|\bar{\gamma}^a, u) && \text{by (5)} \\
&\leq \sum_{a \in \mathcal{A}} P(a) \left[\max_{b \in A} U(b|\bar{\gamma}^a, u) \right] && \text{by optimization} \\
&= \sum_{\bar{\gamma} \in \Gamma(\bar{Q})} \bar{Q}(\bar{\gamma}) \left[\max_{b \in A} U(b|\bar{\gamma}, u) \right] && \text{by (2)} \\
&= G(\bar{Q}|A, u). && \text{by (29).}
\end{aligned}$$

□

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 (18). Since Q is assumed as informative as $\bar{Q}$, Blackwell's Theorem (Blackwell 1953, Theorems 2 and 6) implies the existence of a transition matrix $B : |\Gamma(Q)| \times |\Gamma(\bar{Q})| \to [0, 1]$ defining Q as a mean preserving spread of $\bar{Q}$. For the sake of contradiction, suppose that this B does not preserve the optimality (14) 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:

$$P(a) > 0, \quad B(\gamma|\bar{\gamma}^a) > 0, \quad \gamma \notin \hat{\Gamma}(a|A, u) \tag{31}$$

This contradicts the assumption that Q yields maximal gross utility equal to what is revealed (18) because:

$$\begin{aligned}
\bar{G}(u) &= \sum_{a\in A}\sum_{\omega\in\Omega} u(x(a,\omega))P(a,\omega) && \text{by (17)}\\
&= \sum_{a\in A}\sum_{\omega\in\Omega} u(x(a,\omega))P(a)\bar{\gamma}^a(\omega) && \text{by (1)}\\
&= \sum_{a\in A}\sum_{\omega\in\Omega} u(x(a,\omega))P(a)\sum_{\gamma\in\Gamma(Q)} B(\gamma|\bar{\gamma}^a)\gamma(\omega) && \text{by (13)}\\
&= \sum_{a\in A}\sum_{\gamma\in\Gamma(Q)} P(a)B(\gamma|\bar{\gamma}^a)\sum_{\omega\in\Omega}\gamma(\omega)u(x(a,\omega)) && \text{by rearranging}\\
&= \sum_{a\in A}\sum_{\gamma\in\Gamma(Q)} P(a)B(\gamma|\bar{\gamma}^a)U(a|\gamma,u) && \text{by (5)}\\
&< \sum_{a\in A}\sum_{\gamma\in\Gamma(Q)} P(a)B(\gamma|\bar{\gamma}^a)\left[\max_{b\in A} U(b|\gamma,u)\right] && \text{by (31)}\\
&= \sum_{\gamma\in\Gamma(Q)}\left[\max_{b\in A} U(b|\gamma,u)\right]\sum_{a\in A} P(a)B(\gamma|\bar{\gamma}^a) && \text{by rearranging}\\
&= \sum_{\gamma\in\Gamma(Q)}\left[\max_{b\in A} U(b|\gamma,u)\right]\sum_{\bar{\gamma}\in\Gamma(\bar{Q})} \bar{Q}(\bar{\gamma})B(\gamma|\bar{\gamma}) && \text{by (2)}\\
&= \sum_{\gamma\in\Gamma(Q)}\left[\max_{b\in A} U(b|\gamma,u)\right] Q(\gamma) && \text{by (12)}\\
&= G(Q|A,u) && \text{by (8)}
\end{aligned}$$

($2 \implies 1$) Suppose there exists a transition matrix $B : |\Gamma(Q)| \times |\Gamma(\bar{Q})| \to [0, 1]$ defining Q as a mean and optimality preserving spread (MOPS) of $\bar{Q}$. We use B to construct a mixed strategy $q : \Gamma(Q) \to \Delta(A)$ such that (Q, q) generates the data (4) in an optimal way (7). 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:

$$q(a|\gamma) = \frac{P(a)B(\gamma|\bar{\gamma}^a)}{Q(\gamma)} \tag{32}$$

Note first that q as defined in (32) are mixed strategies by construction since they are nonnegative and summing the numerators across actions yields the denominator:

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

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

$$\begin{aligned}
P_{(Q,q)}(a, \omega) &= \sum_{\gamma\in\Gamma(Q)} q(a|\gamma)Q(\gamma)\gamma(\omega) && \text{by (3)}\\
&= \sum_{\gamma\in\Gamma(Q)} P(a)B(\gamma|\bar{\gamma}^a)\gamma(\omega) && \text{by (32)}\\
&= P(a) \sum_{\gamma\in\Gamma(Q)} B(\gamma|\bar{\gamma}^a)\gamma(\omega) && \text{by rearranging}\\
&= P(a)\bar{\gamma}^a(\omega) && \text{by (13)}\\
&= P(a, \omega). && \text{by (1)}
\end{aligned}$$

Finally, we show that the mixed strategy q identified above only chooses actions at posteriors where they are optimal, as in (7). By construction (32), $q(a|\gamma) > 0$ implies $B(\gamma|\bar{\gamma}^a) > 0$. By the defining property (14) of a mean *and optimality* preserving spread,

$$q(a|\gamma) > 0 \implies \gamma \in \hat{\Gamma}(a|A, u). \tag{33}$$

In that case, we have:

$$\begin{aligned}
g(q, Q) &= \sum_{\gamma\in\Gamma(Q)} \sum_{a\in\mathcal{A}} Q(\gamma)q(a|\gamma)U(a|\gamma, u) && \text{by (6)}\\
&= \sum_{\gamma\in\Gamma(Q)} \sum_{a\in\mathcal{A}} Q(\gamma)q(a|\gamma)\left[\max_{b\in A} U(b|\gamma, u)\right] && \text{by (33)}
\end{aligned}$$

$$\begin{aligned}
&= \sum_{\gamma \in \Gamma(Q)} Q(\gamma) \left[\max_{b \in A} U(b|\gamma, u) \right] \sum_{a \in \mathcal{A}} q(a|\gamma) && \text{by rearrangement} \\
&= \sum_{\gamma \in \Gamma(Q)} Q(\gamma) \left[\max_{b \in A} U(b|\gamma, u) \right] && \text{because } q(\cdot|\gamma) \in \Delta(A) \\
&= G(Q|A, u) && \text{by (29)}
\end{aligned}$$

which implies by definition of the gross value of learning function (8) that q is optimal for choice set A given u.

(1 $\implies$ 3) Suppose there exists an action strategy q such that (Q, q) rationalizes the data according to EU maximization (7). First, observe from rationalization that for all $\bar{\gamma} \in \Gamma(\bar{Q})$ and $\omega \in \Omega$ we have:

$$\begin{aligned}
\bar{Q}(\bar{\gamma})\bar{\gamma}(\omega) &= \sum_{a:\bar{\gamma}^a=\bar{\gamma}} P(a)\bar{\gamma}(\omega) && \text{by (2)} \\
&= \sum_{a:\bar{\gamma}^a=\bar{\gamma}} P(a, \omega) && \text{by (1)} \\
&= \sum_{a:\bar{\gamma}^a=\bar{\gamma}} P_{(Q,q)}(a, \omega) && \text{by (4)} \\
&= \sum_{a:\bar{\gamma}^a=\bar{\gamma}} \sum_{\gamma \in \Gamma(Q)} q(a|\gamma)Q(\gamma)\gamma(\omega) && \text{by (3)} \\
&= \sum_{\gamma \in \Gamma(Q)} \left[\sum_{a:\bar{\gamma}^a=\bar{\gamma}} q(a|\gamma) \right] Q(\gamma)\gamma(\omega) && \text{by rearranging}
\end{aligned}$$

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, Theorems 3 and 5), this implies that information structure Q is as informative as $\bar{Q}$. It remains to show that Q yields maximal gross utility equal to what is revealed (18):

$$\begin{aligned}
G(Q|A, u) &\geq G(\bar{Q}|A, u) && \text{by informativeness} \\
&\geq \bar{G}(u) && \text{by (30)} \\
&= \sum_{a \in A} \sum_{\omega \in \Omega} u(x(a, \omega))P(a, \omega) && \text{by (17)} \\
&= \sum_{a \in A} \sum_{\omega \in \Omega} u(x(a, \omega))P_{(Q,q)}(a, \omega) && \text{by (4)}
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{a\in A}\sum_{\omega\in\Omega} u(x(a,\omega)) \sum_{\gamma\in\Gamma(Q)} q(a|\gamma)Q(\gamma)\gamma(\omega) && \text{by (3)}\\
&= \sum_{\gamma\in\Gamma(Q)} Q(\gamma)\sum_{a\in A} q(a|\gamma)\sum_{\omega\in\Omega}\gamma(\omega)u(x(a,\omega)) && \text{by rearranging}\\
&= \sum_{\gamma\in\Gamma(Q)} Q(\gamma)\sum_{a\in A} q(a|\gamma)U(a|\gamma,u) && \text{by (5)}\\
&= g(q,Q|u) && \text{by (6)}\\
&= G(Q|A,u)
\end{aligned}
$$

where the last step follows from the starting assumption that q was optimal (7) for choice set A given utility u. □

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 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 previous use 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:*

$$G(Q^m|A^m,u) - \tilde{K}^m \geq G(Q^n|A^m,u) - \tilde{K}^n \quad \forall\, 1 \leq m,n \leq M \tag{34}$$

if and only if:

$$\max_{\vec{h}\in H(m,m)} \sum_{j=1}^{J(\vec{h})}[G(Q^{h^{j+1}}|A^{h^j},u) - G(Q^{h^j}|A^{h^j},u)] = 0 \quad \forall\, 1 \leq m \leq M \tag{35}$$

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 (34). Rearrangement yields:

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

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 (35).

Conversely, assume (35). Re-ordering sums for case $m = 1$ implies:

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

or, upon reversing cycles,

$$\max_{\vec{h} \in H(1,1)} \sum_{j=1}^{J(\vec{h})} [G(Q^{h^j}|A^{h^{j+1}}, u) - G(Q^{h^j}|A^{h^j}, u)] = 0. \tag{36}$$

Define as a function of the index $1 \leq m \leq M$,

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

Note that the maximum exists in our case since there are only finitely many paths in $H(1, m)$. For any $1 \leq m, n \leq M$, the construction and (36) imply:

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

Defining $\tilde{K}^m \equiv G(Q^m|A^m, u) - V(m)$ and substituting yields (34). □

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 (18):

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

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 (35). Plugging revealed utility from (18) into this condition and substituting for direct and direct value difference functions (19) and (20) yields the (G-NIC) condition because:

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

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

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 (30) 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 (7), by Lemma 1. Additionally, plugging this equality into (G-NIC) yields cyclic monotonicity (35). In turn, by Lemma 3 this implies the existence of a $\tilde{K} \in \mathbb{R}^M$ satisfying (34). 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:

$$
G(Q^m|A^m, u) - K(Q^m) \geq G(Q|A^m, u) - K(Q) \quad \forall\, 1 \leq m \leq M,\, Q \in \mathcal{Q} \tag{37}
$$

Upon substituting for revealed gross utility by rationalizability within decision problem (Lemma 1), rearranging, and substituting again for direct value difference (19), 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 (21). 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 (22).

Conversely, suppose that the information tuple $\mathbf{Q}$ is rationalizable by some cost function, and that the cost function K satisfies (21) and (22). That the learning cost function K rationalizes learning across decision problems follows from reversing the preceding arguments to conclude (37). $\square$

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 (21) can be extended to a cost function on $\mathcal{Q}$ that also satisfies constraints (22) 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:

$$\tilde{K}^n \geq \tilde{K}^1 + D^{1n}(\mathbf{Q}|u)$$
$$\tilde{K}^n \leq \tilde{K}^1 - D^{n1}(\mathbf{Q}|u)$$

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

$$D^{1n}(\mathbf{Q}|u) \leq \tilde{K}^n \leq -D^{n1}(\mathbf{Q}|u)$$

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 (21) 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 (20) 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 (24) is also an element of the set. □

***Proof of Proposition* 2** This result is immediate by construction upon applying the within-problem characterization (Theorem 1) across decision problem. Distinct from the costly learning model, 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 costly learning with a single information structure of finite cost, so that the models are nested. □

It is useful to 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 (19) of the direct value difference function and definition (16) 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 element-wise by definition (20) of the indirect value difference function:

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

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

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.

References

Afriat, S.N.: The construction of utility functions from expenditure data. Int. Econ. Rev. **8**(1), 67–77 (1967)

Almog, D., Gauriot, R., Page, L., Martin, D.: AI oversight and human mistakes: evidence from centre court. In: arXiv preprint arXiv:2401.16754 (2024)

Almog, D., Martin, D.: Rational inattention in games: experimental evidence. Exp. Econ. 1–28 (2024)

Bhattacharya, V., Howard, G.: Rational inattention in the infield. Am. Econ. J. Microecon. **14**(4), 348–393 (2022)

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

Blackwell, D. (1953). Equivalent comparisons of experiments. Annals Math. Stat. 265–272

Brown, Z.Y., Jeon, J.: Endogenous information and simplifying insurance choice. Econometrica **92**(3), 881–911 (2024)

Caplin, A., Csaba, D., Leahy, J., Nov, O.: Rational inattention, competitive supply, and psychometrics. Quarter. J. Econ. **135**(3), 1681–1724 (2020)

Caplin, A., Dean, M.: Revealed preference, rational inattention, and costly information acquisition. Am. Econ. Rev. **105**(7), 2183–2203 (2015)

Caplin, A., Dean, M., Leahy, J.: Rationally inattentive behavior: characterizing and generalizing Shannon entropy. J. Polit. Econ. **130**(6), 1676–1715 (2022)

Caplin, A., Dean, M., Martin, D.: Search and satisficing. Am. Econ. Rev. **101**(7), 2899–2922 (2011)

Caplin, A., Martin, D.: A testable theory of imperfect perception. Econ. J. **125**(582), 184–202 (2015)

Caplin, A., Martin, D.: Comparison of decisions under unknown experiments. J. Polit. Econ. **129**(11), 3185–3205 (2021)

Caplin, A., Martin, D., Marx, P.: Modeling Machine Learning. Tech. rep, National Bureau of Economic Research (2022)

Caplin, A., Martin, D., Marx, P., Morozova, A., Xu, L.: Testing capacity-constrained learning. Working paper (2024)

Chambers, C.P., Liu, C., Rehbeck, J.: Costly information acquisition. J. Econ. Theory **186**, 104979 (2020)

Dean, M., Neligh, N.: Experimental tests of rational inattention. J. Polit. Econ. **131**(12), 3415–3461 (2023)

Denti, T.: Posterior separable cost of information. Am. Econ. Rev. **112**(10), 3215–3259 (2022)

Duraj, Jetlir, Lin, Yi-Hsuan.: Costly information and random choice. Econ. Theor. **742**, 135–159 (2022)

Floyd, R.W.: Algorithm 97: shortest path. Commun. ACM **5**(6), 345 (1962)

Kamenica, E., Gentzkow, M.: Bayesian persuasion. Am. Econ. Rev. **101**(6), 2590–2615 (2011)

Koopmans, T.C., Beckmann, M.: Assignment problems and the location of economic activities. Econometrica **25**(1), 53–76 (1957)

Lipnowski, E., Ravid, D.: Predicting choice from information costs. In: arXiv preprint arXiv:2205.10434 (2022)

Lu, J.: Random choice and private information. Econometrica **84**(6), 1983–2027 (2016)

Matejka, F., McKay, A.: Rational inattention to discrete choices: a new foundation for the multinomial logit model. Am. Econ. Rev. **105**(1), 272–98 (2015)

Mensch, J., Malik, K.: Posterior-mean separable costs of information acquisition. In: arXiv preprint arXiv:2311.09496 (2023)

Rambachan, A.: Identifying prediction mistakes in observational data. Quarter J. Econ. **139**(3), 1665–1711 (2024)

Rochet, Jean-Charles.: A Necessary and sufficient condition for rationalizability in a quasi-linear context. J. Math. Econ. **16**, 191–200 (1987)

Rockafellar, R Tyrrell: Convex Analysis. Princeton University Press, Princeton (1970)

Rockafellar, R.T.: The multiplier method of Hestenes and Powell applied to convex programming. J. Optim. Theory Appl. **12**(6), 555–562 (1973)

Varian, H.R.: The nonparametric approach to demand analysis. Econometrica: J. Econ. Soc. pp. 945–973 (1982)

Warshall, S.: A theorem on Boolean matrices. J. ACM (JACM) **9**(1), 11–12 (1962)