Rationalizing dynamic choices*

Henrique de Oliveira Rohit Lamba

June 2022

Abstract

An analyst observes an agent take a sequence of actions. The analyst does not have access to the agent's information and ponders whether the observed actions could be justified through a rational, Bayesian model. We show that the observed actions cannot be justified if and only if there is a single deviation argument that leaves the agent better off, regardless of the information. Three applications are presented: a test of the Bayesian model, partial identification of preferences without assumptions on information, and a result showing that, with more risk aversion, more actions can be rationalized.

1 Introduction

As information arrives over time, people may take actions that seemingly go against their own past choices. How can we judge someone's sequence of choices without knowing what they knew? A permissive criterion would allow for *any* sequence of choices that can be explained by the piecemeal arrival of *some* information. The purpose of this paper is to characterize, for a general decision problem, the sequences of choices which can be rationalized by such criteria.

We consider the following model: There is a set of states of the world Ω. The agent starts with a prior $p \in \Delta(\Omega)$, sees a signal s_1 that provides her some information about the actual state of the world, and then chooses an action a_1. The agent then sees another informative signal s_2, chooses an action a_2 and so on, until the final action a_T is chosen. A terminal payoff is realized, represented by an arbitrary function $u : A \times \Omega \to \mathbb{R}$, where A is the set of all action sequences.

Our aim is to characterize the empirical content of this model. To that end, we take the perspective of an outside analyst who knows the agent's utility function u, but does not know

*de Oliveira: Fundação Getúlio Vargas-EESP, henrique.deoliveira@fgv.br; Lamba: Penn State University, rlamba@psu.edu. We are grateful to Ashwin Kambhampati for excellent research assistance. We are indebted to Dilip Abreu, Nageeb Ali, Pierpaolo Battigalli, Benjamin Brooks, Andrew Caplin, Sylvain Chassang, Laura Doval, Daniel Grodzicki, Vijay Krishna, Jay Lu, Stephen Morris, David Pearce, Luciano Pomatto, Doron Ravid, Larry Samuelson, Ran Shorrer, Ron Siegel, Marciano Siniscalchi, Alexander Wolitzky, and seminar participants at University of Chicago, Yale University, University of Rochester, Indian Statistical Institute Delhi, University of Pennsylvania, New York University, University of British Columbia, Princeton University, Pennsylvania Economic Theory Conference and European Summer Symposium in Economic Theory at Gerzensee for their comments.

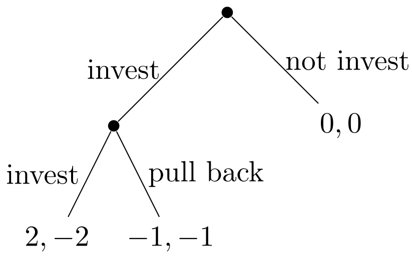
the agent's prior $p \in \Delta(\Omega)$ nor her information process $\pi : \Omega \to \Delta(S)$, where S is the set of all signal sequences. Upon observing some data about the agent's choices, the analyst asks: could this data be generated by optimal Bayesian behavior, for some p and π?

The definition of empirical content depends on what data the analyst can observe. We start with the parsimonious assumption that the analyst observes a single action sequence $(a_1, \ldots, a_T)$. Such an action sequence can then be rationalized by an analyst if there is a p and π and some optimal strategy that chooses the action sequence with positive probability.

To understand the setting and what this definition allows, consider the following simple example:

Example 1. A CEO faces an opportunity to invest in a project with uncertain payoffs: there is a return of 4 if the project meets favorable conditions in the future (good state) and 0 if not (bad state). The project bears fruits on two rounds of investment, and each round of investment costs 1 unit. The CEO has three options: not invest, invest in the first round and pull back in the second, or investment in both periods. The CEO's payoff matrix and decision tree can be summarized as follows:

	not invest	invest & pull back	invest & invest
good	0	-1	2
bad	0	-1	-2



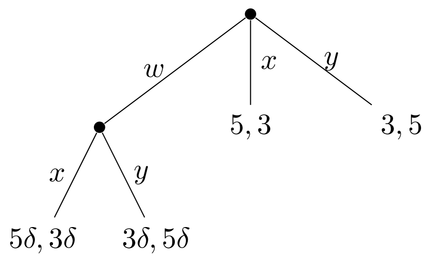
Suppose that we learn that the CEO invested in the first round, incurring the initial cost, but then pulled back. Some might interpret that as evidence of incompetence, saying that in no state can this sequence of actions be justified. They might say that even if the CEO was not sure about the state of the world, not investing would surely have been a better choice. These critics would be ignoring a simple explanation: it might be that the CEO initially received good news about the investment, but after the first round of investment learnt that the project was likely to fail.

In Example 1, the action sequence (*invest*, *pull back*) is what we will call *apparently dominated*—there exists another sequence of actions, (*not invest*, $\emptyset$), under which the agent does strictly better in every state of the world.[1] It will be easy to show that any action sequence that cannot be rationalized is apparently dominated. However, as Example 1 shows, the converse is not true. In fact, in Example 1, all three possible sequences of choices can be rationalized, which illustrates how permissive this first criterion is. But it is not vacuous and can exclude some dynamic choices. For instance, consider the following example:

Example 2. A firm can bet on one of two technologies, X or Y. The firm can also postpone the decision, but by doing so its payoff is discounted by a factor δ, where $0 < \delta < 1$. The payoff matrix and decision tree are as follows:

[1] Generally, an action sequence is apparently dominated if there exists another action sequence (or a lottery over action sequences) that does strictly better in every state of the world.

	x	y	wx	wy
X	5	3	5δ	3δ
Y	3	5	3δ	5δ



Note that both wx and wy are apparently dominated, which does not necessarily rule them out.[2] We learn that the firm has decided to wait instead of making an immediate bet. Under what values of δ can this choice be rationalized? By waiting, the firm can get at most 5δ. By making an immediate decision, the firm is guaranteed to get at least 3. Hence, if $\delta < 3/5$, waiting cannot be rationalized.

But this is not the full story. If the firm makes an immediate decision to randomize equally between x and y, it is guaranteed an expected payoff of 4, no matter the state. Therefore, waiting cannot be rationalized when $\delta < 4/5$. On the other hand, if $\delta \geqslant 4/5$, waiting can be explained by the following information: it could be that the firm starts with an even prior and then fully learns the state of the world in the second period. Thus, waiting can be rationalized precisely when $\delta \geqslant 4/5$.

More generally, an action sequence can be rationalized when there exists a prior p and an information structure π for which an optimizing agent could end up choosing that action sequence with positive probability. Thus, to argue that an action sequence can be rationalized, it is enough to provide a single information structure and prior that prove it to be so; to argue that an action sequence cannot be rationalized, it must be shown that every information structure and prior would fail to rationalize it. In Example 2, we found a single deviation that simultaneously showed that every information structure would fail to rationalize waiting, thereby avoiding direct consideration of the set of all information structures.

The challenge now is: for any arbitrary set of states, actions and utility function, in order to show that an action sequence cannot be rationalized, can we generalize the deviation argument? The construction of this argument through a *deviation rule* forms the core of our paper.

Formally, a deviation rule is an *adapted* mapping from action sequences to lotteries over action sequences, $D : A \to \Delta(A)$. Adaptedness simply requires that deviations today can only be a function of past actions and past deviations, and not of future actions or deviations. In Example 1, if we map (*invest*, *pull back*) to (*not invest*, $\emptyset$), then adaptedness demands that we have to map (*invest*, *invest*) also to (*not invest*, $\emptyset$). As a result, there exists an action sequence and a state of the world for which the deviation makes the agent worse off. In Example 2, the (perhaps intuitively appealing) mapping $wx \mapsto x$, $wy \mapsto y$, $x \mapsto x$ and $y \mapsto y$ is not adapted, and hence not a valid deviation rule. However, the mapping $wx \mapsto \frac{1}{2}x + \frac{1}{2}y$, $wy \mapsto \frac{1}{2}x + \frac{1}{2}y$, $x \mapsto y$ and $y \mapsto y$ is adapted, and is eventually used to show that action sequences wx and

[2]We are using the shorthand wx for (w, x) and wy for (w, y).

wy cannot be rationalized if $\delta < 4/5$.

The concept of deviation rules is bereft of information since it must respect the constraint that the analyst may not know anything about the agent's sequential information structure. For any strategy of the agent $\sigma : S \to \Delta(A)$, the composition mapping $D \circ \sigma = \sigma'$ is a new strategy that tells the agent the following: upon observing a sequence of signals $s = (s_1, ..., s_T)$, if σ had generated a distribution $\sigma(s) \in \Delta(A)$, now instead generate the distribution $\sigma'(s) \in \Delta(A)$. Since σ and D are both adapted, the new strategy σ' is also adapted and hence well defined.

We say that a deviation rule improves upon an action sequence if for every state of the world it strictly increases payoffs along that action sequence without worsening payoffs elsewhere on the decision tree. We then say that the action sequence is *truly dominated* by this proposed deviation rule. In Example 1, the action sequence (*invest*, *pull back*) cannot be improved upon without worsening payoffs along the rest of the decision tree and hence is not truly dominated. On the other hand, in Example 2 the action sequences wx and wy can be improved upon by the deviation rule described above without tinkering with payoffs elsewhere and hence are truly dominated. Our main result, Theorem 1, establishes the following equivalence: An action sequence cannot be rationalized if and only if it is truly dominated.

The theorem can be viewed as a form of duality—it replaces the "for all" quantifier with the "there exists" quantifier and vice-versa. In order to show that an action sequence can be rationalized, the analyst can construct *one* information structure for which the action sequence receives positive weight under an optimal strategy. In order to show that an action sequence cannot be rationalized, the analyst can now construct *one* deviation rule that dominates it.[3]

By delineating the sequences of actions that cannot be rationalized, Theorem 1 settles the question of empirical content when a single action sequence is observed. This is a minimal data requirement that allows us to make predictions even for an individual agent. But this minimality means that the theory may not be rejected. In some cases, such as Example 1, the theory cannot be rejected by any observation. It is then natural to look for finer predictions coming from richer datasets.

Theorems 2 and 3 concern data in the form of entire distributions, such as what can be obtained from a large sample of identical agents with independent information. In this context, Theorem 1 pins down the support of any distribution of action sequences that an analyst could observe in the population. Now, observing the choices of an entire population, can the analyst go further and ask which distributions can be rationalized?

More concretely, in Example 1, we argued that the action sequence (*invest*, *pull back*) can be rationalized. It is easy to see that it cannot be rationalized with probability one, that is,

[3] Note that if $T = 1$, the set of actions which can be rationalized are precisely those that are a best-response to some belief over states. The theorem then reduces to the celebrated Wald-Pearce Lemma (Wald [1949] and Pearce [1984]), which states that the actions which are never a best-response, and hence cannot be rationalized, are strictly dominated by some mixed strategy. Here, our rule would recommend deviating from the dominated action to the dominating mixed strategy, and keep the identity mapping elesewehere. Adaptedness of course has no bite in the static model.

no sequential information structure can induce a rational firm to take this action sequence for sure; it might as well choose (*not invest*, $\emptyset$) no matter the signals it observed. The population interpretation is that there must be an upper bound on the fraction of firms that choose (*invest*, *pull back*) in the data set. Similarly, in Example 2, what is the upper bound on the fraction of agents that choose wx or wy as a function of δ? The next two results help the reader answer these questions by characterizing the family of distributions over action sequences that can be rationalized.

Suppose the analyst observes chosen action sequences along with the associated realized states for a large number of decision problems. For Example 1, this data requirement would collate entries in one of six possible bins:

(*not invest*, $\emptyset$), good	(*invest*, *pull back*), good	(*invest*, *invest*), good
(*not invest*, $\emptyset$), bad	(*invest*, *pull back*), bad	(*invest*, *invest*), bad

The objective now is to explain when a joint distribution $\gamma \in \Delta(A \times \Omega)$ can be rationalized. For its dual counterpart, we say that a deviation rule dominates γ if it generates a strict improvement in expected payoffs. We call this *average dominance*, where the expectation over actions and states is taken for the distribution generated by the composition map between the deviation rule and whatever strategy a representative agent follows. Theorem 2 then provides the dual characterization that a joint distribution cannot be rationalized if and only if it is averagely dominated. This result boils down to a set of inequalities which correspond to obedience constraints familiar from information design (see Bergemann and Morris [2016]), extended here to a dynamic environment.

In some scenarios, the analyst may only observe the set of action sequences but not the realized states. So, in the context of Example 1, the analyst records which of the three possible action sequences were chosen by each firm in the data set, but does not know what were the underlying fundamentals associated with each of those decisions. The objective now is to explain when a marginal distribution $\overline{\gamma} \in \Delta(A)$ can be rationalized. The notion of dominance that pins down the duality is more nuanced. It combines ideas on dominance used for Theorems 1 and 2 resulting in an *intermediate* notion of *dominance*. It considers deviation rules that take the worst-case improvement over states for each action sequence, and then averages these values over action sequences using $\bar{\gamma}$. Theorem 3 states that a distribution $\overline{\gamma}$ cannot be rationalized if and only if it is intermediately dominated.

For Example 1, we can conclude that $2/3$ is the upper bound on the fraction of firms that can rationally be seen making the choice (*invest*, *pull back*). For Example 2, we show that the maximal fraction of agents that can rationally be seen to choose to wait in the first period is given by 0 when $\delta < 4/5$ and by $3 - 2/\delta$ when $\delta \geqslant 4/5$, which converges to 1 as δ converges to 1. There are two key steps here : the construction of the appropriate deviation rule for values greater than the upper bound and a rationalizing information structure for values below the upper bound.

By characterizing the empirical content of the model, the three theorems provide a clear method for testing for Bayesian rationality. If, for instance, more than two-thirds of firms in the data set facing Example 1 choose (*invest*, *pull back*), the Bayesian model is rejected. If instead one *assumes* Bayesian rationality, the results can be used to partially identify missing pieces of information from the agents preferences. For instance, the firm's choice to wait in Example 2 helps identify the cost of waiting to be $\delta \geqslant \frac{4}{5}$. More generally, for a population of firms, we show that a distribution that puts weight $\gamma^w > 0$ on wx (or wy) generates a lower bound given by $\delta \geqslant \max\left\{\frac{2}{3-\gamma^w}, \frac{4}{5}\right\}$.

Besides these potentially practical applications, our framework can also be used conceptually. As an illustration, we show that the set of actions sequences that can be rationalized is an increasing function of risk aversion—the more risk averse the agent, the harder it is to rule out action sequences as explained by some dynamic arrival of information.

The idea of deviation rules has been an important tool in characterizing behavior in decision theory, behavioral economics, information design, and mechanism design. These connections are explored in greater detail in Section 8, which points towards a wider applicability of the concepts and results developed here.

2 Model and definitions

2.1 Notation

A *stochastic map* from X to a finite set Y is a function $\alpha : X \to \Delta(Y)$, where $\Delta(Y)$ is the set of probability distributions over Y. We represent the probability assigned to y at the point x by $\alpha(y|x)$. The composition of two stochastic maps $\alpha : X \to \Delta(Y)$ and $\beta : Y \to \Delta(Z)$ is given by

$$\beta \circ \alpha(z|x) = \sum_{y \in Y} \beta(z|y)\alpha(y|x).$$

We can think of a lottery as a stochastic mapping whose domain is a singleton. Therefore, given $\alpha \in \Delta(Y)$ and $\beta : Y \to \Delta(Z)$, we write

$$\beta \circ \alpha(z) = \sum_{y \in Y} \beta(z|y)\alpha(y)$$

to be the probability with which z is chosen by $\beta \circ \alpha$.

For a real-valued function $u : Y \to \mathbb{R}$ and for a lottery $\alpha \in \Delta(Y)$, we denote by $u(\alpha) = \sum_{y \in Y} \alpha(y)u(y)$ the expected value of $u(.)$ under the distribution α.

Throughout the text, we consider a finite number of time periods $t = 1, \ldots, T$. For a collection of sets $(X_t)_{t=1}^T$, we will use the following notation

$$X^t = \prod_{\tau=1}^{t} X_\tau \qquad X = \prod_{\tau=1}^{T} X_\tau$$

with elements $\mathbf{x}^t \in X^t$ and $\mathbf{x} \in X$. Finally, a stochastic map $\alpha : X \to \Delta(Y)$ is said to be

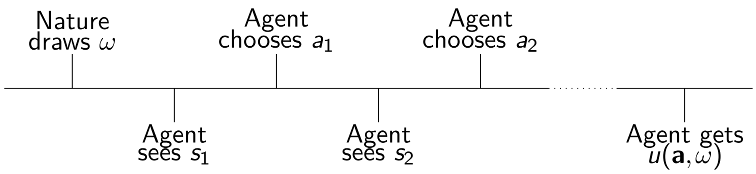


Figure 1: The timeline of signals and actions

adapted if the marginal probability of the first t terms of $\mathbf{y}$ depends only on the first t terms of $\mathbf{x}$; formally, it is adapted if the function

$$\sum_{y_{t+1},\dots,y_T} \alpha(y_1,\dots,y_t,y_{t+1},\dots y_T|x_1,\dots,x_t,x_{t+1},\dots,x_T)$$

is constant in $x_{t+1},\dots,x_T$.

2.2 The Bayesian model

In each time period t, the agent chooses an action a_t from a finite set A_t. Payoffs are determined after period T by a utility function $u(\mathbf{a}, \omega)$, which depends on the entire action sequence $\mathbf{a} = (a_1, \dots, a_T) \in A$ and a potentially unknown state of the world ω drawn from a finite set Ω. There are no other restrictions on the utility function.

The agent is informed about the underlying state of the world over time through a sequence of signals. The timeline of the dynamic decision problem is expressed in Figure 1. Every period, before taking an action, the agent observes a signal that is (potentially) correlated with the state of the world and with the signals she has observed in the past. Formally, the sequence of signals is generated by a *sequential information structure*:

Definition 1. *A* **sequential information structure** *is a sequence of finite sets of signals* $(S_t)_{t=1}^T$ *and a stochastic mapping* $\pi : \Omega \to \Delta(S)$.[4]

We will often refer to the sequential information structure simply as π; the set of signals shall be implicit. The agent's strategy maps each sequence of signals into a lottery over actions every period, with the restriction that the agent cannot base the choice of an action on signals that have not yet been revealed, which we call adaptedness.

Definition 2. *A* **strategy** *for the agent is an adapted stochastic mapping* $\sigma : S \to \Delta(A)$.[5]

Given the sequential information structure π and agent's strategy σ, the probability that the agent takes a given sequence of actions in each state of the world ω is given by $\sigma \circ \pi(\mathbf{a}|\omega)$.

[4] We can equivalently define the sequential information structure period-by-period as follows. Let $\pi = (\pi_t)_{t=1}^T$ be a family of stochastic mappings where $\pi_1 : \Omega \to \Delta(S_1)$, and $\pi_t : \Omega \times S^{t-1} \to \Delta(S_t)\ \forall\ 2 \le t \le T$. With the exception of zero probability events, we can deduce that the two definitions are equivalent. The minor distinction does not affect the agent's utility and is therefore irrelevant for our results. For a proof, see Lemma 3 in de Oliveira [2018].

[5] As with information structures, an equivalent way to think of the agent's strategy is a family of stochastic mappings $\sigma = (\sigma_t)_{t=1}^T$, where $\sigma_1 : S_1 \to \Delta(A_1)$, and $\sigma_t : S^t \times A^{t-1} \to \Delta(A_t)\ \forall\ 2 \le t \le T$. It is possible to deduce one formulation from the other.

Finally, given a prior $p \in \Delta(\Omega)$, she can evaluate her expected payoff:

$$U(\sigma, \pi, p) = \sum_{\omega \in \Omega} p(\omega) \sum_{\mathbf{a} \in A} \sigma \circ \pi (\mathbf{a}|\omega)\, u(\mathbf{a}, \omega).$$

The agent's problem then is to choose an optimal σ given π and p. Throughout the paper, we refer to this model of decision making as the *Bayesian model.*

Our goal is to characterize the empirical content of this model. To that end, we say that an action sequence can be *rationalized* if it can be chosen with positive probability by an optimizing agent with some information structure and some prior.

Definition 3. *An action sequence* $\mathbf{a} \in A$ *can be* **rationalized** *if there exists a triplet* (σ, π, p) *such that:*

1. $\sigma \in \arg\max_{\hat{\sigma}} U(\hat{\sigma}, \pi, p)$ *and*
2. $\sigma \circ \pi \circ p(\mathbf{a}) > \mathbf{0}$.[6]

This definition is permissive in the sense that an action sequence is considered rationalized even if its probability is very small, so long as it is positive. Moreover, because the agent sees a signal before choosing the first action, any two interior prior beliefs p and p' result in the same criterion, since we can always consider a signal distribution which updates from p to p' with positive probability. In that sense, the choice of prior, in addition to the choice of the sequential information structure, arms the analyst with more instruments than she requires to rationalize an action sequence. However, fixing a prior that puts zero probability on some states loses generality, since updated beliefs must also put zero probability on those states.[7]

To deduce that an action sequence cannot be rationalized, the analyst needs to work through all possible pairs (π, p), and show that the corresponding optimal strategy σ will not pick that action sequence with positive probability. Since the set of all sequential information structures is quite large, this poses a challenge. Our main goal is to find an alternative way to characterize the set of action sequences that cannot be rationalized.

3 The static problem

To fix ideas, it is easiest to start from the simple case of $T = 1$. In this static problem, the agent starts with a prior p, observes a signal s, and takes an action a, resulting in a payoff $u(a, \omega)$. Letting

$$q(\omega|s) = \frac{\pi(s|\omega)p(\omega)}{\pi \circ p(s)}$$

[6] Here $\sigma \circ \pi \circ p(\mathbf{a}) = \sum_{\omega} \sigma \circ \pi (\mathbf{a}|\omega)\, \mathbf{p}(\omega)$ (see Section 2.1).

[7] This logic can be pushed further: to determine the set of actions that can be rationalized going forward, the only relevant aspect of a belief is the set of states that have zero probability. So, a behavioral model where agents may violate the martingale condition of beliefs could rationalize the same set of action sequences as the Bayesian model, as long as its belief process agrees with the Bayesian belief process on which states have zero probability. We are grateful to Andrew Caplin for pointing this out to us.

denote the posterior belief of the agent upon seeing s, we can rewrite the agent's expected utility from choosing strategy σ as:

$$U(\sigma, \pi, p) = \sum_{\omega, a, s} u(a, \omega)\sigma(a|s)\pi(s|\omega)p(\omega) = \sum_{\omega, a, s} u(a, \omega)\sigma(a|s)q(\omega|s)\pi \circ p(s). \quad (1)$$

This makes the agent's problem separable in s, so it reads: for each s, choose an action a to maximize

$$\sum_{\omega \in \Omega} u(a, \omega)q(\omega|s). \quad (2)$$

Therefore, an action can be rationalized if and only if it is a *best-response to some posterior belief* q. Hence, to find if an action can be rationalized, we can restrict attention to the case where the agent starts with a "prior q" and learns nothing thereafter. In particular, if an action can be rationalized, there is a triplet (σ, π, p) where σ is optimal and chooses that action with probability 1. To summarize:

Remark 1. *Let $T = 1$. Then the following statements are equivalent:*

1. *a is a best-response to some belief q;*
2. *There exists (σ, π, p) such that σ maximizes (1) with $\sigma(a) > 0$.*
3. *There exists (σ, π, p) such that σ maximizes (1) with $\sigma(a) = 1$.*[8]

An elegant duality result by Wald [1949] and Pearce [1984] characterizes what it means for an action to be rationalized in the static model. The result states that, in a two-player game,

Lemma 1 (Wald-Pearce)**.** *An action is never a best response if and only if it is strictly dominated by some mixed strategy.*

In our context, think of a game where Player 1 is our agent, choosing action a, and Player 2 is Nature, choosing state ω. A mixed strategy $\alpha \in \Delta(A)$ strictly dominates a if and only if $u(\alpha, \omega) > u(a, \omega)$ for all $\omega \in \Omega$, where $u(\alpha, \omega)$ is the expected utility of following that mixed strategy. An action a is then said to be strictly dominated if there exists a mixed strategy α that strictly dominates it.

Given Remark 1, a is never a best response if and only if it cannot be rationalized. Therefore

Corollary 1. *For $T = 1$, an action a cannot be rationalized if and only if it is strictly dominated.*

[8]It is worth trying to extrapolate the contents of Remark 1 to the case of $T > 1$. It is easy to see that the equivalence between parts 2 and 3 no longer holds in Example 1. Specifically, (*invest*, *pull back*) can be rationalized with positive probability, but never with probability 1. Moreover, if we simply invoke a static information structure wherein the agent learns all possible information prior to taking all the actions, the same example shows that parts 1 and 2 of Remark 1 fail to be equivalent as well. In a nutshell, the sequential structure of the problem matters.

The key idea behind the Wald-Pearce lemma is that it is possible to invert the order of quantifiers in the statement "for all $q \in \Delta(\Omega)$, there exists $\alpha \in \Delta(A)$ such that $\mathbb{E}_q [u(a, \omega)] < \mathbb{E}_q [u(\alpha, \omega)]$". This can be seen, for example, by constructing a zero-sum game where nature picks the belief q and the agent picks an alternative action α (possibly mixed). Using the min-max theorem, we get that

$$\min_q \max_\alpha \mathbb{E}_q[u(\alpha, \omega) - u(a, \omega)] = \max_\alpha \min_q \mathbb{E}_q[u(\alpha, \omega) - u(a, \omega)].$$

When a cannot be rationalized, the above expression is positive and bounded away from zero. More specifically, the positivity of the left-hand side is equivalent to a not being rationalized and the positivity of the right hand side is equivalent to it being strictly dominated.

The two theoretical challenges for us therefore are (i) to formulate the right notion of what it means for an action sequence to be dominated in the sequential model, and (ii) to establish the appropriate inversion of quantifiers for our framework. We start first by defining the appropriate notion of domination in the sequential model.

4 Deviation rules and true dominance

4.1 A necessary but not sufficient condition

An obvious notion of dominance that does not rely on information structures is the following: a sequence of actions is "dominated" if there exists another sequence of actions that does strictly better in every state of the world. We will refer to this as *apparent dominance*. Recollect that the payoff from a randomized action sequence $\alpha \in \Delta(A)$ is denoted by $u(\alpha, \omega) = \sum_{\mathbf{a} \in A} \alpha(\mathbf{a}) u(\mathbf{a}, \omega)$, where $\alpha(\mathbf{a})$ refers to the probability of action sequence $\mathbf{a}$ under α.

Definition 4. *An action sequence* $\mathbf{a} \in A$ *is* **apparently dominated** *if there exists a randomized action sequence* $\alpha \in \Delta(A)$ *such that*

$$u(\alpha, \omega) > u(\mathbf{a}, \omega) \ \forall \ \omega \in \Omega.$$

Every action sequence that cannot be rationalized is apparently dominated, making it a necessary condition for our endeavored characterization. That is, if an action sequence is not apparently dominated, we can always find an information structure such that the optimal strategy corresponding to it chooses the action sequence with positive probability. The following Lemma formalizes the claim.

Proposition 1. *Suppose* $\mathbf{a} \in A$ *cannot be rationalized. Then,* $\mathbf{a}$ *must be apparently dominated.*

Proof. Suppose $\mathbf{a}$ is not apparently dominated. By Lemma 1, the Wald-Pearce Lemma, $\mathbf{a}$ must be a best-response to some static "belief p". Letting p be the prior and π be completely uninformative, the best response to (p, π) is the strategy that always chooses $\mathbf{a}$. □

Even though apparent dominance is a demanding condition, it is possible for an apparently dominated action sequence to be rationalized. In Example 1, the action sequence $a_1 = invest$ and $a_2 = pull\ back$ is apparently dominated by the action sequence $a_1 = not\ invest$ and $a_2 = \emptyset$. Yet it is easy to construct an information structure where it will be optimal for the agent to choose (*invest*, *pull back*) with positive probability.[9]

Notice that the apparent dominance of (*invest*, *pull back*) can be established simply by comparing its payoffs with that of *not invest*. The payoffs for (*invest*, *invest*) are therefore irrelevant. Yet, when the state good is very likely, these payoffs are precisely what motivates the agent to do the initial investment. When we see that the agent chose (*invest*, *pull back*), the fact that the agent *could* have ended up choosing (*invest*, *invest*) makes those payoffs relevant.

Therefore, something more than apparent dominance is required for an action sequence to not be rationalized. In addition to improving upon the action sequence under consideration, that "more" needs to evaluate other sequences of actions that the agent might expect to have chosen. This motivates the definition of a deviation rule, which prescribes not only how the agent should deviate in the *observed* action sequence, but in *every* other possible action sequence as well.

4.2 Deviation rules and true dominance

A *deviation rule* is an adapted mapping $D : A \to \Delta(A)$, where recollect that being adapted means that the marginal distribution on A^t, the (potentially random) deviation strategy for the first t periods, depends only on A^t, the first t elements of the original strategy from which the agent is deviating. We can think of the deviation rule as a list of alternative actions the agent would take as a function of the actions she originally intended to take. Importantly, a deviation rule is a fully prescribed plan so that if σ is the original strategy, then $D \circ \sigma(\mathbf{a}|\mathbf{s})$ too is a well-defined strategy.

Now, we are in a position to define the appropriate notion of dominance for our model.

Definition 5. *A deviation rule* $D : A \to \Delta(A)$ **dominates** *an action sequence* $\mathbf{a}$ *if*

1. $u(D(\mathbf{a}), \omega) > u(\mathbf{a}, \omega)$ *for all* $\omega \in \Omega$.

2. $u(D(\mathbf{b}), \omega) \geqslant u(\mathbf{b}, \omega)$ *for all* $\mathbf{b} \in A$ *and* $\omega \in \Omega$.

We say that $\mathbf{a}$ *is* **truly dominated** *if there exists a deviation rule that dominates it.*

Notice that the second part of the definition demands that the payoff induced by the deviation rule shouldn't become worse for any other action sequence in any other state. Moreover, there's no visible time dimension in the definition above; time is implicit in the condition that D must be adapted. For $T = 1$, the same definition applies, but the condition that D is

[9] The first period signal tells the agent that the good state is highly likely, only to reveal in period two through the second signal that the bad state is now more likely.

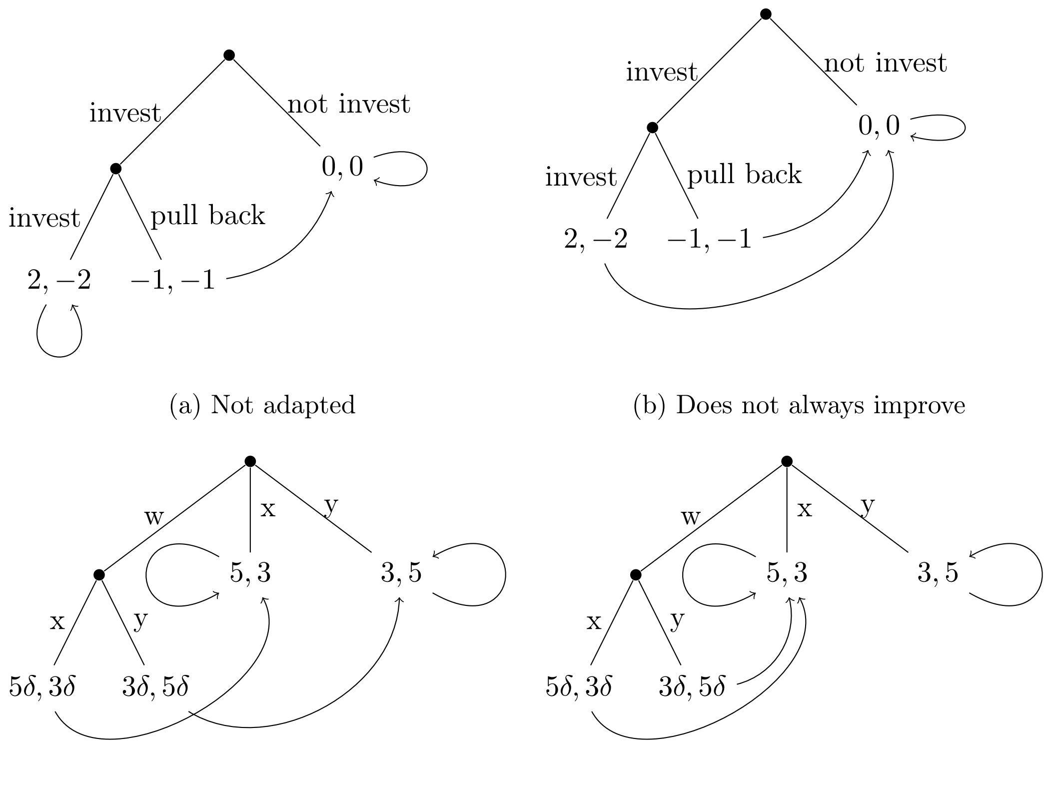


Figure 2: Deviation rules for Examples 1 and 2

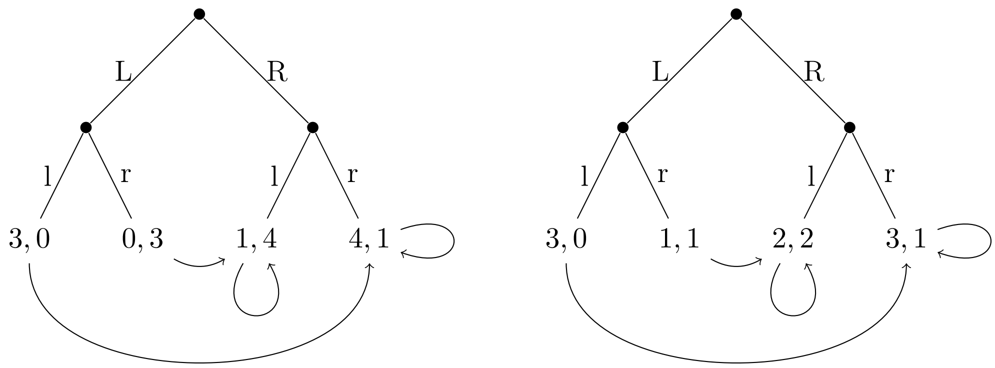
adapted becomes vacuous and so does the second part of the definition. In that case, if a is strictly dominated by α, we can define a deviation rule D_α which takes a to α and does not change any other actions. D_α then dominates a according to the definition above.

When $T > 1$, the adaptedness restriction prevents the construction of such a simple deviation rule—if D specifies a change for the first action in the sequence $\mathbf{a}$, then it must specify the same change for all sequences $\mathbf{b}$ which share that same first action, and so on. The second condition and the embedded notion of adaptedness in the definition impose meaningful restrictions when $T > 1$, encapsulating the distinction between true and apparent dominance.

4.3 Discussion

To better grasp the definitions of deviation rule and true dominance, here we illustrate the concepts in the context of our examples. For the decision trees depicted in Figure 2, each complete sequence of actions corresponds to a terminal node. Thus any mapping from sequences of actions into sequences of actions is depicted as arrows between terminal nodes.

Figures 2a and 2b depict the decision tree for Example 1. Since the sequence of actions (*invest, pull back*) is apparently dominated by *not invest*, we may try to find a deviation rule that dominates (*invest, pull back*). The simplest such proposal would be that the agent should choose *not invest* whenever she was going to choose (*invest, pull back*), as shown in Figure 2a. However, at the time when the agent is choosing to invest, she may not yet know whether she



(a) Adapted and improves upon Ll & Lr (b) Adapted and improves upon Lr

Figure 3: Deviation rules with history dependence

will pull back in the future. The impracticality of this proposal is reflected in the fact that this "deviation rule" is not adapted. If we want the agent to never invest whenever she was going to choose (*invest,pull back*), we must also recommend that she never invest when she was going to choose (*invest, invest*), as in 2b. But although the deviation rule in 2b is now adapted, it worsens payoffs for the action sequence (*invest, invest*) in the good state; thus, it violates part 2 of Definition 5.

Similarly, in the waiting example, the "deviation rule" depicted in Figure 2c is not adapted, since it represents the infeasible advice "whatever you would choose in the second period, choose the same in the first period". The deviation rule in Figure 2d represents the advice "if you were thinking about waiting, choose x instead", which is adapted. When $\delta < \frac{3}{5}$, it dominates wx and wy, but when $\delta > \frac{3}{5}$ it does not dominate wx nor wy, because x may give a strictly lower payoff than wy. For the tightest possible statement, we therefore constructed the deviation rule $wx \mapsto \frac{1}{2}x + \frac{1}{2}y$, $wy \mapsto \frac{1}{2}x + \frac{1}{2}y$, $x \mapsto y$ and $y \mapsto y$ which (simultaneously) truly dominates wx and wy if and only if $\delta < \frac{4}{5}$.

Our examples so far have featured simple first-period deviations. Figure 3 shows how history-dependent deviations may be required to establish that an action sequence is truly dominated. In Figure 3a, both (L, l) and (L, r) are truly dominated by the deviation rule depicted. Despite having the same first period action, (L, l) and (L, r) are deviated to different action sequences: (R, r) and (R, l), respectively; hence the history dependence in deviations. Analogously, (Lr) is shown to be truly dominated in 3b by the same deviation rule, but note that here (L, l) is not truly dominated.[10]

5 The main result

We now state our main result.

Theorem 1. *A sequence of actions cannot be rationalized if and only if it is truly dominated.*

[10]The difference between the two examples in Figure 3 is conceptually interesting and subtle; it is discussed further in Sections 8 and 9.5.

The theorem provides a tight characterization of the set of action sequences that cannot be rationalized. Through its duality formulation, it simplifies their identification by requiring the analyst to construct *one* deviation rule as opposed to treading through the family of *all* sequential information structures.

The steps involved in establishing this result are divided into three subsections. First, we state the *obedience principle*: any sequential information structure is equivalent to a canonical information structure, wherein at each point in time the agent is recommended an action which is in her own interest to follow. Second, we state a *Generalized Separation Lemma* that will allow us to invert the order of quantifiers. Third, we put all the arguments together and prove the result.

5.1 Obedience principle

The set of all possible signals can be a very large space to work with. We can in fact restrict attention to a set of canonical signal structures, without any loss of generality. In keeping with the tradition in mechanism design, we call this result the obedience principle. It is analogous to the obedience principle in Myerson [1986], Forges [1986], Kamenica and Gentzkow [2011], and Bergemann and Morris [2016].

First, we define the subset of canonical sequential information structures. In what follows, let Id_A refer to the identity mapping from A to A.

Definition 6. *(σ, π, p) is an* **obedient triple** *if $S = A$ and $\sigma = Id_A$.*

An obedient triple is given by a prior, an information structure which recommends an action, and a strategy of the agent which always obeys the recommendation. When an action sequence can be rationalized with an obedient triple, we say that it has an obedient rationalization. We can now state and prove the obedience principle.

Lemma 2 (Obedience principle)**.** *If* $\mathbf{a}$ *can be rationalized then it has an obedient rationalization.*

Proof. Suppose that $\mathbf{a}$ is rationalized by (σ, π, p). We show that $\mathbf{a}$ is also rationalized by $(Id_A, \sigma \circ \pi, p)$. First, note that

$$\sigma \circ \pi \circ p\,(\mathbf{a}) = Id_A \circ (\sigma \circ \pi) \circ p\,(\mathbf{a})$$

by associativity of composition, and that Id_A is adapted. Hence, if $\mathbf{a}$ is chosen with positive probability under (σ, π, p), it also is under $(Id_A, \sigma \circ \pi, p)$. Now we must show that Id_A will be optimal for $(\sigma \circ \pi, p)$ whenever σ is optimal for (π, p). Suppose that an alternate strategy $D : A \to \Delta(A)$ does better than Id_A when facing $(\sigma \circ \pi, p)$. In terms of payoff, it is easy to check that $U\,(Id_A, \sigma \circ \pi, p) = U\,(\sigma, \pi, p)$ and $U\,(D, \sigma \circ \pi, p) = U\,(D \circ \sigma, \pi, p)$. So if $(D, \sigma \circ \pi, p)$ gives a higher expected payoff than $(Id_A, \sigma \circ \pi, p)$, then the deviation $(D \circ \sigma, \pi, p)$ gives a higher payoff than (σ, π, p) as well, implying that σ was not optimal. □

5.2 Generalized Separation Lemma

The standard min-max theorem used to prove the Wald-Pearce lemma in Section 3 cannot be directly applied, one of the reasons being that the argument under "max" is not a compact set. To get around this problem, we state and prove a version of the hyperplane separation result that helps us flip the order of quantifiers in the final step of the proof of Theorem 1.

Lemma 3 (Generalized Separation Lemma)**.** *Let $X \subset \mathbb{R}^m$ be an evenly convex polyhedron and $Y \subset \mathbb{R}^n$ be a polytope.*[11] *If $f : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$ is affine in each variable, then the following statements are equivalent:*

1. *For every $x \in X$, there exists a $y \in Y$ such that $f(x, y) > 0$;*
2. *There exists $y \in Y$ such that, for every $x \in X$, $f(x, y) > 0$.*

In the mathematical hierarchy of ideas, all min-max type results are essentially separating hyperplane theorems, which in turn are duality results. In our setup, the hyperplane argument is somewhat simplified by the fact that we operate in a finite setting, and thus, the problem reduces to a linear program, which is essentially what Lemma 3 captures.

5.3 Proof of Theorem 1

The "only if" direction: if **a** *is truly dominated, it cannot be rationalized.* Let D be a deviation rule which shows that **a** is truly dominated. We show that any strategy that plays **a** with positive probability cannot be optimal. Indeed, given an arbitrary (σ, π, p), we can define an alternative strategy $\tilde{\sigma} = D \circ \sigma$. Now consider how the expected payoff of the agent changes by switching from σ to $\tilde{\sigma}$. Let γ denote the joint distribution over $(\mathbf{b}, \omega)$ which is induced by (σ, π, p). The difference in payoffs then becomes

$$U(\tilde{\sigma}, \pi, p) - U(\sigma, \pi, p) = \mathbb{E}_\gamma[u(D(\mathbf{b}), \omega) - u(\mathbf{b}, \omega)].$$

For each $(\mathbf{b}, \omega)$, this difference is non-negative, with strict inequality for $\mathbf{b} = \mathbf{a}$. Hence if γ puts positive probability on **a**, the overall difference will be strictly positive, meaning that the agent benefits strictly from deviating to $\tilde{\sigma}$. The exact inequalities that show $\tilde{\sigma}$ to be an improvement over σ are presented in Claim 1 in the appendix.

The "if" direction: if **a** *cannot be rationalized, it is truly dominated.* Given an action sequence **a** which cannot be rationalized, we must find a deviation rule D which shows that it is truly dominated. Letting $\Gamma(\mathbf{a}) = \{(\sigma, p, \pi) | \sigma \circ \pi \circ p(\mathbf{a}) > 0\}$, we can write the statement "**a** cannot be rationalized" as

$$\forall\ (\sigma, \pi, p) \in \Gamma(\mathbf{a})\ \exists\ \hat{\sigma} \text{ s.t. } U(\hat{\sigma}, \pi, p) > U(\sigma, \pi, p).$$

[11] An evenly convex polyhedron is a set defined by a finite number of linear inequalities, weak or strict. A polytope is the convex hull of finite number of points or, equivalently, a compact set defined by finitely many linear inequalities, weak or strict.

By the Obedience Principle (Lemma 2), the statement "**a** cannot be rationalized" is equivalent to the statement "**a** cannot be rationalized by an obedient triple". This means that we can, without loss, restrict attention to $\pi : \Omega \to \Delta(A)$ and to $\sigma = Id_A$ in the statement above. Moreover, given that the set of signals is now A, all other strategies $\hat{\sigma}$, are simply the set of all deviation rules $D : A \to \Delta(A)$. Incorporating these, we get the equivalent statement

$$\forall p \in \Delta(\Omega) \; \& \; \pi : \Omega \to \Delta(A) \text{ s.t. } \pi \circ p(\mathbf{a}) > 0 \; \exists \; D : A \to \Delta(A) \text{ s.t. } U(D, \pi, p) > U(Id_A, \pi, p).$$

Our goal is to switch the order of quantifiers in this statement, which would then produce the deviation rule we seek. Notice that trying to use a min-max theorem to achieve this inversion of quantifiers would run into multiple problems. That is, if we wrote

$$\inf_{\substack{(\pi,p) \text{ s.t.} \\ \pi\circ p(\mathbf{a})>0}} \max_{D} \; [U\,(D, \pi, p) - U\,(Id_A, \pi, p)]$$

the objective function would not be linear in the vector (π, p), the set we are minimizing over would not be compact, and the value of the infimum would actually be zero. Therefore, we invoke the Generalized Separation Lemma to do the needful. To use the lemma, we simplify the problem further.

First, the objective function can made linear through a simple change of variables: Let $\gamma \in \Delta(A \times \Omega)$ be the joint distribution on $A \times \Omega$ induced by the pair (π, p). That is, $\gamma(\mathbf{b}, \omega) = \pi(\mathbf{b}|\omega)p(\omega)$. The set of joint distributions we are considering consists of those γ whose marginal probability on **a** is strictly positive. Doing this, the objective function becomes

$$\mathbb{E}_\gamma[u(D(\mathbf{b}), \omega) - u(\mathbf{b}, \omega)],$$

which is bilinear in (γ, D).

Finally, the statement that **a** cannot be rationalized can be rewritten as

$$\forall \, \gamma \in \Delta(A \times \Omega) \text{ with } \gamma(\mathbf{a}) > 0, \exists \; D : A \to \Delta(A) \text{ s.t. } \; \mathbb{E}_\gamma[u(D(\mathbf{b}), \omega) - u(\mathbf{b}, \omega)] > 0.$$

Let $X = \{\gamma \in \Delta(A \times \Omega) \text{ s.t. } \gamma(\mathbf{a}) > 0\}$, $Y = \{D : A \to \Delta(A) \text{ s.t. } D \text{ is adapted}\}$, and $f \equiv \mathbb{E}_\gamma[u(D(\mathbf{b}), \omega) - u(\mathbf{b}, \omega)]$. In Claim 2 in the appendix, we show that X is an evenly convex polyhedron, Y is a polytope and f is an affine function in γ and D. Thus, invoking Lemma 3, we can flip the order of quantifiers to get:

$$\exists \; D : A \to \Delta(A) \text{ s.t. } \forall \, \gamma \in \Delta(A \times \Omega) \text{ with } \gamma(\mathbf{a}) > 0 : \; \mathbb{E}_\gamma[u(D(\mathbf{b}), \omega) - u(\mathbf{b}, \omega)] > 0.$$

Let D^* be the deviation rule that solves this problem. By construction, D^* dominates **a**, and thus **a** is truly dominated.

5.4 One deviation to rule them all

Although we proved Theorem 1 by showing that there exists a deviation rule for each truly dominated action sequence, it is easy to find a single deviation rule that simultaneously dominates every truly dominated action sequence. Analogously, it is possible to find a single information structure that simultaneously rationalizes every action sequence that is not truly dominated. Here we prove the existence of such a deviation rule and information structure:

Corollary 2. *Let $\mathcal{A} \subseteq A$ denote the set of truly dominated action sequences. Then*

1. *there exists a deviation rule D that simultaneously dominates every action sequence in $\mathcal{A}$ and*

2. *there exists a triple (σ, π, p) that simultaneously rationalizes every action sequence in $A \backslash \mathcal{A}$.*

Put together, this means that the entire set of truly dominated action sequences can be characterized by one pair, consisting of a single information structure and a single deviation rule. To prove part 1, we simply take D to be a strict convex combination of each deviation rule corresponding to each action sequence that is truly dominated. To prove part 2, we take a strict convex combination of all obedient rationalizations of the elements in $A \backslash \mathcal{A}$.

Broadly, the idea for both constructions boils down to the fact that Definition 3 required the action to be picked with any positive probability. Therefore, as long as the deviation rule that was picked to truly dominate an action sequence is picked with positive probability in part 1 of Corollary 2, it will improve upon expected payoff of the agent. Similarly, if the information structure that rationalizes an action sequence is picked with positive probability in part 2 of Corollary 2, the "compound lottery" will also pick that action sequence will positive probability.

6 Rationalizing distributions

In the previous section, we characterized the empirical content of the Bayesian model when the analyst observes a single action sequence. We now consider a situation where the analyst has information about a large population of agents, so that his data consists of an entire distribution of chosen action sequences. Two duality results are presented which assume varying levels of data availability. Deviation rules, once again, play a central role in the characterizations. We end the section with a brief discussion on a comparison between these results with Theorem 1.

6.1 Distributions over actions and states

At first, we assume that the analyst has access to a rich dataset which records both action sequences and corresponding realized states. That is, the analyst observes an entire joint distribution $\gamma \in \Delta(A \times \Omega)$. The distributions that can result from a Bayesian model can then be defined by a modification of Definition 3.

Definition 7. *A distribution $\gamma \in \Delta(A \times \Omega)$ can be* **rationalized** *if there exists (σ, π, p) such that*

1. $\sigma \in \arg\max_{\hat{\sigma}} U(\hat{\sigma}, \pi, p)$, *and*

2. $\gamma(\mathbf{a}, \omega) = \sigma \circ \pi(\mathbf{a}|\omega) \cdot p(\omega)\ \forall\ \mathbf{a} \in A,\ \omega \in \Omega.$[12]

Thus a joint distribution γ is rationalized if there exists a sequential information structure and a prior such that, in best responding to them, the agent's optimal strategy generates γ. In the context of Example 1, we are assuming here that a large number of firms face some prior and sequential information structure, and in best responding to it, a joint distribution over the two states and three possible action sequences is produced, which the analyst seeks to rationalize.

Our notion of dominance must take into account the distributional structure of information available to the analyst. Earlier, we required the deviation rule to improve upon every action sequence and state. Since the criterion of rationalization is now stronger, the notion of dominance must be weaker. Therefore, we look at the *average improvement* brought about by a deviation rule.

Definition 8. *A deviation rule $D : A \to \Delta(A)$* **dominates** *a distribution $\gamma \in \Delta(A \times \Omega)$ if*

$$\sum_{\mathbf{a},\omega} [u(D(\mathbf{a}), \omega) - u(\mathbf{a}, \omega)]\, \gamma(\mathbf{a}, \omega) > 0.$$

We say that γ is **dominated on average** *if there exists a deviation rule that dominates it.*

To understand the appropriateness of this notion of dominance, we can again follow the logic of the obedience principle (Lemma 2): If γ can be rationalized, it must be possible to find an obedient rationalization. Under such an obedient rationalization, any alternative strategy $\hat{\sigma}$ is an adapted mapping from A to A, and hence a deviation rule. If, however, γ cannot be rationalized, then the candidate obedient strategy is not optimal, and there must be an alternative strategy—a deviation rule—that improves upon it. This proves the following dual characterization.

Theorem 2. *A distribution $\gamma \in \Delta(A \times \Omega)$ cannot be rationalized if and only if it is dominated on average.*

Proof. The "only if" direction: if γ is dominated on average, it cannot be rationalized. Let D be a deviation rule which shows that γ is dominated on average. We show that any strategy

[12] Note that for a fixed $\gamma \in \Delta(A \times \Omega)$, the prior p is necessarily its marginal on Ω. This is implicit in part 2 of the definition. We keep the choice of the triplet (σ, π, p) in the definition to maintain consistency with Definition 3.

that produces γ as the joint distribution over $A \times \Omega$ cannot be optimal. Suppose (σ, π, p) induces γ, and define an alternative strategy $\tilde{\sigma} = D \circ \sigma$. The difference in payoffs then becomes

$$U(\tilde{\sigma}, \pi, p) - U(\sigma, \pi, p) = \mathbb{E}_\gamma[u(D(\mathbf{b}), \omega) - u(\mathbf{b}, \omega)].$$

which is strictly positive since D dominates γ. Hence γ cannot be induced by an optimal strategy and cannot be rationalized.

The "if" direction: if γ cannot be rationalized, it is dominated on average. Fix a $\gamma \in \Delta(A \times \Omega)$ which cannot be rationalized. We must find a deviation rule D which shows that it is dominated on average. Letting $\Gamma(\gamma) = \{(\sigma, \pi, p) | \gamma(\mathbf{a}, \omega) = \sigma \circ \pi(\mathbf{a}|\omega) \cdot p(\omega)\}$, we can write the statement "γ cannot be rationalized" as

$$\forall\ (\sigma, \pi, p) \in \Gamma(\gamma)\ \exists\ \hat{\sigma} \text{ s.t. } U(\hat{\sigma}, \pi, p) > U(\sigma, \pi, p).$$

Following the same steps as in the proof of Theorem 1, we use the obedience principle to rewrite the above statements as

$$\exists\ D : A \to \Delta(A) \quad \text{s.t.} \quad \mathbb{E}_\gamma[u(D(\mathbf{b}), \omega) - u(\mathbf{b}, \omega)] > 0.$$

which shows that there exists a D that dominates γ, and hence γ is dominated on average.

□

An equivalent way of thinking about Theorem 2 is this: a distribution γ can be rationalized if, for all deviation rules $D : A \to \Delta(A)$,

$$\sum_{\mathbf{a}, \omega} [u(\mathbf{a}, \omega) - u(D(\mathbf{a}), \omega)] \gamma(\mathbf{a}, \omega) \geqslant 0.$$ [13]

Now, fix p to be the marginal of γ on Ω, $S = A$, and $\sigma = Id_A$. Then, noting that $\gamma(\mathbf{a}, \omega) = \pi(\mathbf{a}|\omega) p(\omega)$, the above inequality gives us a unique obedient triplet that rationalizes γ.

The result can therefore be seen as a counterpart to obedience constraints in information design (see surveys by Bergemann and Morris [2019] and Kamenica [2019]). In the lexicon of that literature, all distributions $\gamma \in \Delta(A \times \Omega)$ that satisfy the above inequality for all deviation rules can be supported as a "Bayes correlated equilibrium" of our decision problem.[14]

Theorem 2 also sheds light on how to find information structures that rationalize particular action sequences. An action sequence $\mathbf{a}$ can be rationalized if there exists a distribution γ that can be rationalized and puts positive probability on $\mathbf{a}$. Any such distribution can be interpreted directly as an information structure that signals action recommendations. Since

[13] It can be noted that the use of mixed deviation rules in this statement is redundant, if the set of inequalities hold for all pure deviation rules $D : A \to A$, the statement is still true.

[14] Just as the objective of the linear program in information design is to identify the set of binding obedience constraints, the objective of the above result is to identify the critical deviation rule.

Theorem 2 characterizes all distributions that can be rationalized, we need only to look at those that put positive probability on **a** to find all obedient triples that rationalize **a**.

6.2 Distributions over actions

In some scenarios, observing the realized state of the world might be difficult or even impossible for the analyst. For instance, the analyst may observe the investment decisions made by the population of firms in Example 1, but may not observe whether the underlying market forces were good or bad for them. In that case, we can ask whether a given distribution over action sequences can be rationalized.

Since a triple (σ, π, p) defines a joint distribution $\gamma \in \Delta(A \times \Omega)$, we consider the set of such joint distributions that is consistent with a given marginal $\overline{\gamma} \in \Delta(A)$,

$$\Gamma(\bar{\gamma}) = \left\{\gamma \in \Delta\left(A \times \Omega\right) | \sum_{\omega} \gamma\left(\mathbf{a}, \omega\right) = \bar{\gamma}\left(\mathbf{a}\right) \right\}.$$

The definition of rationalizing a distribution of action sequences then corresponds to Definition 7, but where we allow any joint distribution with the given marginal $\overline{\gamma}$.

Definition 9. *A distribution* $\bar{\gamma} \in \Delta(A)$ *can be* **rationalized** *if there exists* (σ, π, p) *such that*

1. $\sigma \in \arg\max_{\hat{\sigma}} U\left(\hat{\sigma}, \pi, p\right)$, *and*
2. $\sigma \circ \pi \circ p = \bar{\gamma}$.

To prove that $\overline{\gamma}$ cannot be rationalized, we must show that it is impossible to find any $\gamma \in \Gamma(\overline{\gamma})$ that can be rationalized in the sense of Definition 7. One way to do this is to find a deviation rule that works simultaneously for all distributions in $\Gamma(\overline{\gamma})$. This idea leads us to the concept of *intermediate domination.*

Definition 10. *A deviation rule* $D : A \to \Delta\left(A\right)$ **dominates** *a distribution* $\overline{\gamma} \in \Delta(A)$ *if*

$$\sum_{a} \min_{\omega} \left[u\left(D\left(\mathbf{a}\right), \omega\right) - u\left(\mathbf{a}, \omega\right)\right] \overline{\gamma}\left(\mathbf{a}\right) > 0.$$

We say that $\overline{\gamma}$ *is* **intermediately dominated** *if there exists a deviation rule that dominates it.*

So we say that D dominates $\overline{\gamma}$ if the average improvement is positive, even when we choose the worst possible state for each action sequence. The requirement of Definition 10 is thus intermediate to the notions of true dominance, which looks at the worst improvement across states and actions, and average dominance, which looks at the average improvement across states and actions.

If $\overline{\gamma}$ is intermediately dominated, the deviation rule that dominates $\overline{\gamma}$ demonstrates that any $\gamma \in \Gamma(\overline{\gamma})$ cannot be rationalized, and thus $\overline{\gamma}$ cannot be rationalized. As the reader might suspect at this point, the converse also holds.

Theorem 3. *A distribution $\overline{\gamma} \in \Delta(A)$ cannot be rationalized if and only if it is intermediately dominated.*

Proof. As in the case of Theorem 1, one direction is easy. Suppose D dominates $\overline{\gamma}$ in the sense of Definition 10, and by contradiction γ can be rationalized by some triplet (σ, π, p). Then, the alternative strategy $\hat{\sigma} = D \circ \sigma$ gives the agent a strictly higher expected payoff when facing (π, p). Thus, it must be that $\overline{\gamma}$ cannot be rationalized.

Conversely, suppose that $\overline{\gamma}$ cannot be rationalized. Then, by Theorem 2, for every $\gamma \in \Delta(A \times \Omega)$ with marginal $\gamma_A = \overline{\gamma}$, there exists a deviation rule D such that

$$\sum_{a,\omega} \left[u\left(D\left(\mathbf{a}\right), \omega\right) - u\left(\mathbf{a}, \omega\right)\right] \gamma\left(\mathbf{a}, \omega\right) > 0.$$

By Lemma 3, we can invert the order of quantifiers: there exists a deviation rule D such that for all $\gamma \in \Delta(A \times \Omega)$ with marginal $\gamma_A = \overline{\gamma}$, we have:

$$\sum_{a,\omega} \left[u\left(D\left(\mathbf{a}\right), \omega\right) - u\left(\mathbf{a}, \omega\right)\right] \gamma\left(\omega|\mathbf{a}\right) \overline{\gamma}\left(\mathbf{a}\right) > 0.$$

Now, we are free to choose $\gamma(\omega|\mathbf{a})$ arbitrarily given $\mathbf{a}$, which means that the inequality above must hold for every $\gamma(\omega|\mathbf{a})$. This happens precisely when

$$\sum_{a} \min_{\omega} \left[u\left(D\left(\mathbf{a}\right), \omega\right) - u\left(\mathbf{a}, \omega\right)\right] \overline{\gamma}\left(\mathbf{a}\right) > 0.$$

□

Noting the system of inequalities that define the three notions of dominance can help understand the intermediate nature of this result. Condition (b) in Definition 5 implies that for true dominance, we need

$$\min_{\mathbf{b}\in A} \min_{\omega\in\Omega} \left[u\left(D\left(\mathbf{b}\right), \omega\right) - u\left(\mathbf{b}, \omega\right)\right] \geqslant 0$$

with a strict inequality for $\mathbf{b} = \mathbf{a}$, the particular action sequence being dominated. Intermediate dominance (Definition 10) modifies this identity by taking the average over the distribution of action sequences using the marginal $\bar{\gamma}$:

$$\sum_{b} \min_{\omega} \left[u\left(D\left(\mathbf{b}\right), \omega\right) - u\left(\mathbf{b}, \omega\right)\right] \overline{\gamma}\left(\mathbf{b}\right) > 0.$$

And, finally, average dominance (Definition 8) takes average over both action sequences and states using the knowledge of the joint distribution γ:

$$\sum_{\mathbf{b},\omega} \left[u\left(D\left(\mathbf{b}\right), \omega\right) - u\left(\mathbf{b}, \omega\right)\right] \gamma\left(\mathbf{b}, \omega\right) > 0.$$

In Section 7, we illustrate how Theorem 3 can be applied to pin down the maximal probability with which an apparently dominated action sequence can be taken in the context of Examples 1 and 2. Before that we provide a brief comparative discussion of the three main results in terms of the empirical contents of the respective Bayesian models they characterize.

6.3 Discussion

The richer predictions afforded by Theorem 2 and 3 come at the backdrop of several assumptions on the environment that we now discuss.

Theorems 2 and 3 assumed that the analyst could observe a whole distribution of action sequences. What does that mean? A natural interpretation is that we have data on choices for a large population of agents and the analyst observes the empirical distribution of choices for that population. Under that interpretation, we also assumed that all agents in that population have the same utility function, prior, and information structure. Moreover, we also need to assume that the signals seen by each agent are independent. It is only then we can conclude by a standard law of large numbers argument that the empirical distribution should be close to the theoretical distribution generated by the triple (σ, π, p). No such assumptions are required for Theorem 1.

The independence assumption, in particular, is quite important. Consider what would be the effect of relaxing it in Example 1. If we allow any correlation, we could have that every agent faces a problem with the same relevant state of the world (G or B) and sees the same public signal. It is perfectly plausible that the public signal be first good and then bad, leading all agents in the population to choose the action sequence (*invest*, *pull back*). But then we could observe 100% of the population choose an apparently dominated action sequence, which would seem to violate Theorem 3. Without independence, all we can say is that the observed distribution must have its support on action sequences that are not truly dominated, highlighting the relevance of Theorem 1 even when population data is available.

Hence, which theorem is most appropriate is not exclusively a function of the kind of data available, but also of the assumptions that the analyst is willing to make on the data generating process. In that sense, we let the analyst be the judge of which result is the most useful for the question and data at hand.

7 Applications

We now present three applications of our framework and results. First, we show that Theorem 3 can be employed to determine the maximal probability with which an apparently dominated action sequence can be rationalized. Second, we show how a systematic investigation of deviation rules can be helpful in partially identifying a parameter in the agent's preferences. Third, we use Theorem 1 to show that set of truly dominated action sequences increases with risk aversion—the more risk averse the agent, the more action sequences can be rationalized.

7.1 Maximal probability of apparently dominated actions

If we have data on a population of agents and see many of them taking an apparently dominated action sequence, we may question whether this behavior is consistent with our Bayesian model. The machinery developed in Section 6 can be used to answer this question. We ask what is the maximum probability of taking that action sequence that can still be rationalized (since an apparently dominated action sequence can not be chosen with probability one). If the fraction of agents taking this apparently dominated action sequence exceeds the upper bound, then the Bayesian model, under the appropriate assumptions discussed in Section 6.3, is rejected.[15]

We do this in two steps. Suppose the highest probability is $\gamma \in [0, 1]$.[16] First, a lower bound is obtained for γ by constructing a specific information structure, and then an upper bound is obtained by devising a specific deviation rule. An educated guess for each side makes these bounds coincide, leading to a precise value for γ. We operationalize these ideas in the context of the two examples from the introduction.

Revisiting Example 1. We know that all three actions sequences can be rationalized. Since (*invest*, *pull back*) is apparently dominated, it cannot be rationalized with probability 1, rather with some probability $\gamma \in (0, 1)$. For the lower bound on γ, suppose we start with a uniform prior: both good and bad states are equally likely. Consider a sequential information structure that gives no information in the first period, and in the second period it gives information according to the following conditional probability system:

	g	b
good	α	$1-\alpha$
bad	0	1

When the state is good, signal g is generated with probability α and signal b is generated with probability $1-\alpha$, and when the state is bad, signal b is generated for sure. Thus, conditional on investing in the first period, the agent will choose to continue investing in the second period if the signal is g. It can be checked by applying Baye's rule that the agent will pull back upon seeing signal b if and only if $\alpha \geq \frac{2}{3}$.[17] So, assume $\alpha \geq \frac{2}{3}$.

The agent's expected payoff if he chooses to invest in the first period is given by:

$$\frac{1}{2} \cdot (-1) + \frac{1}{2} \cdot [2\alpha + (1-\alpha)(-1)] = \frac{3}{2}\alpha - 1$$

[15]This is also related to a strand in the behavioral economics literature that studies seemingly dominated dynamic choices, eg. DellaVigna and Malmendier [2006] and Grubb and Osborne [2015]. What fractions of such choices in the data set would surely reject the Bayesian model? See Section 8 for more details.

[16]Note that $\gamma = 0$ for a truly dominated action sequence (Theorem 1), and it is easy to see that it equals 1 for an action sequence that is not apparently dominated. The interesting case, that is $\gamma \in (0, 1)$, arises when the action sequence is apparently dominated but not truly dominated.

[17]Note $\mathbb{P}(\text{good}|b) = \frac{1-\alpha}{2-\alpha}$. It is optimal to choose to pull back if $2 \cdot \mathbb{P}(\text{good}|b) + (-2) \cdot \mathbb{P}(\text{bad}|b) \leq -1$, which is the case when $\alpha \geq \frac{2}{3}$.

Clearly, it is optimal to invest in the first period if $\alpha \geq \frac{2}{3}$. Finally, the probability with which the agent will choose (*invest*, *pull back*) under this information structure is

$$\frac{1}{2} \cdot 1 + \frac{1}{2} \cdot (1 - \alpha),$$

which, given the constraints on α, is maximized at $\alpha = \frac{2}{3}$. Thus, the distribution generated over the three action sequences $\{(\textit{not invest}, \emptyset), (\textit{invest}, \textit{pull back}), (\textit{invest}, \textit{invest})\}$ with $\alpha = \frac{2}{3}$ is $(0, \hat{\gamma}, 1 - \hat{\gamma})$ where $\hat{\gamma} = \frac{2}{3}$. We can conclude that $\gamma \geq \frac{2}{3}$.

Next, we characterize the upper bound on γ using deviation rules. Without loss of generality, consider a distribution of the form $(0, \hat{\gamma}, 1 - \hat{\gamma})$, and consider the following deviation rule:

	not invest	*invest & pull back*	*invest & invest*
D	*not invest*	*not invest*	*not invest*

The expression in Theorem 3 is given by:

$$\hat{\gamma} \cdot 1 + (1 - \hat{\gamma}) \cdot (-2) > 0 \quad \text{that is} \quad 3\hat{\gamma} - 2 > 0.$$[18]

Thus, any $\hat{\gamma} \geq \frac{2}{3}$ is "dominated" by this deviation rule which means that $\gamma \leq \hat{\gamma} \leq \frac{2}{3}$.

Collectively, we can conclude that the highest probability with which the action sequence (*invest*, *pull back*) can be rationalized in Example 1 is $\gamma = \frac{2}{3}$.

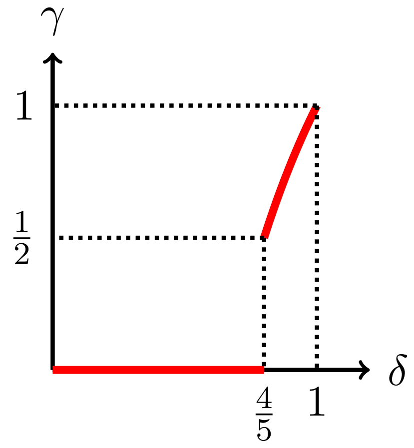
Revisiting Example 2. We are interested in the question, what is the maximum probability with which wx can rationalized as a function of δ? Call this number γ_δ. As before, we first construct a lower bound using information structures and then an upper bound using deviation rules, and choose wisely so that these coincide.

The calculations are a bit more involved for we want to report the probability as an arbitrary function of δ. It is immediate from previous discussions that $\gamma_\delta = 0$ for $\delta < \frac{4}{5}$, since wx can only be rationalized for $\delta \geq \frac{4}{5}$. It is also clear that $\gamma_\delta < 1$ for $\delta < 1$, and it exactly equal to 1 for $\delta = 1$. In the appendix, we show that in fact:

$$\gamma_\delta = \left(3 - \frac{2}{\delta}\right) \cdot \mathbb{1}\left(\delta \geq \frac{4}{5}\right),$$

where $\mathbb{1}$ is the indicator function. The adjoining figure plots this value of γ_δ. The construction of the information structure and the deviation rules that delivers $\left(3 - \frac{2}{\delta}\right)$ as the lower and upper bounds respectively are provided in the appendix.

[18]It is clear from this expression why it is without loss of generality to consider a distribution of the form $(0, \hat{\gamma}, 1 - \hat{\gamma})$. For any positive weight on $(\textit{not invest}, \emptyset)$, we can move the mass to the other two actions sequences, since the deviation maps $(\textit{not invest}, \emptyset)$ to itself and we are evaluating the highest probability with which $(\textit{invest}, \textit{pull back})$ can be rationalized.



7.2 Partial identification of preferences

Suppose that the utility function of the agent depends on a parameter $\delta \in \mathbb{D}$, known to the agent but unknown to the analyst. The set of action sequences which can be rationalized would then depend on δ. Therefore, if the analyst observes a certain action sequence being taken, she can rule out values of δ that would be inconsistent with that observation.

Let $A(\delta) \subset A$ denote the action sequences which can be rationalized for a given value of δ, and in case the analyst can observe a distribution over action sequences, let $B(\delta) \subset \Delta(A)$ denote the set of distributions over action sequences which can be rationalized for a given value of δ. Upon observing an action sequence $\mathbf{a}$, the analyst would deduce that the true value of δ must lie in the set

$$\mathbb{D}(\mathbf{a}) = \{\delta | \mathbf{a} \in A(\delta)\}$$

and, upon observing a distribution $\gamma \in \Delta(A)$ over action sequences, the analyst would deduce that the true value of δ must lie in the set

$$\mathbb{D}(\gamma) = \{\delta | \gamma \in B(\delta)\}.$$

It can be immediately noted that

$$\mathbb{D}(\mathbf{a}) \subseteq \mathbb{D}(\gamma) \ \forall \mathbf{a} \in \text{Supp}(\gamma),$$

where $\text{Supp}(\gamma)$ denotes the support of the distribution γ.

We can use deviation rules as a method of approaching the set of possible parameters, $\mathbb{D}(\mathbf{a})$ or $\mathbb{D}(\gamma)$, as follows. Given a deviation rule D, define the set $\mathbb{D}_D(\mathbf{a})$ of utility parameters for which D does *not* dominate $\mathbf{a}$, in the sense of Theorem 1. Similarly, define $\mathbb{D}_D(\gamma)$ to be the set of utility parameters for which D does *not* dominate $\mathbf{a}$, in the sense of Theorem 3. These sets, $\mathbb{D}_D(\mathbf{a})$ and $\mathbb{D}_D(\gamma)$, are typically much easier to characterize than $A(\delta)$ and $B(\delta)$ respectively. Moreover, the following result precisely characterizes the extent of identification in our framework.

Corollary 3. $\mathbb{D}(\mathbf{a}) \subset \mathbb{D}_D(\mathbf{a})$ *for any deviation rule* D*, and* $\mathbb{D}(\mathbf{a}) = \bigcap_D \mathbb{D}_D(\mathbf{a})$*. Similarly,* $\mathbb{D}(\gamma) \subset \mathbb{D}_D(\gamma)$ *for any deviation rule* D*, and* $\mathbb{D}(\gamma) = \bigcap_D \mathbb{D}_D(\gamma)$*.*

Proof. $\mathbb{D}(\mathbf{a}) \subset \mathbb{D}_D(\mathbf{a})$ by construction, and $\mathbb{D}(\mathbf{a}) = \bigcap_D \mathbb{D}_D(\mathbf{a})$ follows immediately from Theorem 1. Similarly, $\mathbb{D}(\gamma) \subset \mathbb{D}_D(\gamma)$ by construction, and $\mathbb{D}(\gamma) = \bigcap_D \mathbb{D}_D(\gamma)$ follows immediately from Theorem 3. □

This method was already used in Example 2 when we showed that the multiplicative waiting cost could not be greater than $4/5$ upon observing that the agent chose to wait, that is $\mathbb{D}(wx) = \left[\frac{4}{5}, 1\right]$, and we directly identified the set by constructing the "binding" deviation rule that recommends randomizing 50-50 in the first period instead of waiting. Similarly, building off on the argument above in Section 7.1, upon observing a distribution γ that puts weight $\gamma^w > 0$ on wx (or wy), the analyst can conclude that $\mathbb{D}(\gamma^w) = \left[\max\{\frac{2}{3-\gamma^w}, \frac{4}{5}\}, 1\right]$. So, for $0 < \gamma^w \leqslant 1/2$, we can only conclude that $\delta \geqslant \frac{4}{5}$. However, when $\gamma > 1/2$, this bound tells us that we must have $\delta \geqslant \frac{2}{3-\gamma^w}$.

7.3 Impact of risk aversion

The set of action sequences that can be rationalized is inextricably connected with the agent's utility function. It may be possible to ask how the set of action sequences that can be rationalized changes as the utility function of the agent is changed in a systematic way. Here, we show that the set of actions that can be rationalized increases with risk aversion. Thus, if we can rule out an action sequence for an agent with a utility function u, we can also rule out that action sequence for all agents who have a utility function v which is less risk averse than u.

Recall that v is less risk averse than u if and only if there exists an increasing and convex function $f : \mathbb{R} \to \mathbb{R}$ such that $v = f \circ u$. Using this fact, Weinstein [2016] and Battigalli, Cerreia-Vioglio, Maccheroni, and Marinacci [2016] show that the set of rationalizable strategies increases with risk aversion. Using Theorem 1, the same logic can be applied here.

Corollary 4. *Let* $v, u : A \times \Omega \to \mathbb{R}$ *be two utility functions, with* v *less risk averse than* u*. If* $\mathbf{a}$ *cannot be rationalized for* u*, then it cannot be rationalized for* v*.*

Proof. Let f be an increasing convex function such that $v = f \circ u$. By Theorem 1, $\mathbf{a}$ cannot be rationalized for v if and only if there exists a deviation rule $D : A \to \Delta(A)$ such that $u(D(\mathbf{b}), \omega) \geqslant u(\mathbf{b}, \omega)$ for all $\mathbf{b}$ and ω, with a strict inequality for $\mathbf{b} = \mathbf{a}$. The same deviation rule will work for v, since

$$\begin{aligned} v\left(D\left(\mathbf{b}\right),\omega\right) &= \sum_{\mathbf{b}'\in A} f\circ u(\mathbf{b}',\omega)D(\mathbf{b}'|\mathbf{b}) \\ &\geqslant f\left[\sum_{\mathbf{b}'\in A} u(\mathbf{b}',\omega)D(\mathbf{b}'|\mathbf{b})\right] \\ &= f\left(u\left(D(\mathbf{b}),\omega\right)\right) \\ &\geqslant f\left(u\left(\mathbf{b},\omega\right)\right) \\ &= v\left(\mathbf{b},\omega\right) \end{aligned}$$

where the second inequality is strict for $\mathbf{b} = \mathbf{a}$. Hence $\mathbf{a}$ is truly dominated for v, and therefore cannot be rationalized for v. □

The corollary is stated purely in terms of what actions can be rationalized. But the proof is in terms of deviation rules, and we are not aware of any proof that would work directly in the information space. Note that we used both directions of Theorem 1: first, the hard direction to show that there exists a deviation rule for u, then the easy direction to show that $\mathbf{a}$ cannot be rationalized for v.

8 Final remarks

This paper characterizes the empirical content of a standard Bayesian model for a general dynamic decision problem. Theorems 1, 2 and 3 vary the data requirement for the analyst and produce different implications for the Bayesian model, but they all rely on the same unifying feature: the idea of deviation rule, invoked each time we define dominance.

Theorem 1 gives the most parsimonious test of the model, by asking when can an action sequence be chosen by an optimizing agent. In contrast, Theorem 3 assumes that an entire distribution over action sequences can be observed, whereas Theorem 2 assumes further that states can also be observed. Put this way, it might seem that the results can be ranked, with Theorem 1 giving the weakest predictions, Theorem 3 intermediary, and Theorem 2 the strongest. However, these stronger predictions also come with stronger assumptions on the data generating process, e.g. that the signals in the population are generated independently. Because of these accompanying assumptions, we see no obvious way to rank the theorems, leaving to the analyst the task of choosing the appropriate result.

The tradition of empirically testing the Bayesian model goes at least as far back as Tversky and Kahneman [1971], Kahneman and Tversky [1972], and Grether [1980]. These tests usually involved carefully designed experiments that induced counter-intuitive predictions. In contrast, our interest here is in understanding the predictions of the model using limited field data, without any knowledge of the information that is available to agents.

Shmaya and Yariv [2016] also explore the empirical implications of the Bayesian assumption. They consider an experimental setting where the sequence of realized signals and the agent's mapping from signals to actions are observable to the analyst, but the agent has a subjective signal generating process in mind. They show that without making any assumptions on the signal generating process "anything goes"—all mappings from signal histories to actions may be optimal for a Bayesian agent. Theirs is a simple decision problem, where the agent reports the most likely state given her beliefs. We find that with richer settings the sequential decision model can indeed make predictions, even when the analyst can only observe the realized path of actions (as opposed to the entire mapping from signals to actions).

There is a long tradition in economics of recovering parameters from observed choices, most notably in the revealed preference literature (see Chambers and Echenique [2016]). While most of the literature focuses on identifying utility functions, we take them as given and ask when can choices be explained via information. Caplin and Martin [2015] provide a necessary condition to identify utility functions by rationalizing observed choices in (static) Bayesian models of decision making. The condition, called *no improving action switches*, states that no systematic reassignment of actions can lead to a higher expected utility. Caplin and Dean [2015] further show that in a stochastic choice environment this condition is necessary for an action to be rationalized when agents are allowed to acquire information through arbitrary cost functions. No improving action switches is analogous to the ideas of deviation rule and true dominance, with the difference that deviation rules include the adaptedness condition in order to respect the sequentiality of the problem.

The deviation method of proving that some behavior does not conform to the standard rational model has been used before in behavioral economics. DellaVigna and Malmendier [2006] showed with field data that gym goers switching from paying a subscription to paying per visit would save them money in expectation. Grubb and Osborne [2015] showed that cell phone users switching plans and keeping their usage the same would reduce their expected costs. Under some assumptions on their utility functions, these observations were used to argue that consumers violated the rational model. These arguments were persuasive partly because of the simplicity of the deviations used, so that a full characterization of empirical content was not necessary. We hope that, by formalizing this deviation method, new applications may arise in settings where the necessary deviations may not be as simple.

In the (axiomatic) decision theory literature, papers often seek to identify utilities and information simultaneously (see, for example, Dillenberger, Lleras, Sadowski, and Takeoka [2014], Piermont, Takeoka, and Teper [2016], and Lu [2016]). However, identifying this rich space of parameters requires much richer data as well. Frick, Iijima, and Strzalecki [2019] is an example that is close to our model. They study a dynamic random utility model, with one possibility being that the agent learns about their utility over time.

The question of which action sequences can be rationalized can also be expressed in terms of communication equilibria (see Myerson [1986] and Forges [1986], and more recently Sugaya and Wolitzky [2018]).[19] The reformulated question becomes: what action sequences can occur with positive probability in a communication equilibrium of a single-player game? Under this interpretation, our Obedience Principle (Lemma 2) is a particular case of the revelation principle of Myerson and Forges (see Propositions 1, 2 and 3 in Sugaya and Wolitzky [2018]), though our restricted context allows for a much simpler proof. Myerson introduced the notion of codominated actions, which also extends the notion of a dominated action in a static multiplayer game to a multi-stage game. Although it seems reasonable to conjecture that the codomination procedure would eliminate all truly dominated action sequences under generic payoffs, it only gives a sufficient condition for true dominance in general—there may be actions which are not codominated, but are never chosen with positive probability in any communication equilibrium.[20] For example, in Figure 3b, no actions are codominated, but the sequence of actions Lr is truly dominated.

The results on distributions, Theorems 2 and 3, can be interpreted as an application to information design (Kamenica and Gentzkow [2011], Bergemann and Morris [2016]). In this case, the question is whether it is possible for a persuader to choose a dynamic information process that incentivizes decisions by the agent to induce a specific distribution. Other papers have analyzed information design problems in a dynamic environment, notably Ely and Szydlowski [2019] and Doval and Ely [2020]. In fact, Example 1, and rationalizing an apparently dominated action sequence more generally, is related to the idea of using information as a carrot in Ely and Szydlowski [2019]. It incentivizes the choice of an action sequence using a dynamic information structure, which cannot be incentivized with any positive probability using a static one. To the best of our knowledge, our paper is the first to point out the importance of deviation rules in dynamic information design, and their role in characterizing incentive constraints in non-separable problems.

In recent work, Makris and Renou [2021] extend Bergemann and Morris [2016]'s analysis of the Bayes correlated equilibrium to a general dynamic environment. They also consider the restriction to the case of one player—a decision problem. For this, they use deviation rules to define a notion of "sure dominance", which is the analogue of our "true dominance" for the problem of experimentation (when signals can depend on actions). Their duality result is another indication that the approach used here can be applied in a variety of related contexts.

Information is also modeled as a general dynamic stochastic process in Chambers and Lambert [2021], and there too the agent takes a sequence of actions. Their focus however is on finding the right utility function in order to elicit the overall information structure. In the context of Myersonian mechanism design, Rahman [2010a,b] shows that Rochet [1987]'s

[19] Formally, in communication games, the mediator starts with no information, and only acquires information through the incentivized reports of players. But this framework can easily be extended to one where the mediator starts with information by adding to the game a *dummy player* who takes no actions, has constant payoff, but knows the state of the world since the beginning of the game. All that this dummy player does is report the state to the mediator at the beginning of the game. Alternatively, we can follow Makris and Renou [2021] who consider a mediator who directly knows the state of the world.

[20] By generic, we mean that no two action sequences can give the same payoffs in the same state. This is a strong restriction, violated by many applications of interest.

characterization of incentive compatibility can be simplified to a set of detectable deviations and adapted to include both static multidimensional and dynamic problems. The idea of deviations there and the constructive proof through a zero-sum game between the principal and agents parallels Theorem 1 here. Recently in an exercise related to ours, Deb and Renou [2021] characterize the empirical content of a model with multiple agents, binary states, and common sequential learning.

Finally, it is tempting to frame the solution to our problem in a recursive or inductive form. We approached the problem in a classical fashion, defining a general utility function, and introducing a general mapping as a deviation where all the dynamics are packed into the adaptedness condition. In the appendix, we show that a natural backward inductive approach can be useful in thinking about true dominance, and that eventually our notion of deviation rules subsumes the inductive construction. In fact, we further explain that the final stages of backward induction assume the complexity of the original problem and a naive application can lead to a mis-identification of the set of action sequences that can rationalized.

9 Appendix

9.1 Proof of Lemma 3

Statement: Let $X \subset \mathbb{R}^m$ be an evenly convex polyhedron and $Y \subset \mathbb{R}^n$ be a polytope (both non-empty). If $f : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$ is affine in each variable, then the following statements are equivalent:

1. For every $x \in X$, there exists a $y \in Y$ such that $f(x, y) > 0$;

2. There exists a $y \in Y$ such that, for every $x \in X$, $f(x, y) > 0$.

Proof. (2) $\Rightarrow$ (1) is obvious, since we can just pick, for every x, the same y that is given in (2).

We now prove (1) $\Rightarrow$ (2). Since Y is a polytope, it is the convex hull of its extreme points: $Y = conv(y_1, \ldots, y_J)$. Each j determines an affine function $f(x, y_j)$. Thus there is a $J \times n$ matrix G and a $J \times 1$ vector g such that we can write $f(x, y_j) = G_j x - g_j$ for every j.

Now, since X is an evenly convex polyhedron, it is define by a finite number of linear inequalities. Write those as $Ax \leqslant a$ and $Bx < b$. We can thus rewrite statement (1) as

1. There is no solution to the system of inequalities

$$Ax \leqslant a \quad Bx < b \quad Gx \leqslant g.$$

By Motzkin's Transposition Principle (Motzkin [1936]), this system does not have a solution if and only if there exist vectors $\alpha, \beta, \gamma \geqslant 0$ satisfying at least one of the two systems:

(a) $\alpha A + \beta B + \gamma G = 0$ and $\alpha a + \beta b + \gamma g < 0$

(b) $\alpha A + \beta B + \gamma G = 0$ and $\alpha a + \beta b + \gamma g \leqslant 0$ and $\beta \neq 0$.

Notice that if $\gamma = 0$, any of these two systems gives us a contradiction to the system of inequalities $Ax \leqslant a$ and $Bx < b$. Since we assume that X is non-empty, we must have $\gamma \neq 0$. We may normalize γ and the other variables so that $\sum_j \gamma_j = 1$. Now take any $x \in X$. From system (a), we get

$$\gamma G x = -\alpha A x - \beta B x \geqslant -\alpha a - \beta b > \gamma g$$

while from system (b) we get

$$\gamma G x = -\alpha A x - \beta B x > -\alpha a - \beta b \geqslant \gamma g.$$

Either way, we have that $\gamma\,(Gx - g) > 0$ for every $x \in X$. This finishes the proof, as can be seen by writing out the expression in its original terms: Letting $y = \sum_j \gamma_j y_j$, we have

$$f(x, y) = \sum_j \gamma_j\,(G_j x - g_j) = \gamma\,(Gx - g) > 0$$

for any $x \in X$. □

9.2 Completing the proof of Theorem 1

There are two pieces in the proof of Theorem 1, that are missing in the main text in Section 5.3, which we complete here. The first is the following claim that completes the "only if" direction of the proof, and the second fits the final step that flips the order of quantifiers to the Generalized Separation Lemma, that is, Lemma 3.

Claim 1. *Fix (π, p). If $\mathbf{a}$ is truly dominated by D and strategy σ is such that $\sigma \circ \pi \circ p(\mathbf{a}) > 0$, then the strategy $\tilde{\sigma} = D \circ \sigma$ provides a strictly higher expected payoff.*

Proof. We show that the agent strictly benefits from switching from the strategy σ to the strategy $\tilde{\sigma}$:

$$\begin{aligned}
U(\sigma,\pi,p) &= \sum_{\mathbf{b}\in A\setminus\{\mathbf{a}\}}\sum_{\omega\in\Omega} u(\mathbf{b},\omega)\,\sigma\circ\pi(\mathbf{b}|\omega)\,p(\omega) \\
&\quad + \sum_{\omega\in\Omega} u(\mathbf{a},\omega)\,\sigma\circ\pi(\mathbf{a}|\omega)\,p(\omega) \\
&< \sum_{\mathbf{b}\in A\setminus\{\mathbf{a}\}}\sum_{\omega\in\Omega} u(D(\mathbf{b}),\omega)\,\sigma\circ\pi(\mathbf{b}|\omega)\,p(\omega) \\
&\quad + \sum_{\omega\in\Omega} u(D(\mathbf{a}),\omega)\,\sigma\circ\pi(\mathbf{a}|\omega)\,p(\omega) \\
&= \sum_{\mathbf{b}\in A}\sum_{\omega\in\Omega} u(D(\mathbf{b}),\omega)\,\sigma\circ\pi(\mathbf{b}|\omega)\,p(\omega) \\
&= \sum_{\mathbf{b}\in A}\sum_{\omega\in\Omega} u(\mathbf{b},\omega)\,D\circ\sigma\circ\pi(\mathbf{b}|\omega)\,p(\omega) \\
&= \sum_{\mathbf{b}\in A}\sum_{\omega\in\Omega} u(\mathbf{b},\omega)\,\tilde{\sigma}\circ\pi(\mathbf{b}|\omega)\,p(\omega) \\
&= U(\tilde{\sigma},\pi,p)
\end{aligned}$$

Note that weak part inequality above follows from part 2 of Definition 5 (i.e. true dominance) and the strictness of it follows from part 1 of the definition. □

Claim 2. *Let* $X = \{\gamma \in \Delta(A\times\Omega)$ *s.t.* $\gamma(\mathbf{a}) > 0\}$, $Y = \{D : A \to \Delta(A)$ *s.t.* D *is adapted*$\}$, *and define* $f : X \times Y \to \mathbb{R}$ *as* $f(\gamma, D) = \mathbb{E}_\gamma[u(D(\mathbf{b}),\omega) - u(\mathbf{b},\omega)]$. *Then,*

1. *X is an evenly convex polyhedron;*
2. *Y is a polytope; and*
3. *f is an affine function in γ and D.*

Proof. 1. This follows from the definition of evenly convex polyhedron, since X is defined by a finite number of inequalities, namely

   (a) $0 \leqslant \gamma(\mathbf{a},\omega) \leqslant 1$;

   (b) $\sum_{\mathbf{a},\omega}\gamma(\mathbf{a},\omega) = 1$; and

   (c) $\sum_{\omega}\gamma(\mathbf{a},\omega) > 0$.

2. Here we consider Y as a subset of $\mathbb{R}^{A\times A}$. Note that Y is defined by a finite number of inequalities, namely

   (a) $0 \leqslant D(\mathbf{b}|\mathbf{a}) \leqslant 1$ for all $\mathbf{a},\mathbf{b} \in A$;

   (b) $\sum_{\mathbf{b}\in A} D(\mathbf{b}|\mathbf{a}) = 1$ for all $\mathbf{a} \in A$; and

(c) The adaptedness restrictions, namely

$$\sum_{b_{t+1},\ldots,b_T} D(b_1,\ldots,b_t,b_{t+1},\ldots b_T|a_1,\ldots,a_t,a_{t+1},\ldots,a_T) = \sum_{b_{t+1},\ldots,b_T} D(b_1,\ldots,b_t,b_{t+1},\ldots b_T|a_1,\ldots,a_t,c_{t+1},\ldots,c_T)$$

for all $a_1,\ldots a_T$, $b_1,\ldots b_t$, and $c_{t+1},\ldots,c_T$. Hence Y is a polyhedron. It is also bounded, as can be seen from the inequalities in (a). Therefore, by Theorem 1.1 in Ziegler [2012], Y is a polytope.

3. This follows immediately if we expand the definition of f:

$$f(\gamma, D) = \sum_{\mathbf{a},\mathbf{b},\omega} [u(\mathbf{a},\omega) - u(\mathbf{b},\omega)] D(\mathbf{a}|\mathbf{b})\gamma(\mathbf{b},\omega).$$

□

9.3 One deviation to rule them all

Proof of Proposition 2. Let $\mathcal{A}$ denote the set of truly dominated action sequences. For each $\mathbf{a} \in \mathcal{A}$, pick a deviation rule that dominates $\mathbf{a}$, and let D be a strict convex combination of all those deviation rules. Then it is straightforward to note that D simultaneously dominates all $\mathbf{a} \in \Sigma$.

For each $\mathbf{a} \notin \mathcal{A}$, take an obedient triple $(Id_A, \pi_{\mathbf{a}}, p_{\mathbf{a}})$ that rationalizes it. We can combine the prior and information structure into a single joint distribution $\gamma_{\mathbf{a}} \in \Delta(A \times \Omega)$ by defining $\gamma_{\mathbf{a}}(\mathbf{b},\omega) = \pi_{\mathbf{a}}(\mathbf{b}|\omega)\, p_{\mathbf{a}}(\omega)$. Using this notation, the statement that $(Id_A, \pi_{\mathbf{a}}, p_{\mathbf{a}})$ rationalizes $\mathbf{a}$ can be translated to the following two conditions:

1. For all $D : A \to \Delta(A)$, we have

$$\sum_{\mathbf{b},\omega} [u(\mathbf{b},\omega) - u(D(\mathbf{b}),\omega)]\, \gamma_{\mathbf{a}}(\mathbf{b},\omega) \geqslant 0$$

2. $\sum_{\omega} \gamma_{\mathbf{a}}(\mathbf{a},\omega) > 0$

Now, let $\gamma = \sum_{\mathbf{a}\notin\mathcal{A}} \mu_{\mathbf{a}}\gamma_{\mathbf{a}}$ be a strict convex combination of all the $\gamma_{\mathbf{a}}$, i.e. $\mu_{\mathbf{a}} > 0$ and $\sum_{\mathbf{a}} \mu_{\mathbf{a}} = 1$. Then it is still true that

$$\sum_{\mathbf{b},\omega} [u(\mathbf{b},\omega) - u(D(\mathbf{b}),\omega)]\, \gamma(\mathbf{b},\omega) \geqslant 0$$

for every deviation rule D. Moreover, $\sum_{\omega} \gamma(\mathbf{a},\omega) > 0$ for every $\mathbf{a} \notin \mathcal{A}$. This means that γrationalizes all $\mathbf{a} \notin \mathcal{A}$ simultaneously. That is, defining $\pi(\mathbf{b}|\omega) = \gamma(\mathbf{b}|\omega)$ and $p(\omega) = \sum_{\mathbf{b}} \gamma(\mathbf{b},\omega)$, we have that (Id_A, π, p) rationalizes all $\mathbf{a} \notin \mathcal{A}$ simultaneously.

□

9.4 Solving for the highest probability of wx in Example 2

As before, we will do this in two steps: construct a lower bound through information structures and an upper bound through deviation rules, and show these coincide.

For a lower bound, start with a uniform prior, and consider a sequential information structure that gives no information in the first period, and in the second period it gives information according to the following conditional probability system:

	a	b
X	1	0
Y	α	$1-\alpha$

So when the agent sees signal b, she's sure that the state is Y; when she sees signal a, she puts some probability greater than $\frac{1}{2}$ on the state being X (assuming $\alpha > 0$). Her optimal strategy is then to choose x after a signal of a and to choose y after a signal of b. Her expected payoff from that strategy is

$$\frac{1}{2}5\delta + \frac{1}{2}\left(\alpha 3\delta + (1-\alpha)\, 5\delta\right) = 5\delta - \alpha\delta.$$

For waiting to be optimal in the first period, we must have that $5\delta - \alpha\delta \geqslant 4$, giving us the highest possible value of α to be $\alpha^* = 5 - \frac{4}{\delta}$. Thus, the probability that the agent chooses wx under this information structure is the same as the probability that it results in a signal of a, namely

$$\frac{1}{2} + \frac{1}{2}\alpha^* = 3 - \frac{2}{\delta}.$$

To show that this is precisely the maximum probability that wx can be chosen, we use Theorem 3. Consider the following two deviation rules: D_x, which takes w to x and D_y, which takes w to y. Now, define $D_\lambda = \lambda D_x + (1-\lambda)\, D_y$, and write the gains from deviating in each state:

deviation from wx	$u\left(D\left(wx\right), X\right) - u\left(wx, X\right)$	$u\left(D\left(wx\right), Y\right) - u\left(wx, Y\right)$
D_x	$5-5\delta$	$3-3\delta$
D_y	$3-5\delta$	$5-3\delta$
D_λ	$\lambda\left(5-5\delta\right) + (1-\lambda)(3-5\delta)$	$\lambda\left(3-3\delta\right) + \left(1-\lambda\right)\left(5-3\delta\right)$

deviation from wy	$u\left(D\left(wy\right), X\right) - u\left(wy, X\right)$	$u\left(D\left(wy\right), Y\right) - u\left(wy, Y\right)$
D_x	$5-3\delta$	$3-5\delta$
D_y	$3-3\delta$	$5-5\delta$
D_λ	$\lambda\left(5-3\delta\right) + \left(1-\lambda\right)\left(3-3\delta\right)$	$\lambda\left(3-5\delta\right) + \left(1-\lambda\right)\left(5-5\delta\right)$

We would like to pick λ so that the deviation is quite favorable when the agent is choosing wx, even in the worst case state. So we choose $\lambda = \frac{1+\delta}{2}$ so that the payoffs under wx are equated. Under that deviation rule, the expected worst-case benefit of deviating is

$$\left[4-4\delta\right]\gamma\left(wx\right) + \left[4-6\delta\right]\gamma\left(wy\right)$$

which is positive precisely when

$$\gamma(wx) > \left(3 - \frac{2}{\delta}\right)(\gamma(wx) + \gamma(wy)).$$

In particular, if $\gamma(wx) > 3 - \frac{2}{\delta}$, the agent would benefit from this deviation.

9.5 A backward induction approach

As explained in Section 8, one can attempt a solution to our problem through a backward induction approach. Here we formalize this approach and comment on its significance. We begin by defining a subproblem.

Definition 11. *We refer to the collection of action sets $A_1, \ldots, A_T$, together with the utility function $u : A \times \Omega \to \mathbb{R}$, as the agent's* decision problem. *The* subproblem *obtained by fixing $(a_1, \ldots, a_t)$ is the subcollection of action sets $A_{t+1}, \ldots, A_T$, together with the utility function $v : A_{t+1} \times \cdots \times A_T \times \Omega \to \mathbb{R}$ defined by*

$$v(a_{t+1}, \ldots, a_T, \omega) = u(a_1, \ldots, a_t, a_{t+1}, \ldots, a_T, \omega).$$

Thus, a sequence $(a_{t+1}, \ldots, a_T)$ is truly dominated in the subproblem if there exists a deviation rule $D : A_{t+1} \times \cdots \times A_T \to \Delta(A_{t+1} \times \cdots \times A_T)$ such that

$$\begin{aligned} v(D(a_{t+1}, \ldots, a_T), \omega) &> v(a_{t+1}, \ldots, a_T, \omega) \quad \text{for all } \omega \in \Omega \text{ and} \\ v(D(b_{t+1}, \ldots, b_T), \omega) &\geqslant v(b_{t+1}, \ldots, b_T, \omega) \quad \text{for all } b_{t+1} \in A_{t+1}, \ldots, b_T \in A_T, \omega \in \Omega. \end{aligned}$$

The following simple result shows how we can derive conclusions about a decision problem by looking at particular subproblems.

Proposition 2. *If $(a_{t+1}, \ldots, a_T)$ is truly dominated in the subproblem obtained by fixing $(a_1, \ldots, a_t)$, then $(a_1, \ldots, a_T)$ is truly dominated in the original problem.*

Proposition 2 gives a method of finding truly dominated sequences by backward induction. We first fix $(a_1, \ldots, a_{T-1})$ and then find which actions a_T are truly dominated in the single-period problem that follows.[21] Let $\tilde{A}_T$ be the last period actions that survived (that is, can be rationalized), and now fixing $(a_1, \ldots, a_{T-2})$ we find which sequences $(a_{T-1}, a_T) \in A_{T-1} \times \tilde{A}_T$ are truly dominated in this two-period problem, and so on. This exercise helps the analyst in two ways. First, it directly simplifies her search for the set of action sequences that cannot be rationalized, and second, intuitively speaking, it informs her that the construction of deviation rules for other action sequences should not take these truly dominated action sequences in their support. For example, if action sequence if the (a_{T-1}, a_T) is truly dominated in the subproblem, the analyst immediately knows that whole action sequence $\mathbf{a}$ is truly dominated

[21] Since this is a "static" problem, it is the same as looking for actions which are strictly dominated by some other action.

in the original problem, and moreover, that in order to construct a deviation rule for any other action sequence **b** that may be truly dominated, the associated deviation rule does not have to put any weight on **a**.

It is straightforward to observe that the idea of deviation rule subsumes this type of induction arguments– they are implicit in the construction of a deviation rule used to establish true dominance.

There are two further caveats to making backward induction the primary approach in solving our problem. First, in its final stages the method described above can be almost as complex as the original problem. Second, a naive application of it might lead to misidentification of the set of action sequences that can be rationalized, as the following conjecture exposits:

Conjecture 1. *Suppose that (i) the sequence of actions* $(a_1, \ldots, a_t)$ *can be chosen with positive probability, and (ii)* $(a_{t+1}, \ldots, a_T)$ *can be rationalized in the subproblem obtained by fixing* $(a_1, \ldots, a_t)$*. Then,* $(a_1, \ldots, a_T)$ *can be rationalized.*

This conjecture is false; the decision tree in Figure 3b is a simple counterexample. Both L and R can be chosen in the first period, and in the decision that follows action L, both l and r can be chosen. This could lead the analyst to believe that (L, r) can rationalized. However, the deviation rule depicted in the figure shows otherwise. The problem with naive inductive reasoning is that a choice of L can only be rationalized if the agent is sure about the state being the first one; choosing r would then require an inconsistent belief. Such indifference in payoffs (both (L, l) and (R, r) yield 3 in the first state), requires a comparison of the full sequence of actions. Thus, in general, it is apt to define true dominance along the entire sequence of actions, and employ the induction argument to construct simple deviation rules whenever possible.

References

P. Battigalli, S. Cerreia-Vioglio, F. Maccheroni, and M. Marinacci. A note on comparative ambiguity aversion and justifiability. *Econometrica*, 84(5):1903–1916, 2016.

D. Bergemann and S. Morris. Bayes correlated equilibrium and the comparison of information structures in games. *Theoretical Economics*, 11(2):487–522, 2016.

D. Bergemann and S. Morris. Information design: a unified perspective. *Theoretical Economics*, 57(1):44–95, 2019.

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

A. Caplin and D. Martin. A testable theory of imperfect perception. *The Economic Journal*, 125:184–202, 2015.

C. P. Chambers and F. Echenique. *Revealed preference theory.* Cambridge University Press, 2016.

C. P. Chambers and N. S. Lambert. Dynamic belief elicitation. *Econometrica*, 89(1):375–414, 2021.

H. de Oliveira. Blackwell's informativeness theorem using diagrams. *Games and Economic Behavior*, 109:126–131, 2018.

R. Deb and L. Renou. Dynamic choice and common learning. University of Toronto and Queen's Mary University of London, 2021.

S. DellaVigna and U. Malmendier. Paying not to go to the gym. *American Economic Review*, 96(3):694–719, 2006.

D. Dillenberger, J. S. Lleras, P. Sadowski, and N. Takeoka. A theory of subjective learning. *Journal of Economic Theory*, 153(1):287–312, 2014.

L. Doval and J. Ely. Sequential information design. *Econometrica*, 88(6):2575–2608, 2020.

J. C. Ely and M. Szydlowski. Moving the goalposts. *Journal of Political Economy*, 128(2): 468—506, 2019.

F. Forges. An approach to communication equilibria. *Econometrica*, 54(6):1375–1385, 1986.

M. Frick, R. Iijima, and T. Strzalecki. A note on comparative ambiguity aversion and justifiability. *Econometrica*, 87(6):1941–2002, 2019.

D. M. Grether. Bayes rule as a descriptive model: The representativeness heuristic. *Quarterly Journal of Economics*, 95(3):537–557, 1980.

M. Grubb and M. Osborne. Cellular service demand: Biased beliefs, learning, and bill shock. *American Economic Review*, 105(1):234–271, 2015.

D. Kahneman and A. Tversky. Subjective probability: A judgment of representativeness. *Cognitive Psychology*, 3(3):430–454, 1972.

E. Kamenica. Bayesian persuasion and information design. *Annual Review of Economics*, 11: 249–272, 2019.

E. Kamenica and M. Gentzkow. Bayesian persuasion. *American Economic Review*, 101(6): 2590–2615, 2011.

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

M. Makris and L. Renou. Information design in multi-stage games. University of Kent and Queen Mary University of London, 2021.

T. Motzkin. Beiträge zur theorie der linearen ungleichungen. *PhD thesis* Azriel, Jerusalem, 1936.

R. Myerson. Multistage games with communication. *Econometrica*, 54(2):323–358, 1986.

D. Pearce. Rationalizable strategic behavior and the problem of perfection. *Econometrica*, 52 (4):1029–1050, 1984.

E. Piermont, N. Takeoka, and R. Teper. Learning the Krepsian state: Exploration through consumption. *Games and Economic Behavior*, 100:69–94, 2016.

D. Rahman. Dynamic implementation. University of Minnesota, 2010a.

D. Rahman. Detecting profitable deviations. University of Minnesota, 2010b.

J.-C. Rochet. A necessary and sufficient condition for rationalizability in a quasi-linear context. *Journal of Mathematical Economics*, 16(2):191–200, 1987.

E. Shmaya and L. Yariv. Experiments on decisions under uncertainty: A theoretical framework. *American Economic Review*, 106(7):1775–1801, 2016.

T. Sugaya and A. Wolitzky. The revelation principle in multistage games. Stanford University and Massachusetts Institute of Technology, 2018.

A. Tversky and D. Kahneman. Belief in the law of small numbers. *Psychological Bulletin*, 76 (2):105–110, 1971.

A. Wald. Statistical decision functions. *Annals of Mathematical Statistics*, 20(2):165–205, 1949.

J. Weinstein. The effect of changes in risk attitude on strategic behavior. *Econometrica*, 84 (5):1881–1902, 2016.

G. M. Ziegler. *Lectures on polytopes*, volume 152. Springer Graduate texts in Mathematics, 2012.