

Rationalizing dynamic choices*

Henrique de Oliveira

Rohit Lamba

March 2019

Abstract

Consider an analyst who observes an agent taking a sequence of actions. The analyst ponders whether the sequence of actions observed could have been taken by a rational, Bayesian agent. Although the analyst observes the chosen actions, he does not have direct access to the agent's information and must therefore consider a multitude of possibilities. Could *some* gradual release of information have led the agent to *optimally* take that sequence of actions?

We show that a sequence of actions cannot be rationalized by any information structure if and only if it can be proved to be *dominated* via a *deviation argument*. This argument prescribes a way of deviating that would leave the agent better off in any possible scenario, regardless of the information she might have. As an application of this characterization, we show that an increase in the agent's risk-aversion leads to less predictive power—more sequences of actions can be rationalized. We also show results that simplify the analyst's search for a deviation argument and demonstrate how these arguments can be used to partially identify utility parameters without making assumptions on the agent's information.

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 actions which can be rationalized by such a criterion.

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

*de Oliveira: Penn State University, henrique@psu.edu; 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, 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, and Pennsylvania Economic Theory Conference for their comments.

of the world, and then chooses an action a_1 , following which the agent sees another informative signal s_2 , chooses an action a_2 and so on, until the agent chooses a final action a_T . A terminal payoff is then realized, represented by an arbitrary function $u : A \times \Omega \rightarrow \mathbb{R}$, where A is the set of all action sequences.

An analyst knows the utility function u , and observes the chosen sequence of actions (a_1, \dots, a_T) . The analyst does not know the agent's prior $p \in \Delta(\Omega)$ nor her signal generating process $\pi : \Omega \rightarrow \Delta(S)$, where S is the set of all sequences of signals. The analyst asks: does there exist any p and (π, S) for which the chosen sequence of actions (a_1, \dots, a_T) could have been optimal?

To understand the setting and what this criterion 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 final payoffs 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, 0*), 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 the 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 $0 < \delta < 1$. Payoffs are as follows:

¹Generally, 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. 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 $\delta \geq 4/5$, 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 $\delta \geq 4/5$.

The full solution for Example 2 was obtained in two steps: for $\delta \geq 4/5$, we constructed one information structure under which the firm finds it optimal to choose w ; for $\delta < 4/5$, we showed that the firm would not choose w because deviating to an immediate 50-50 bet on x and y gives a higher expected payoff independent of the information that might have arrived in the next period. Alternatively, we could have shown that w cannot be chosen when $\delta < 4/5$ by showing that the information structure we constructed is the most favorable towards choosing w among all possible information structures.

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, we have to show 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 actions to lotteries over actions, $D : A \rightarrow \Delta(A)$. 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, 0)*, then adaptedness demands that we have to map *(invest, invest)*

also to $(not\ invest, \emptyset)$. As a result, the deviation is not *uniformly* better, and so $(invest, pull\ back)$ can in fact be rationalized. 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 wy cannot be rationalized if $\delta < 4/5$.

We say that a deviation rule uniformly improves upon an action sequence if it strictly increases payoff along that action sequence without worsening payoffs elsewhere on the decision tree. If an action sequence can be uniformly improved upon by a deviation rule then it is *truly dominated*. In Example 1, the action sequence $(invest, pull\ back)$ cannot be uniformly improved upon and hence is not truly dominated, whereas in Example 2 the action sequences wx and wy can be uniformly improved upon by the deviation rule described above and hence are truly dominated. Our main result establishes the following equivalence:

Theorem. 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.

What is the static counterpart of the theorem? The agent takes only one action and then payoffs are realized. 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, the rule would recommend deviating from the dominated action to the dominating mixed strategy. In this sense, our result is a dynamic generalization of the Wald-Pearce Lemma.

The concept of deviation rules is bereft of information since it must work for all sequential information structures. For any strategy of the agent $\sigma : S \rightarrow \Delta(A)$, the composition mapping $D \circ \sigma$ uniformly improves upon the action sequence that cannot be rationalized.² Moreover, the adaptedness property condenses all potential inductive arguments on how deviations from the agent's strategy can be constructed that respect the sequentiality of the problem.

To determine whether an action sequence can be rationalized, one can investigate directly, using the definition, or indirectly, using deviation rules. Which solution is simpler may depend on the particular problem, but in Section 6 we show three results that together drastically simplify

²Note that $D \circ \sigma$ refers to the strategy that is obtained by first deciding what an agent following σ would have done, and then deviating from that as prescribed by D .

the task of looking for deviation rules. The first result formalizes an intuitive process of backward induction. The second result restricts the search to deviation rules that partition action sequences in two classes: repulsive and absorbing. The agent is never recommended to deviate away from absorbing sequences, and always recommended to deviate away from repulsive sequences. In particular, "chains of deviations" are unnecessary. Finally, we show that it is possible to find a single deviation rule that simultaneously dominates every truly dominated action sequence.

There are two ways of thinking about the applicability of the framework. First, it provides a minimal test of Bayesian rationality. Could investing in the first round only in Example 1 or waiting in Example 2 be justified by any possible learning? And second, assuming Bayesian rationality, it can help the analyst (or the econometrician) partially identify missing pieces of information from the agent's preferences. For instance, the firm's choice to wait in Example 2 helps identify the cost of waiting to be $\delta \geq \frac{4}{5}$. In fact, in more detailed examples, the notion of true dominance restricts an unknown parameter to satisfy a system of inequalities which encloses it in smaller set of possibilities.

Finally, we apply our main theorem to 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. Using our theorem and an elegant observation by Weinstein [2016], the proof of this result is straightforward.

2 Model and definitions

2.1 Notation

A *stochastic mapping* from X to a finite set Y is a function $\alpha : X \rightarrow \Delta(Y)$, 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 $\alpha : X \rightarrow \Delta(Y)$ and $\beta : Y \rightarrow \Delta(Z)$ 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 $\beta : Y \rightarrow \Delta(Z)$, 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 (α, β) .

For a real-valued function $u : Y \rightarrow \mathbb{R}$ 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 $u(\cdot)$ under the distribution α .

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

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

with elements $\mathbf{x}^t \in X^t$ and $\mathbf{x} \in X$. Finally, a stochastic map $\alpha : X \rightarrow \Delta(Y)$ is said to be *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 Agent's problem

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 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 $\pi : \Omega \rightarrow \Delta(S)$.³

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 $\sigma : S \rightarrow \Delta(A)$.⁴

³We 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 $\pi_1 : \Omega \rightarrow \Delta(S_1)$, and $\pi_t : \Omega \times S^{t-1} \rightarrow \Delta(S_t) \forall 2 \leq t \leq T$. 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].

⁴As 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 $\sigma_1 : S_1 \rightarrow \Delta(A_1)$, and $\sigma_t : S^t \times A^{t-1} \rightarrow \Delta(A_t) \forall 2 \leq t \leq T$. It is possible to deduce one formulation from the other.

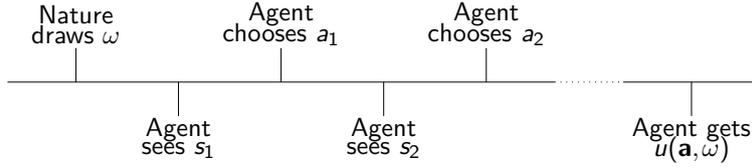


Figure 1: The timeline of signals and actions

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 $\sigma \circ \pi(\mathbf{a}|\omega)$. 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 . We say that an action sequence can be *rationalized* if it can be chosen with positive probability by an optimizing agent for 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. $\sigma \circ \pi \circ p(\mathbf{a}) > 0$.⁵

This definition is permissive in the sense that an action sequence is considered to be 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.⁶

To deduce that an action sequence cannot be rationalized, the analyst 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.

⁵Here $\sigma \circ \pi \circ p(\mathbf{a}) = \sum_{\omega} \sigma \circ \pi(\mathbf{a}|\omega) p(\omega)$ (see Section 2.1).

⁶This 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.

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

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

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 in order 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 $\sigma(a) = 1$.*

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.

3.1 The Wald-Pearce Lemma

An elegant 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.*

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, positivity of the left hand side is equivalent to a not being rationalized and 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.

3.2 A generalization of Wald-Pearce

The inversion of quantifiers result that we will need can be interpreted as an intuitive set-generalization of the Wald-Pearce Lemma, as follows. Consider a two-player game with finitely many strategies, where player 1 chooses $a \in A$ and player 2 chooses $b \in B$. Let player 1's payoffs be given by $u(a, b)$. We first generalize the definitions invoked in the Wald-Pearce Lemma.

Definition 4. *Let $\tilde{B} \subset B$.*

1. $a \in A$ is a best-response for \tilde{B} if there exists $\beta \in \Delta(B)$ such that $\beta(\tilde{B}) > 0$ and $u(a, \beta) \geq u(a', \beta) \forall a' \in A$.
2. $a \in A$ is strictly dominated at \tilde{B} if there exists $\alpha \in \Delta(A)$ such that $u(\alpha, b) \geq u(a, b) \forall b \in \tilde{B}$, with a strict inequality $\forall b \in \tilde{B}$.

In this language, saying that a is a best-response for the entire set B is the same as saying that a is a best-response to some belief, and saying that a is strictly dominated at the entire set B

just means that a is strictly dominated. Hence, these definitions generalize the ones used in the Wald-Pearce Lemma.

Now, if a is never a best-response for a subset $\tilde{B} \subset B$, it may still be that a is a best response to some belief on B . But in that case that belief must put probability zero on the set \tilde{B} for even a tiny probability of player 2 choosing an action from \tilde{B} would invalidate a as a possible choice. The only way this can happen is if a is strictly dominated at \tilde{B} . The generalization of the Wald-Pearce Lemma thus follows.

Lemma 2. *a is never a best-response for \tilde{B} if and only if a is strictly dominated at \tilde{B} .*

The min-max formulation of Lemma 2 would be:

$$\inf_{\beta: \beta(\tilde{B}) > 0} \max_{\alpha} \mathbb{E}[u(\alpha, \beta) - u(a, \beta)] = \max_{\alpha} \inf_{\beta: \beta(\tilde{B}) > 0} \mathbb{E}[u(\alpha, \beta) - u(a, \beta)].$$

The value of the above expression is in fact always zero. Unlike the Wald Pearce Lemma and its application to Corollary 1, the min-max approach to prove Lemma 2 does not work. Because of the strict inequality under "inf" the set of β s under consideration is not compact, and hence the min-max theorem cannot be applied. The result however is still true. A separating hyperplane argument (using a generalization of Farkas's lemma due to Motzkin) is invoked to prove it, details of which are provided in the appendix.

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

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

Perhaps unsurprisingly, 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} . \square

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 $a_1 = \textit{invest}$ and $a_2 = \textit{pull back}$ is apparently dominated by the action sequence $a_1 = \textit{not invest}$ and $a_2 = \emptyset$. Yet it is easy to construct an information structure where it will be optimal for the agent to choose $(\textit{invest}, \textit{pull back})$ with positive probability: 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.

Notice that the apparent dominance of $(\textit{invest}, \textit{pull back})$ can be established simply by comparing its payoffs with that of $\textit{not invest}$. The payoffs for $(\textit{invest}, \textit{invest})$ 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 $(\textit{invest}, \textit{pull back})$, the fact that the agent *could* have ended up choosing $(\textit{invest}, \textit{invest})$ makes those payoffs relevant.

Therefore, we need more than just apparent dominance for it to be impossible for an action sequence to be rationalized. In addition to improving upon the action sequence under consideration, that "more" also 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 $D : A \rightarrow \Delta(A)$, 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 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 6. *A deviation rule $D : A \rightarrow \Delta(A)$ dominates an action sequence \mathbf{a} if*

1. $u(D(\mathbf{a}), \omega) > u(\mathbf{a}, \omega)$ for all $\omega \in \Omega$.
2. $u(D(\mathbf{b}), \omega) \geq u(\mathbf{b}, \omega)$ for all $\mathbf{b} \in A$ and $\omega \in \Omega$.

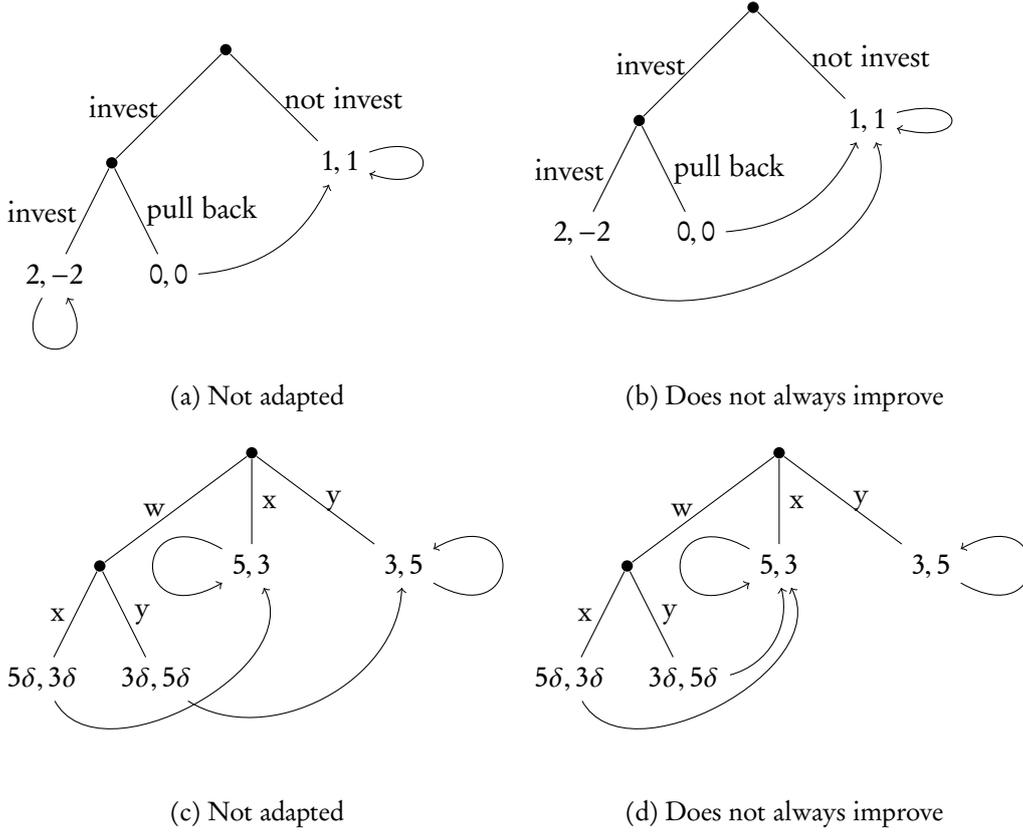


Figure 2: Deviation rules for Examples 1 and 2

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

Notice that there's no visible time dimension in the definition above; time is implicit in the condition that D must be adapted (see Section 4.3 for examples). For $T = 1$, the same definition applies, but the condition that D is adapted becomes vacuous. In that case, if \mathbf{a} is strictly dominated by α , we can define a deviation rule D_α which takes \mathbf{a} to α and does not change any other actions. D_α then dominates \mathbf{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. While the second condition and the embedded notion of adaptedness in the definition have no bearing when $T = 1$, they impose meaningful restrictions when $T > 1$, encapsulating the distinction between true dominance 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. As in a (single player) extensive-form game, the sequences of actions can be depicted as a decision tree. Each complete sequence of actions corresponds to a terminal node. Thus any mapping from sequences of actions into sequences of actions can be

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 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 no longer leads to a uniform improvement in payoffs; it violates part 2 of Definition 6.

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 the same in the second period, choose 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. It may be described as the prescription: "Never choose L; always switch to R instead. If you were going to choose L, switch your second-period choice; if you were going to choose R, stick to your second-period choice". Analogously, Lr is shown to be truly dominated in 3b by the same deviation rule, but note that here Ll is not truly dominated. We will discuss further the significance of the difference between the two examples in Figure 3 in Section 6.

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.*

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.

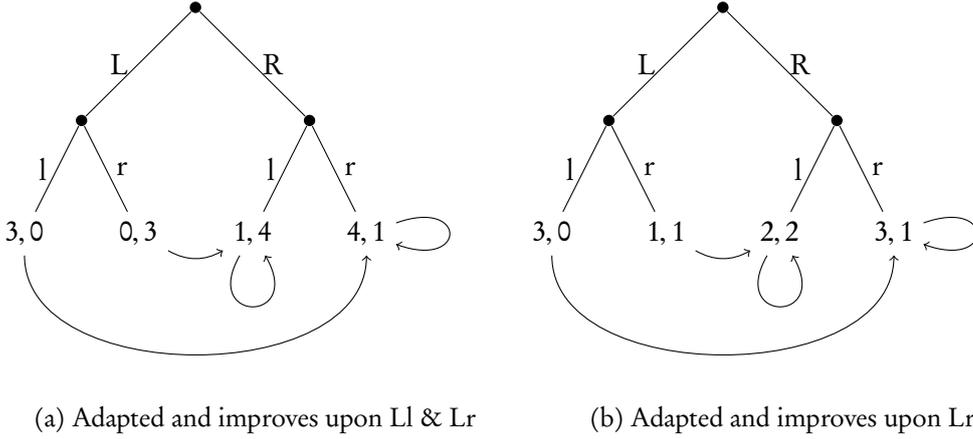


Figure 3: Deviation rules with history dependence

The steps involved in establishing this result are divided into two 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 translate the problem into a game between the analyst and a coalition between the agent and nature, and then use the generalization of the Wald-Pearce Lemma to invert quantifiers.

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 7. (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 always 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 3 (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. 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 $D : A \rightarrow \Delta(A)$ 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. \square

5.2 Proof of Theorem 1

The "only if" direction: if \mathbf{a} is truly dominated, it cannot be rationalized. Let D be a deviation rule which shows that \mathbf{a} is truly dominated. We show that any strategy that plays \mathbf{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}, ω) 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}, ω) , this difference is non-negative, with strict inequality for $\mathbf{b} = \mathbf{a}$. Hence if γ puts positive probability on \mathbf{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 \mathbf{a} cannot be rationalized, it is truly dominated. Given an action sequence \mathbf{a} which cannot be rationalized, we must find a deviation rule D which shows that it is truly dominated. Letting $\Gamma(\mathbf{a}) = \{(\sigma, p, \pi) \mid \sigma \circ \pi \circ p(\mathbf{a}) > 0\}$, we can write the statement " \mathbf{a} cannot be rationalized" as

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

By the Obedience Principle (Lemma 3), the statement " \mathbf{a} cannot be rationalized" is equivalent to the statement " \mathbf{a} cannot be rationalized by an obedient triple". This means that we can, without loss, restrict attention to $\pi : A \rightarrow A$ 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 $D : A \rightarrow \Delta A$. Incorporating these, we get the equivalent statement

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

Our goal is to switch the order of quantifiers in this statement, which would produce the

strategy $\hat{\sigma}$ and hence 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, outlined previously. 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 wouldn't 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.

To make the objective function linear requires only 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 \mathbf{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 (D, γ) .

In the spirit of Section 3.2, to get around the other problems, we construct an auxiliary game with the eventual purpose of appealing to Lemma 2. Consider a finite game where Player 1 chooses mixed strategy D and Player 2 chooses the mixed strategy γ , where the pure strategy counterparts are pure deviation rules and elements of $A \times \Omega$ respectively.⁷ Showing that the mixed strategies in this game are indeed deviation rules is analogous to showing the equivalence between mixed and behavioral strategies in extensive form games. The proof of this is presented in Lemma 5 in the appendix.

Player 1's expected payoff from a mixed strategy profile is the expression above. We can interpret this as a zero-sum game of the analyst against nature, where nature chooses both the state and a sequence of actions of the agent, and the analyst chooses a deviation rule.

Now, the statement that \mathbf{a} cannot be rationalized can be rewritten as

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

Note that 0 is the payoff of the strategy $D = Id_A$, which is available to Player 1. Hence, the statement above means that the strategy Id_A is never a best-response for the set $\{\mathbf{a}\} \times \Omega$ (see Definition 4). By Lemma 2, Player 1 must have available some strategy D^* which strictly dominates Id_A at the set $\{\mathbf{a}\} \times \Omega$. This is equivalent to saying that D^* is a deviation rule which dominates \mathbf{a} .

⁷A pure deviation rule is an adapted mapping $D : A \rightarrow A$, where adaptedness requires that for $\mathbf{b} = D(\mathbf{a})$, (b_1, \dots, b_t) can only be a function of (a_1, \dots, a_t) for all $t = 1, 2, \dots, T$.

6 Relevant deviations rules

Suppose we are given a sequence of actions \mathbf{a} and we want to find a deviation rule that dominates it, if that's possible. Searching within all adapted mappings $D : A \rightarrow \Delta(A)$ might be a daunting task, but in practice there are many ways of making this search easier. Here we discuss three results that facilitate this search.

6.1 Backward induction

It is tempting to frame the solution to our problem in a recursive or inductive form. Here 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. The following simple result shows how we can derive conclusions about a decision problem by looking at particular subproblems.

Proposition 2 (Informal). *If (a_{t+1}, \dots, a_T) is truly dominated in the subproblem obtained by fixing (a_1, \dots, a_t) , then (a_1, \dots, 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, \dots, a_{T-1}) and then find which actions a_T are truly dominated in the single-period problem that follows.⁸ Let \tilde{A}_T be the last period actions that survived (that is, can be rationalized), and now fixing (a_1, \dots, 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, as we will prove in the next subsection, 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 (a_{T-1}, a_T) is truly dominated in the subproblem, the analyst immediately knows that whole action sequence \mathbf{a} is truly dominated in the original problem, and moreover, that in order to construct a deviation rule for any other action sequence \mathbf{b} that may be truly dominated, the associated deviation rule does not have to put any weight on \mathbf{a} .

There are however two 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 mis-identification of the set of action sequences that can be rationalized, as the following conjecture exposes:

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

⁸Since this is a "static" problem, it is the same as looking for actions which are strictly dominated by some other action.

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 be 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.

6.2 Simple deviations

Is there a systematic way of ruling out poor candidate deviation rules? Intuitively speaking, if \mathbf{a} is truly dominated, it seems futile to recommend deviating to the tree which contains \mathbf{a} . Moreover, it also seems unnecessary to form chains of deviations: if it makes sense to deviate towards some action sequence \mathbf{b} , then perhaps it also makes sense not to deviate away from it. Here we formalize these intuitions by thinking of a deviation rule as a Markov chain. Recollect that for $D(\mathbf{b}) \in \Delta(A)$, we refer to $D(\mathbf{a}|\mathbf{b}) \in [0, 1]$ as the weight put on \mathbf{a} by the probability distribution $D(\mathbf{b})$.

Definition 8. *Given a deviation rule D , we say that*

1. \mathbf{a} is *repulsive* if $D(\mathbf{a}|\mathbf{b}) = 0$ for all $\mathbf{b} \in A$;
2. \mathbf{a} is *absorbing* if $D(\mathbf{a}|\mathbf{a}) = 1$.

The term *absorbing* is directly borrowed from the Markov chains taxonomy, and the notion of *repulsiveness* is closely related to the idea of accessibility. In Markov chains, we say a "state" \mathbf{a} is *accessible* from state \mathbf{b} in n steps if $D^n(\mathbf{a}|\mathbf{b}) > 0$, where D^n represents the application of the deviation rule n times. The "state" \mathbf{a} is then said to be *inaccessible* from \mathbf{b} if $D^n(\mathbf{a}|\mathbf{b}) = 0$ for all n . It is easy to see that \mathbf{a} is repulsive if and only if it is inaccessible from all other "states" \mathbf{b} . We can now operationalize this terminology to prune the decision tree of truly dominated action sequences.

Definition 9. *A deviation rule D removes (a_1, \dots, a_t) if*

1. \mathbf{b} is *repulsive* whenever $(b_1, \dots, b_t) = (a_1, \dots, a_t)$, and
2. \mathbf{b} is *absorbing* whenever $(b_1, \dots, b_t) \neq (a_1, \dots, a_t)$.

Theorem 2. *If $\mathbf{a} = (a_1, \dots, a_T)$ is a truly dominated action sequence, then there exists a deviation rule D that dominates \mathbf{a} and removes (a_1, \dots, a_t) for some $t \geq 1$.*

As a thought experiment, let us apply this result to Example 1 from the introduction. Suppose we're trying to show that $(invest, pull\ back)$ is truly dominated. There are a total of 15 pure decision rules to consider, and in principle we would need to look for a dominating deviation rule among all their mixtures.⁹ Using the theorem above, we can simplify this search dramatically. First, suppose that D removes $a_1 = invest$. There is a single deviation rule that does that, namely the one that always deviates to $not\ invest$ (see Figure 4a). Now, suppose that D removes $(a_1, a_2) = (invest, pull\ back)$. Then $(invest, invest)$ will be absorbing, and by adaptedness it is the only candidate for a deviation from $(invest, pull\ back)$. Therefore, we have again a single candidate for D , namely the one that recommends deviating from $(invest, pull\ back)$ to $(invest, invest)$ and does not recommend any other deviation (see Figure 4b). Thus the theorem tells us that we only have to consider two deviations.

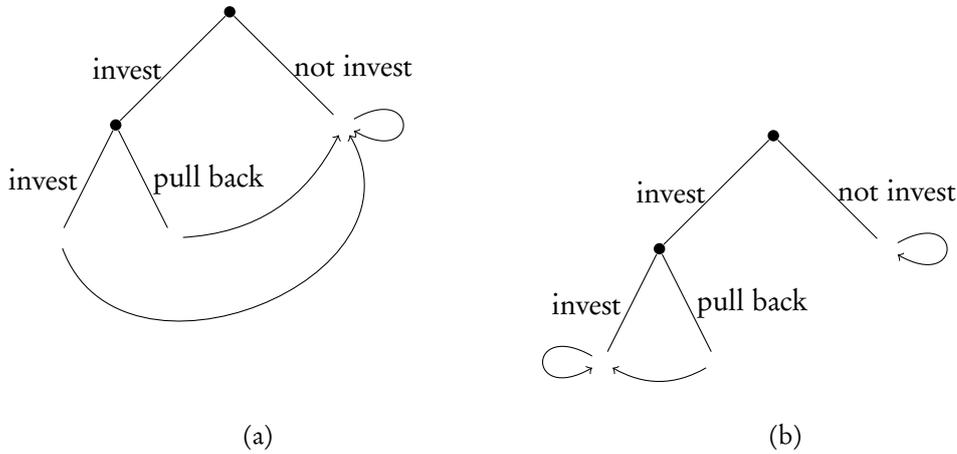


Figure 4: Relevant deviations for the investment example

6.3 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 in fact easy to find a single deviation rule that simultaneously dominates every truly dominated action sequence. This can be done rather simply by defining a new deviation rule which is a strict convex combination of the ones found for each truly dominated action sequence. Even more, it is possible to find a deviation rule that not only dominates every truly dominated action sequence, but also never recommends playing an action sequence that is truly dominated. We can formalize these claims as follows.

Proposition 3. *There exists a deviation rule D such that*

1. *every truly dominated action sequence is repulsive, and*
2. *every truly dominated action sequence is dominated by D .*

⁹There are 3 action sequences so a total of 27 possible (pure) deviation mappings, but not all of them are adapted. It can be checked that exactly 15 combinations are possible for mappings that are adapted.

7 Applications

We now present two applications of our framework and results. First, 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. Second, we modify Example 2 to show how a systematic investigation of deviation rules can be helpful to partially identify a parameter in the agent’s preferences.

7.1 Impact of risk aversion

The set of action sequences that can be rationalized is inextricably connected with the agent’s utility function. In our earlier examples, we interpreted the payoffs as objectively given. 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 which 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 $f : \mathbb{R} \rightarrow \mathbb{R}$ such that $v = f \circ u$. Using this fact, Weinstein [2016] shows that the set of rationalizable strategies increases with risk aversion. Using Theorem 1, the same logic can be applied here.

Corollary 2. *Let $v, u : A \times \Omega \rightarrow \mathbb{R}$ 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 $D : A \rightarrow \Delta(A)$ such that $u(D(\mathbf{b}), \omega) \geq u(\mathbf{b}, \omega)$ for all \mathbf{b} and ω , with a strict inequality for $\mathbf{b} = \mathbf{a}$. The same deviation rule will work for v , since

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

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 . \square

Note that 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 our theorem: first, the hard direction to show that there exists a deviation rule for μ , then the easy direction to show that \mathbf{a} cannot be rationalized for ν .

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 $A(\delta)$ denote the 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

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

We can use deviation rules as a method of approaching the set $\mathbb{D}(\mathbf{a})$ of possible parameters, as follows. Given a deviation rule D , we can define the set $\mathbb{D}_D(\mathbf{a})$ of utility parameters for which D does *not* dominate \mathbf{a} . This set (or its complement) is typically much easier to characterize than $A(\delta)$. Moreover, the following result precisely characterizes the extent of identification in our permissible framework.

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

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. \square

This method was already used in Example 2 when we showed that the cost could not be greater than $4/5$ upon observing that the agent chose to wait. The only deviation rule needed there was the one that recommended randomizing 50-50 in the first period instead of waiting. Although it was fairly straightforward to find the relevant deviation rule in that example, in some cases a more a systematic analysis of deviation rules may be needed.

Consider, for instance, a modified version of Example 2 which adds a second project, and where the decision to wait is split into two possibilities:

| | x | y | x' | y' | wx | wy | $w'x'$ | $w'y'$ |
|---|-----|-----|------|------|-----------|-----------|-----------|-----------|
| X | 5 | 3 | 6 | 0 | 5δ | 3δ | 6δ | 0 |
| Y | 3 | 5 | 0 | 6 | 3δ | 5δ | 0 | 6δ |

Here, the agent can choose to invest in one of two projects in the first period: one with outcomes (x, y) and another with outcomes (x', y') , or choose to wait and select which one of the two projects to keep available (w or w'). A total of eight action sequences are possible.

Suppose that the analyst observes that the agent has chosen wx . Thus, we can rule out values of δ such that there exists a deviation rule that dominates wx . By Theorem 2, we know that we can restrict attention to deviation rules that *remove* wx or w . There is a single pure deviation rule that removes wx —taking it to wy —and there are six pure deviation rules that remove w : two that take w to x or y , and four that take w to w' . Theorem 2 says that we can, without loss of generality, restrict attention to deviation rules that are mixtures of these seven.

The deviation rule of Example 2 would still tell us that we cannot have $\delta < 4/5$: take wx and wy to a 50-50 bet on x and y and keep all the other choices unchanged. But when $\delta \geq 4/5$, the argument we used to establish that wx could be rationalized relied on the information structure which conveyed full information in the second period. If the agent expects to have full information in the second period, w' would now be superior to w . This indicates that deviation rules which take w to w' could be relevant.

So we consider two deviation rules:

| a | wx | wy |
|----------|-------------------------------|-------------------------------|
| D_1 | $\frac{1}{2}x + \frac{1}{2}y$ | $\frac{1}{2}x + \frac{1}{2}y$ |
| D_2 | $w'x'$ | $w'y'$ |

where it is to be understood that both deviation rules keep action sequences other than wx and wy unchanged. We can write down the change in payoff for each deviation rule $u(D(\mathbf{a}), \omega) - u(\mathbf{a}, \omega)$ for each state:¹⁰

| ω | X | | Y | |
|---------------------------------|--------------------------------|---------------------|---------------------|--------------------------------|
| a | wx | wy | wx | wy |
| No Deviation | 0 | 0 | 0 | 0 |
| D_1 | $4 - 5\delta$ | $4 - 3\delta$ | $4 - 3\delta$ | $4 - 5\delta$ |
| D_2 | δ | -3δ | -3δ | δ |
| $\alpha D_1 + (1 - \alpha) D_2$ | $\alpha(4 - 6\delta) + \delta$ | $4\alpha - 3\delta$ | $4\alpha - 3\delta$ | $\alpha(4 - 6\delta) + \delta$ |

Notice that D_2 alone does not dominate any sequence, since $u(D_2(wy), X) - u(wy, X) = u(D_2(wx), Y) - u(wx, Y) = -3\delta < 0$. But when we consider a randomization between D_1 and D_2 , we may be able to exclude some values of δ that are greater than $4/5$. Indeed, a mixed deviation

¹⁰Notice that, by Theorem 2, we can assume without loss of generality that the deviation rule recommends no deviations for action sequences which are not wx or wy . This means that the change in payoffs would be zero for the omitted columns.

rule $\alpha D_1 + (1 - \alpha) D_2$ dominates wx when the system of inequalities

$$\begin{aligned}\alpha(4 - 6\delta) + \delta &> 0 \\ 4\alpha - 3\delta &\geq 0\end{aligned}$$

has a solution $\alpha \in [0, 1]$, which happens precisely when $\delta < 8/9$. Therefore we learn that when wx is chosen, we must have $\delta \geq 8/9$. In fact, this bound is tight. A prior that puts equal probability on both states and a sequential information structure which gives posteriors of $1/4$ or $3/4$ in the second period rationalize wx for $\delta \geq 8/9$. Thus we have shown that $\mathbb{D}(wx) = [\frac{8}{9}, 1]$.

In conclusion, deviation rules essentially construct a system of inequalities which help identify the set $\mathbb{D}(\mathbf{a})$ for an observed sequence of actions \mathbf{a} .

8 Final remarks

This paper provides a theory of what it means for a sequence of actions to be rationalized when information observed by the agent is not available to the outside analyst. Allowing for any possible sequential information structure, we asked when an action sequence can be chosen with positive probability by an optimizing agent. To answer this question we (i) introduced the notions of deviation rule and true dominance, (ii) proved a duality result that precisely characterizes the set of action sequences that cannot be rationalized, and (iii) provided a theoretical treatment of deviation rules as Markov chains that helps simplify the search for action sequences that cannot be rationalized.

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. The exercise of identifying information using data on observed choices has been part of recent work on axiomatic choice theory (see, for example, Dillenberger, Lleras, Sadowski, and Takeoka [2014], Piermont, Takeoka, and Teper [2016], Lu [2016], and Dillenberger, Krishna, and Sadowski [2018]). However, the techniques involved and the data requirements in that literature are quite different. Closer to our setup, information is modeled as a general dynamic stochastic process in Chambers and Lambert [2018], 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 spirit of revealed preference analysis, 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 re-assignment 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.

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 their entire mapping from signals to actions).

The question of which action sequences can be rationalized can be expressed in terms of communication equilibria (see Myerson [1986] and Forges [1986], and more recently Sugaya and Wolitzky [2018] and Makris and Renou [2019]).¹¹ 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 3) 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.¹² For example, in Figure 3b, no actions are codominated, but the sequence of actions A_r is truly dominated.

A Bayesian persuasion or information design approach (following Kamenica and Gentzkow [2011], Bergemann and Morris [2016], Ely [2017] and Doval and Ely [2018]), extended straightforwardly to our sequential environment could be of independent interest. Beyond characterizing the set of action sequences that can be rationalized by some information structure, we can con-

¹¹Formally, 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 [2019] who consider a mediator who directly knows the state of the world.

¹²By 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.

struct bounds on the probability with which an action sequence can be taken. For instance, in Example 1, it can be shown that the highest probability with which a Bayesian agent can take the action sequence (*invest, pull back*), for any prior and sequential information structure, is given by 30 percent. Therefore, one can conclude that under certain assumptions, if more than 30 percent of agents in a large enough population take the action sequence (*invest, pull back*), these choices would reject the Bayesian model of information processing. These and related questions are promising areas for future work.

9 Appendix

The appendix presents a formal proofs of Lemma 2 and completes the missing claims in the proof of Theorem 1.

9.1 Generalized Wald-Pearce

To generalize the Wald-Pearce Lemma, we invoke a generalization of Farka's lemma due to Motzkin: it says that for any system of linear inequalities that does not have a solution, we can find a direct contradiction from a suitably chosen linear combinations of those inequalities. For our purposes, the following implication of Motzkin's Transposition Theorem will be sufficient.

Lemma 4. *For matrices A and B , the system of inequalities*

$$Ax \geq 0, \quad Bx > 0$$

does not have a solution in vectors x and y if and only if there exist vectors α and β such that

$$\alpha^T A + \beta^T B = 0, \quad \alpha, \beta \geq 0, \quad \beta \neq 0.$$

The original reference for Motzkin's Transposition Theorem, and Lemma 4 which follows as a corollary, is Motzkin [1936]. A more accessible proof can be found in Nemirovski and Roos [2009]. We can now prove Lemma 2, that is the generalization of the Wald-Pearce Lemma.

Statement of Lemma 2: Consider a finite game between player 1 with strategy set A and player 2 with strategy set B . Fix $a \in A$ and $\tilde{B} \subset B$. Then, a is never a best-response for \tilde{B} if and only if a is strictly dominated at \tilde{B} .

Proof of Lemma 2. a is never a best-response for \tilde{B} if and only there is no solution $\beta \in \mathbb{R}^B$ to the

following system of inequalities:

$$\sum_{b \in B} \beta(b) > 0 \quad (3)$$

$$\beta(b) \geq 0 \quad \forall b \in B \quad (4)$$

$$\sum_b [u(a, b) - u(a', b)] \beta(b) \geq 0 \quad \forall a' \in A \quad (5)$$

$$\beta(b) \leq 1 \quad \forall b \in B. \quad (6)$$

$$\sum_{b \in B} \beta(b) = 1 \quad (7)$$

If the system which includes (3), (4), and (5) has a solution, then so does the system which includes (6) and (7) as well. This can be seen by normalizing β by dividing it by $\sum_{b \in B} \beta(b)$, which we already know must be strictly positive from conditions 3 and 4. Hence, we can ignore the last two inequalities.

Now, we can apply Motzkin's Transposition Principle (specifically Lemma 4) to the system given by (3), (4), and (5). This system does not have a solution if and only if there exist $k \in \mathbb{R}_{++}$, $v \in \mathbb{R}_+^B$, and $w \in \mathbb{R}_+^A$ such that

$$v_b + \sum_{a' \in A} [u(a, b) - u(a', b)] w_{a'} + k \mathbb{1}_{B'}(b) = 0 \quad \forall b \in B.$$

Equivalently, we can say that there exists $w \in \mathbb{R}_+^A$ such that, for all $b \in B$,

$$\sum_{a' \in A} [u(a, b) - u(a', b)] w_{a'} \leq 0,$$

with a strict inequality when $b \in B'$. Since that inequality is strict sometimes, we have that $w \neq 0$. Defining $\alpha(a) = \frac{w_a}{\sum w_{a'}}$, we have that α strictly dominates a at B' . \square

9.2 Completing the proof of Theorem 1

There are two pieces in the proof of Theorem 1, referred to in Section 5.2, that we prove here.

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 of "if" direction. 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 6 (i.e. true dominance) and the strictness of it follows from part 1 of the definition. \square

Lemma 5. *The set of all adapted mappings is the convex hull of the set of all pure adapted mappings.*

Proof. This result follows from Kuhn's Theorem (see Kuhn [1953] and Theorem 4.1 in Myerson [1991]). Consider a game where the two players alternate moves:

1. Player 1 chooses $x_1 \in X_1$,
2. Player 2 chooses $y_1 \in Y_1$,
3. Player 1 chooses $x_2 \in X_2$,
4. Player 2 chooses $y_2 \in Y_2$,
5. etc.

The information sets for Player 2 correspond to the histories (x^t, y^{t-1}) for $t = 1, \dots, T$. A behavioral strategy for Player 2 maps each history (x^t, y^{t-1}) into a mixture over Player 2's possible moves Y_t . We can represent this as a collection of stochastic maps, $\alpha_t : X^t \times Y^{t-1} \rightarrow \Delta(Y_t)$ for $t = 1, \dots, T$. A pure strategy can then be formally represented as a collection of functions $f_t : X^t \times Y^{t-1} \rightarrow Y_t$. But this representation specifies what Player 2 would do in histories which are prevented by the strategy itself. Following Kuhn [1953], say that two pure strategies of Player 2 are equivalent if, for every strategy of Player 1, they induce the same probability distribution over outcomes (terminal nodes). Two pure strategies that specify the same decisions on paths that

they may reach are equivalent, so we can represent an equivalence class of such pure strategies by a collection of functions $g_t : X^t \rightarrow Y_t$, or equivalently, by an adapted function $g : X \rightarrow Y$.

Now, a stochastic mapping $\alpha : X \rightarrow \Delta(Y)$ can be equivalently represented by a behavioral strategy $(\alpha_t)_{t=1}^T$, where $\alpha_t : X^t \times Y^{t-1} \rightarrow \Delta(Y_t)$ (for details, see de Oliveira [2018], Lemma 3). Kuhn's Theorem states that there exists a mixed strategy which induces, for every distribution over X , the same distribution over Y as $(\alpha_t)_{t=1}^T$, and hence α . By the remark above, we can write this mixed strategy as a probability distribution over adapted functions $g : X \rightarrow Y$. \square

9.3 Proofs for relevant deviations

9.3.1 Backward induction

When stating Proposition 2, we informally referred to a "subproblem". We begin by defining this concept precisely.

Definition 10. *We refer to the collection of action sets A_1, \dots, A_T , together with the utility function $u : A \times \Omega \rightarrow \mathbb{R}$, as the agent's decision problem. The subproblem obtained by fixing (a_1, \dots, a_t) is the subcollection of action sets A_{t+1}, \dots, A_T , together with the utility function $v : A_{t+1} \times \dots \times A_T \times \Omega \rightarrow \mathbb{R}$ defined by*

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

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

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

Proof of Proposition 2. Here we will use the following notation: given $\mathbf{b} = (b_1, \dots, b_T)$, we will let $\mathbf{b}|_t = (b_1, \dots, b_t)$.

Let $D : A_{t+1} \times \dots \times A_T \rightarrow \Delta(A_{t+1} \times \dots \times A_T)$ be a deviation rule for the subproblem such that D dominates (a_{t+1}, \dots, a_T) . Then we can define $F : A \rightarrow \Delta(A)$ to recommend the same deviations as D in the subproblem, and recommend no deviations elsewhere. Formally,

$$F(\mathbf{c}|\mathbf{b}) = \begin{cases} D(c_{t+1}, \dots, c_T | b_{t+1}, \dots, b_T) & \text{if } \mathbf{b}|_t = \mathbf{a}|_t = \mathbf{c}|_t \\ I(\mathbf{c}|\mathbf{b}) & \text{otherwise} \end{cases}$$

where I is the identity ($= 1$ if $\mathbf{c} = \mathbf{b}$, and zero otherwise). Also, recall that for a stochastic mapping, $F(\mathbf{c}|\mathbf{b})$ is the probability that $F(\mathbf{b}) \in \Delta(A)$ puts on the action sequence \mathbf{c} . Now, we claim that F is adapted and also dominates \mathbf{a} . To show that F is adapted, we must show that $\sum_{c_{s+1}, \dots, c_T} F(\mathbf{c}|\mathbf{b})$ does not depend on (b_{s+1}, \dots, b_T) . We show this separately for $s \geq t$ and $s < t$.

So fix $\mathbf{b} = (b_1, \dots, b_T)$ and $\mathbf{c}|_s = (c_1, \dots, c_s)$ and suppose $s \geq t$. Then $\mathbf{b}|_t$ and $\mathbf{c}|_t$ are uniquely determined, and every term in the sum for F is given by D or every term in the sum is given by I . Hence we have

$$\sum_{c_{s+1}, \dots, c_T} F(\mathbf{c}|\mathbf{b}) = \begin{cases} \sum_{c_{s+1}, \dots, c_T} D(c_{t+1}, \dots, c_T | b_{t+1}, \dots, b_T) & \text{if } \mathbf{b}|_t = \mathbf{a}|_t = \mathbf{c}|_t \\ \sum_{c_{s+1}, \dots, c_T} I(\mathbf{c}|\mathbf{b}) & \text{otherwise} \end{cases}$$

In either of those cases, the sum on the right does not depend on (b_{s+1}, \dots, b_T) , since both D and I are adapted.

Now suppose $s < t$. If $\mathbf{c}|_s \neq \mathbf{a}|_s$ or $\mathbf{b}|_t \neq \mathbf{a}|_t$, then we already know that $F(\mathbf{c}|\mathbf{b}) = I(\mathbf{c}|\mathbf{b})$, so again $\sum_{c_{s+1}, \dots, c_T} F(\mathbf{c}|\mathbf{b}) = \sum_{c_{s+1}, \dots, c_T} I(\mathbf{c}|\mathbf{b})$ does not depend on (b_{s+1}, \dots, b_T) . But if $\mathbf{c}|_s = \mathbf{a}|_s$ and $\mathbf{b}|_t = \mathbf{a}|_t$, some terms of the sum have $\mathbf{c}|_t = \mathbf{a}|_t$ and others have $\mathbf{c}|_t \neq \mathbf{a}|_t$. In that case, we can write

$$\begin{aligned} \sum_{c_{s+1}, \dots, c_T} F(\mathbf{c}|\mathbf{b}) &= \sum_{\{\mathbf{c}|\mathbf{c}|_t = \mathbf{a}|_t\}} D(c_{t+1}, \dots, c_T | b_{t+1}, \dots, b_T) + \sum_{\{\mathbf{c}|\mathbf{c}|_t \neq \mathbf{a}|_t\}} I(\mathbf{c}|\mathbf{b}) \\ &= \sum_{c_{t+1}, \dots, c_T} D(c_{t+1}, \dots, c_T | b_{t+1}, \dots, b_T) + \sum_{\{\mathbf{c}|\mathbf{c}|_t \neq \mathbf{a}|_t\}} I(\mathbf{c}|\mathbf{a}) \\ &= 1 + 0 \\ &= \sum_{c_{s+1}, \dots, c_T} I(\mathbf{c}|\mathbf{b}). \end{aligned}$$

For the third equality, the first term is 1 because we are summing over the entire support of that distribution, and the second term is zero because when $\mathbf{c}|_t \neq \mathbf{a}|_t$ the identity map gives a value of zero. So we have shown that, whenever $s < t$, we have $\sum_{c_{s+1}, \dots, c_T} F(\mathbf{c}|\mathbf{b}) = \sum_{c_{s+1}, \dots, c_T} I(\mathbf{c}|\mathbf{b})$. Summarizing all cases, we have shown

$$\sum_{c_{s+1}, \dots, c_T} F(\mathbf{c}|\mathbf{b}) = \begin{cases} \sum_{c_{s+1}, \dots, c_T} D(c_{t+1}, \dots, c_T | b_{t+1}, \dots, b_T) & \text{if } s \geq t \text{ and } \mathbf{b}|_t = \mathbf{a}|_t = \mathbf{c}|_t \\ \sum_{c_{s+1}, \dots, c_T} I(\mathbf{c}|\mathbf{b}) & \text{otherwise.} \end{cases}$$

In either case, the sum on the right-hand side does not depend on (b_{s+1}, \dots, b_T) , since D and I are adapted. Hence, F is adapted as well.

To see that F dominates \mathbf{a} , notice that $u(F(\mathbf{b}), \omega) = u(I(\mathbf{b}), \omega) = u(\mathbf{b}, \omega)$ if $\mathbf{b}|_t \neq \mathbf{a}|_t$. When $\mathbf{b}|_t = \mathbf{a}|_t$, we have that $u(F(\mathbf{b}), \omega) \geq u(\mathbf{b}, \omega)$ for all \mathbf{b} and with strict inequality for $\mathbf{b} = \mathbf{a}$, since there F recommends the same deviation as D . \square

9.3.2 Simple deviations

We will need a few lemmata before proving Proposition 2. The proof is by induction, and the first lemma is a version of Proposition 2 involving only the first period action.

Lemma 6. *If $\mathbf{a} = (a_1, \dots, a_T)$ is a truly dominated action sequence, then there exists a deviation rule D that dominates \mathbf{a} , satisfying $D(\mathbf{b}) = I(\mathbf{b})$ whenever $b_1 \neq a_1$.*

Proof. The argument here is similar to the proof of Proposition 2. Let F be a deviation rule that dominates \mathbf{a} and let

$$D(\mathbf{b}) = \begin{cases} F(\mathbf{b}) & \text{when } b_1 = a_1 \\ I(\mathbf{b}) & \text{when } b_1 \neq a_1 \end{cases}.$$

It is easy to check that D is also adapted and dominates \mathbf{a} . To show adaptedness, we can show, as in Proposition 2, that

$$\sum_{c_2, \dots, c_T} D(\mathbf{c}|\mathbf{b}) = \begin{cases} \sum_{c_2, \dots, c_T} F(\mathbf{c}|\mathbf{b}) & \text{if } b_1 = a_1 \\ \sum_{c_2, \dots, c_T} I(\mathbf{c}|\mathbf{b}) & \text{if } b_1 \neq a_1 \end{cases}$$

so adaptedness of D follows from adaptedness of F and I . To see that D dominates \mathbf{a} , notice that

$$u(D(\mathbf{b}), \omega) = \begin{cases} u(F(\mathbf{b}), \omega) & \text{when } b_1 = a_1 \\ u(\mathbf{b}, \omega) & \text{when } b_1 \neq a_1 \end{cases}.$$

□

The key idea behind the lemmata that follow is to regard a deviation rule as the transition probabilities of a Markov chain, with the set of action sequences A as the set of "markov states" (not to be confused with the "states of nature", $\omega \in \Omega$). We can then label action sequences according to their properties as markov states:

Definition 11. *Let D be a deviation rule. We say that an action sequence $\mathbf{a} \in A$ is*

1. recurrent if, starting from \mathbf{a} , the probability of eventually returning to \mathbf{a} is one;
2. transient if, starting from \mathbf{a} , the probability of eventually returning to \mathbf{a} is less than one;
3. absorbing if $D(\mathbf{a}|\mathbf{a}) = 1$;
4. repulsive if, for every $\mathbf{b} \in A$, $D(\mathbf{a}|\mathbf{b}) = 0$.

Note that absorbing implies recurrent and repulsive implies transient, but the converse in each case is not true. This is standard terminology for Markov chains, with the exception of the definition of "repulsive" (see Kemeny, Snell, and Knapp [1976]). The idea that follows relates

payoff dominance of action sequences to their properties as Markov states. The intuition here is that if D dominates \mathbf{a} then it must have a tendency to move the agent away from \mathbf{a} , so that \mathbf{a} will be transient.

Lemma 7. *Let D be a deviation rule that dominates \mathbf{a} . Then \mathbf{a} is a transient state for D .*

Proof. If \mathbf{a} is a recurrent state, then there exists a stationary probability $\alpha \in \Delta(A)$ such that $D \circ \alpha = \alpha$ and $\alpha(\mathbf{a}) > 0$ (Theorem 6.9 in Kemeny, Snell, and Knapp [1976]). Now define a distribution $\gamma \in \Delta(A \times \Omega)$ by $\gamma(\mathbf{b}, \omega) = \alpha(\mathbf{b}) p(\omega)$, where $p \in \Delta(\Omega)$ is arbitrary. Then $\gamma(\mathbf{a}) > 0$ and $\mathbb{E}_\gamma [u(D(\mathbf{b}), \omega)] = \mathbb{E}_\gamma [u(\mathbf{b}, \omega)]$, which contradicts the fact that D dominates \mathbf{a} . \square

The following fundamental construction is what allows us to turn transient action sequences to the stronger property of being repulsive. Here we use the notation D^k to mean the composition of D with itself k times.

Lemma 8. *Let D be any deviation rule. Then*

$$D^\infty(\mathbf{c}|\mathbf{b}) = \lim_n \frac{1}{n} \sum_{k=1}^n D^k(\mathbf{c}|\mathbf{b})$$

is a well defined deviation rule. Moreover,

1. $D^\infty = D \circ D^\infty = D^\infty \circ D = D^\infty \circ D^\infty$
2. *If \mathbf{b} is dominated by D , then it is also dominated by D^∞ ;*
3. *If \mathbf{b} is transient for D , then it is also repulsive for D^∞ ;*
4. *If \mathbf{b} is absorbing for D , then it is also absorbing for D^∞*

Proof. The proof that the limit exists and of (1) follows that of the Ergodic Theorem for Markov Chains, (see Theorem 6.1 in Kemeny, Snell, and Knapp [1976]). The proof of (4) follows straightforwardly, by construction. Here we prove parts (2) and (3).

If D dominates \mathbf{a} , then for every $\mathbf{b} \in A$ and $\omega \in \Omega$,

$$\begin{aligned} u(D^k(\mathbf{b}), \omega) &= \sum_{\mathbf{c}} u(D(\mathbf{c}), \omega) D^{k-1}(\mathbf{c}|\mathbf{b}) \\ &\geq \sum_{\mathbf{c}} u(\mathbf{c}, \omega) D^{k-1}(\mathbf{c}|\mathbf{b}) \\ &\vdots \\ &\geq u(D(\mathbf{b}), \omega) \end{aligned}$$

This shows that D^k dominates \mathbf{a} as well. Combining the inequalities for different k and taking the limit, we conclude that $u(D^\infty(\mathbf{b}), \omega) \geq u(D(\mathbf{b}), \omega)$ as well, so D^∞ dominates \mathbf{a} .

To prove (3), suppose \mathbf{b} is transient. We must show that $D^\infty(\mathbf{b}|\mathbf{a}) = 0$ for all $\mathbf{a} \in A$. If \mathbf{a} is recurrent, then we must have $D^k(\mathbf{b}|\mathbf{a}) = 0$ for all k , otherwise \mathbf{b} would be recurrent as well (see Lemma 4.23 in Kemeny, Snell, and Knapp [1976]). If \mathbf{a} is transient, then we have that $\lim_k D^k(\mathbf{b}|\mathbf{a}) = 0$ (by Proposition 5.3 in Kemeny, Snell, and Knapp [1976]). Hence, in either case, we have that $D^\infty(\mathbf{b}|\mathbf{a}) = 0$, so \mathbf{b} is repulsive for D^∞ . \square

A given deviation rule D induces a deviation rule up to each period t , which we will denote by D_t . That is,

$$D_t(c_1, \dots, c_t | b_1, \dots, b_t) = \sum_{c_{t+1}, \dots, c_T} D(c_1, \dots, c_T | b_1, \dots, b_T)$$

where b_{t+1}, \dots, b_T can be chosen arbitrarily, since D is adapted.

Lemma 9. *If $D(\mathbf{b}) = I(\mathbf{b})$ whenever $b_1 \neq a_1$ and $D_1(a_1|a_1) < 1$, then $D_1^\infty(a_1|a_1) = 0$.*

Proof. Let $\lambda = D_1(a_1|a_1) < 1$ and notice that

$$D_1^2(a_1|a_1) = \sum_{b_1} D_1(a_1|b_1) D_1(b_1|a_1) = D_1(a_1|a_1) D_1(a_1|a_1)$$

since $D_1(a_1|b_1) = 0$ whenever $b_1 \neq a_1$. From this, we deduce that $D_1^k(a_1|a_1) = \lambda^k$. Hence

$$D_1^\infty(a_1|a_1) = \lim_n \frac{1}{n} \sum_{k=0}^{n-1} D_1^k(a_1|a_1) = \lim_n \frac{1}{n} \left(\frac{1 - \lambda^n}{1 - \lambda} \right) = 0.$$

\square

Proof of Proposition 2. Suppose $\mathbf{a} = (a_1, \dots, a_T)$ is truly dominated. We want to show that there exists a deviation rule D that dominates \mathbf{a} and there exists a t such that D removes (a_1, \dots, a_t) . We proceed by induction on T . The idea is that starting from a deviation rule D that dominates \mathbf{a} , we will show that D^∞ removes (a_1, \dots, a_t) for some t .

Let $T = 1$. Then, obviously there exists a deviation rule D such that $D(b_1) = I(b_1)$ whenever $b_1 \neq a_1$ and $D(a_1|a_1) < 1$. Thus from Lemma 9, we have $D_1^\infty(a_1|a_1) = 0$.

Next, suppose that result holds for $T - 1$ where $T \geq 2$. We want to show that it is true for T . Using Lemma 6, let D dominate \mathbf{a} such that $D(\mathbf{b}) = I(\mathbf{b})$ whenever $b_1 \neq a_1$. Now there are two possible cases to consider: $D_1(a_1|a_1) < 1$ and $D_1(a_1|a_1) = 1$.

If $D_1(a_1|a_1) < 1$, then from Lemma 9 we know that $D_1^\infty(a_1|a_1) = 0$, and thus $D^\infty(\mathbf{b}|\mathbf{a}) = 0$ for $b_1 = a_1$. Moreover, recollect from Lemma 8 part (4), we know that $D^\infty(\mathbf{b}) = I(\mathbf{b})$ for $b_1 \neq a_1$. Therefore, we can conclude that D^∞ removes a_1 .

Now, suppose $D_1(a_1|a_1) = 1$, which means that the deviation rule D takes every sequence starting with a_1 to another sequence starting with a_1 . Therefore, it naturally defines a deviation

rule for the subproblem that fixes a_1 . Further, it is easy to see that the induced deviation rule shows (a_2, \dots, a_T) to be truly dominated in the subproblem (just reverse the construction in the proof of Proposition 2).

Note that the subproblem is of length $T - 1$, so using the induction hypothesis, we know that there exists a deviation rule F and a t such that F dominates (a_2, \dots, a_T) and F removes (a_2, \dots, a_t) . Thus, in original problem the deviation rule G , defined by $G(\mathbf{b}) = I(\mathbf{b})$ whenever $b_1 \neq a_1$, and $G(\mathbf{c}|\mathbf{b}) = F(c_2, \dots, c_T | b_2, \dots, b_T)$ if $b_1 = a_1$, dominates \mathbf{a} and removes (a_1, \dots, a_t) . \square

9.3.3 One deviation to rule them all

Proof of Proposition 3. Let Σ denote the set of truly dominated action sequences. For each $\mathbf{a} \in \Sigma$, pick a deviation rule that dominates \mathbf{a} , and let D be a strict convex combination of all those deviation rules. Then D simultaneously dominates all $\mathbf{a} \in \Sigma$ and by Lemma 8, so does D^∞ . By Lemma 7 every $\mathbf{a} \in \Sigma$ is transient for D . By Lemma 8, part (3), every $\mathbf{a} \in \Sigma$ is repulsive for D^∞ . \square

References

- D. Bergemann and S. Morris. Bayes correlated equilibrium and the comparison of information structures in games. *Theoretical Economics*, 11(2):487–522, 2016.
- 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. Georgetown University and Stanford University, 2018.
- H. de Oliveira. Blackwell’s informativeness theorem using diagrams. *Games and Economic Behavior*, 109:126–131, 2018.
- D. Dillenberger, J. S. Lleras, P. Sadowski, and N. Takeoka. A theory of subjective learning. *Journal of Economic Theory*, 153(1):287–312, 2014.
- D. Dillenberger, R. V. Krishna, and P. Sadowski. Subjective information choice processes. University of Pennsylvania and Florida State University and Duke University, 2018.

- L. Doval and J. Ely. Sequential information design. Caltech and Northwestern University, 2018.
- J. Ely. Beeps. *American Economic Review*, 107(1):31–53, 2017.
- F. Forges. An approach to communication equilibria. *Econometrica*, 54(6):1375–1385, 1986.
- E. Kamenica and M. Gentzkow. Bayesian persuasion. *American Economic Review*, 101(6):2590–2615, 2011.
- J. G. Kemeny, J. L. Snell, and A. W. Knapp. *Denumerable Markov chains*. Springer-Verlag, 1976.
- H. W. Kuhn. Extensive games and the problem of information. In H. W. Kuhn and A. W. Stiglitz, Tucker, editors, *Contributions to the Theory of Games, Vol II*, pages 193–216. Princeton University Press, 1953.
- 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, 2019.
- 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.
- R. Myerson. *Game theory: Analysis of Conflict*. Harvard university press, 1991.
- A. Nemirovski and K. Roos. Motzkin transposition theorem. In C. A. Floudas and P. M. Pardalos, editors, *Encyclopedia of Optimization*, pages 2345–2348. Springer, Boston, MA, 2009.
- 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.
- 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. 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.