# Learning from a Black Box

Shaowei Ke[*] Brian Wu[†] Chen Zhao[‡§]

May 2022

Please click here for the newest version.

**Abstract**

We study how a decision maker learns when she does not know how recommendations are generated. We introduce three types of behavioral postulates on the updating rule. Some of them characterize the contraction rule. The contraction rule maps each recommendation to its trustworthiness and to one belief consistent with the recommendation, and forms the posterior by mixing the prior with the recommended belief weighted by the trustworthiness measure. Under some other postulates, the updating rule features a form of conservatism. No updating rules, however, can satisfy all the postulates simultaneously. Applications with persistent recency bias and belief divergence are provided.

[*]Department of Economics, University of Michigan. Email: shaoweik@umich.edu.

[†]Strategy Department, Stephen M. Ross School of Business, University of Michigan. Email: wux@umich.edu.

[‡]Faculty of Business and Economics, University of Hong Kong. Email: czhao@hku.hk.

[§]We are grateful to David Ahn, Tilman Börgers, David Dillenberger, Matthew Kovach, Heng Liu, Yang Lu, David Miller, Gerelt Tserenjigmid, and seminar and conference participants at UMich, Toronto, Hitotsubashi Summer Institute, Midwest Economic Theory Conference (MSU), SMU, Virginia Tech, HKUST, UCSC, UT Austin, QUB, PSU, Princeton, ANU, IMSI Decision Making under Uncertainty, Copenhagen Business School, INSEAD, and USC Marshall for helpful discussions.

# 1 Introduction

It is becoming increasingly common that people make decisions with the help of a complex machine learning algorithm. For example, when we are deciding where to eat, Yelp may offer some recommendation. The recommendation may immediately influence our view about how our taste matches with various restaurants, before we make any choices. After Deepmind's AlphaGo became the first computer program to defeat professional Go players, many Go players started to take recommendations from machine learning computer programs such as KataGo and Leela when practicing and learn from them.

Such recommendations are often generated based on datasets with billions of variables and algorithms with millions of parameters. It is nearly impossible for people to understand how the recommendations are generated. Indeed, not even the programmers of the algorithms themselves understand what the algorithm has learned from the data and what the internal logic of the mapping from the input to the output is. Even if a programmer claims to understand how the algorithm works, it is unlikely that she can explain the rationale or theory behind the recommendations to the decision maker.[1]

For these reasons, a complex machine learning algorithm is often called a *black box* (BB). In this paper, we will use this terminology more broadly. If the decision maker does not understand how a recommender generates its recommendations, we will call the recommender a BB.

Specifically, we consider a decision maker who has a prior belief and confronts a menu. A menu is a set of acts (mappings from states to utility outcomes). From the

[1] A growing literature studies how to make complex machine learning algorithms more interpretable and explainable. See Guidotti, Monreale, Ruggieri, Turini, Giannotti, and Pedreschi (2018) for a recent survey.

menu, a BB, which is assumed to be benevolent to the decision maker, recommends an act to the decision maker. Based on the recommendation, the decision maker may revise her belief. Compared with a standard model in which the decision maker receives a recommendation from a Bayesian expert, the main difference in our case is that the decision maker knows precious little about how the BB's recommendation depends on the unobserved state. Take a complex product recommendation algorithm as an example. The unobserved state encodes how well each product matches with the decision maker. The decision maker has some belief about the state. Given a set of products, the algorithm attempts to recommend the best one to the decision maker based on what the algorithm "knows" about the decision maker's taste.[2] Observing the recommendation, the decision maker updates her belief about the state, even though she knows little about how the recommendation is generated.

One can certainly take the classic Bayesian decision-theoretic approach and posit that the decision maker still uses Bayes' rule to learn from the BB. Such a theory, however, is unlikely to be realistic in our context. First, the decision maker knows little about how the recommendation is generated, and thus lacks the joint prior needed to perform Bayesian updating. Second, the most informative recommendations from a BB are often unexpected, to which Bayes' rule may not be applicable.[3] Thus, non-Bayesian learning models that impose less informational demand on the decision maker as Bayes' rule and have the flexibility to incorporate unexpected recommendations will be our goal.

To search for such a model, we introduce an axiomatic theory of learning from

[2]In practice, this set may be determined by, for example, what products the online platform has to offer, and some searching criteria that the decision maker provides. These details are not crucial to our findings. We simply assume that this set is exogenously given.

[3]In the game of Go, the most influential moves recommended by AI are, almost by definition, complete surprises to the profession. For instance, "move 37" in the second game between AlphaGo and the human champion Lee Sedol was considered by many as strange, surprising, and likely a mistake at the moment, but was later celebrated as a beautiful move that shaped the future of Go.

a BB. The primitive of our theory is an *updating rule*. An updating rule takes the decision maker's prior belief and a recommendation—a pair of a recommended act and a menu—and maps them to a posterior belief. We do not make assumptions about how the BB generates the recommendation. Rather, we introduce behavioral postulates on the updating rule that will reveal the decision maker's perception about the BB and how she wants to learn from it.

Three types of behavioral postulates are introduced. The first is (weak and strong) confidence monotonicity, which resembles the independence axiom from expected utility theory in our setup. It says that if a belief $p$ is less confident than a belief $q$ about an act $a$ being better than an act $b$, after learning from the same recommendation, $p$'s posterior should continue to be less confident than $q$'s posterior about $a$ being better than $b$.

The second postulate, partial compliance, states that for any recommendation, there is always some prior belief such that the decision maker will be convinced by the BB's recommendation, even though her prior does not agree with it. If this postulate fails, there will be some recommendation such that whenever the decision maker's prior belief is inconsistent with that recommendation, no matter how close the prior belief is to the recommendation, the decision maker never follows it.

The last postulate, (weak and strong) long-run compliance, uses long-run learning behavior to inform short-run learning behavior. Roughly speaking, it requires that the decision maker comply with the black box in the long run upon receiving repeated recommendations from a sequence of BBs who make independent recommendations of the same quality.

Our first main result shows that strong confidence monotonicity and weak long-run compliance characterize what we call the *contraction rule*. The contraction rule has two components. One is a function that maps each recommendation to one recommended belief that is consistent with the recommendation. The other is a function that maps each recommendation to a measure of how much the decision maker trusts the recommendation. The decision maker's posterior is given by mixing her prior with the recommended belief, weighted by the measure of trustworthiness. If, in addition, every recommendation is reduced to an interior recommended belief, then the contraction rule satisfies partial compliance.

The contraction rule is similar in spirit to one of the most widely used non-Bayesian learning models, the DeGroot (1974) model.[4] The decision maker in this model forms the posterior by computing the weighted sum of her prior belief and the other people's beliefs. The weights measure the trust between different people. Under the contraction rule, the decision maker forms the posterior by computing the weighted sum of her prior belief and a subjectively chosen belief that is consistent with the recommendation. The weight measures the trustworthiness of the recommendation.

Next, we show that if weak confidence monotonicity, partial compliance, and strong long-run compliance are imposed on the updating rule instead, the updating rule must exhibit some form of conservatism. In particular, if the decision maker believes that an act $b$ is weakly better than an act $a$, but receives the recommendation that $a$ is weakly better than $b$ repeatedly from a sequence of BBs who make independent recommendations of the same quality, then after several repetitions, the decision maker must believe that both acts are equally good. In other words, she does not abandon her initial prior to believe that $a$ is strictly better than $b$. Many non-Bayesian updating rules from the literature are indeed conservative. See Section 5 for a discussion.

[4]For its applications and discussions in economics, see, among others, DeMarzo, Vayanos, and Zwiebel (2003); Golub and Jackson (2010, 2012); Jackson (2011); and Jackson and Zenou (2015).

One may wonder what happens if we impose strong confidence monotonicity, partial compliance, and strong long-run compliance on the updating rule. Our last main result shows that strong confidence monotonicity and strong long-run compliance are incompatible with each other. Therefore, at least one of these two postulates must be weakened before being imposed on the updating rule. Contraction rules and conservatism, however, are not incompatible. We characterize when a contraction rule is conservative.

Finally, we illustrate some implications of the contraction rule in simple examples. We show that the contraction rule may lead to persistent and reinforcing recency bias and belief divergence.[5]

## 1.1 Related Literature

In our setup, each recommendation induces a set of beliefs that are consistent with the recommendation (see Section 4). In the decision theory literature, Zhao (2021) and Dominiak, Kovach, and Tserenjigmid (2021) consider primitives that map the decision maker's prior belief and information that takes the form of a set of beliefs to a posterior. They propose updating rules that select, from the given set of distributions, the posterior belief closest to her prior according to some subjective divergence measure. Both papers interpret information as a constraint with which the decision maker's posterior must be consistent. In information theory, there is also a literature that considers belief updating when new information imposes constraints on the probability distribution; see, for example, Shore and Johnson (1980); Skilling (1988); and Caticha (2004). In contrast to all the papers above, we allow the decision maker

[5] In economics, recency bias has been studied in, for example, Fudenberg, Levine, et al. (2014); Bansal and Shaliastovich (2010); and Zhao (2021). There is a large literature on divergence, and many theories of divergence are based on learning. See, among others, Baliga, Hanany, and Klibanoff (2013); Bowen, Dmitriev, and Galperti (2022); Eguia and Hu (2022); and Perego and Yuksel (2022).

to not fully trust the BB, and thus her posterior may not be immediately consistent with the recommendation. Moreover, the format of our information is a pair of a recommended act and a menu, which differs from a set of beliefs. In particular, we offer an example in Section 4 to show why two recommendations may induce the same set of beliefs that are consistent with the recommendations and yet the decision maker may respond differently to those recommendations.

There is a much larger literature on non-Bayesian updating using standard information, i.e., the occurrence of an event, and on learning under ambiguity. For axiomatic decision models, see, for example, Epstein (2006); Ortoleva (2012); Zhao (2020); Cheng (2021); Kovach (2021); and Suleymanov (2021). Of these studies, Epstein (2006) and Kovach (2021) characterize the prior-biased updating rule: The decision maker's behavioral posterior is a convex combination of her prior and the Bayesian posterior. Our contraction rule also features a convex combination between the decision maker's prior and the recommended belief, but is defined on a different primitive.

One of the most widely used non-Bayesian updating rules in social learning is introduced by DeGroot (1974). Our main representation of the updating rule, the contraction rule, is similar to the DeGroot social learning model. In the DeGroot model, in each period, decision makers' prior beliefs are summarized in a prior belief vector. It is multiplied by a matrix that measures trust between different pairs of decision makers to derive the decision makers' posterior belief vector. The trust measures only depend on the identities of the decision makers.

In a recent paper, Cheng (2022) also studies how a decision maker should learn when she has limited knowledge about how the information is generated. The decision maker is a maxmin expected utility maximizer (see Gilboa and Schmeidler (1989)), and she faces ambiguity over the data-generating process behind the sequential information she receives. It is shown that widely used learning rules for the maxmin expected utility model are suboptimal. Cheng proposes a new learning rule that exhibits desirable features. In contrast to Cheng's approach, Chen (2022) takes existing ambiguity models and updating rules and studies the equilibrium of an information cascade model. The main result shows that an information cascade occurs almost surely under the maxmin expected utility model and the full Bayesian updating rule (Pires (2002)), when all priors are considered plausible.

One component of the contraction rule selects a belief that is consistent with the recommendation. Chambers and Hayashi (2010) and Damiano (2006) also model how a decision maker selects her belief from a set of objectively possible probability measures. In these studies, the decision maker does not have a prior to begin with, and thus the selection rule depends only on the set of probability measures. The contraction rule, in contrast, will combine the selected belief with the prior belief to form the new belief.

In a recent paper, Park and Tayawa (2022) show that (an old version of) the contraction rule is sensitive to the order following which recommendations arrive, which is consistent with our applications about the recency bias. They propose a new updating rule in a restricted domain that is order independent.

It is well recognized in the machine learning literature that relying on BBs to make decisions may cause biases (Pedreshi, Ruggieri, and Turini (2008); Barocas and Selbst (2016)); legal liability issues (Kingston (2016); Bathaee (2018)); and severe consequences (Wexler (2017); Nunes, Reimer, and Coughlin (2018)). To open the BB, the literature follows two directions: (i) ex ante designing interpretable models to make predictions (see, for example, Doshi-Velez and Kim (2017) and, in an economic context, Ke, Zhao, Wang, and Hsieh (2022)); (ii) ex post seeking to explain the predictions made by BBs (see, for example, Ribeiro, Singh, and Guestrin (2016) and

Guidotti et al. (2018)). In this paper, we keep the BB closed but investigate how a decision maker incorporates its recommendations into her beliefs. Our results add to the literature by highlighting the general difficulty of learning from BBs: The decision maker faces a trade-off between some desirable properties since she does not understand how the recommendations are generated.

Our paper proceeds as follows. In Section 2, we introduce our setup and primitive. Section 3 presents the behavioral postulates. Section 4 characterizes the contraction rule, and Section 5 studies conservative updating rules and introduces a negative result. In Section 6, several applications of the contraction rules are analyzed. Section 7 concludes.

## 2 Preliminaries

Let the state space $\Theta$ be a compact metric space, $\Sigma$ be the Borel $\sigma$-algebra defined on $\Theta$, and $\Delta(\Theta)$ be the set of all Borel probability measures defined on $\Theta$. We endow $\Delta(\Theta)$ with the topology of weak convergence. Generic elements of $\Delta(\Theta)$, denoted by $p, q, r$, are called beliefs. An *act* is a continuous function that maps each state into a utility outcome. Let the set of all available acts be $\mathcal{A}$, with generic elements denoted as $a, b, c$. Under belief $p \in \Delta(\Theta)$, the decision maker's expected utility of an act $a \in \mathcal{A}$ is $U(a,p) = \int_\Theta a(\theta)p(d\theta)$. A nonempty finite set of acts $A \subseteq \mathcal{A}$ is a *menu* if for any $a \in A$, there exist $p \in \Delta(\Theta)$ such that $U(a,p) > U(b,p)$ for any $b \in A$ with $b \neq a$. In other words, there are no redundant acts in a menu. Let $\mathcal{M}$ be the set of all menus and $A, B, C$ denote generic menus.

The decision maker has a prior belief and encounters an exogenously given menu. From the menu, the BB recommends an act to the decision maker. We assume that the BB is benevolent, i.e., it always recommends the best act under its own belief.

Based on the recommendation, the decision maker then updates her belief. Putting this setup in an example, think of a decision maker who receives a recommendation from an algorithm. The unobserved state is the taste type of the decision maker, which specifies how well each product matches with her. The algorithm may genuinely attempt to recommend the best product based on its “understanding” about the decision maker’s taste. From the recommendation, the decision maker learns something useful about the state of the world.

How does the BB in our model differ from a Bayesian expert in standard models? Figure 1 compares the Bayesian network of a standard model with that of our model.[6] In standard models, the decision maker knows the conditional distribution of the head node given the tail node for each edge of the Bayesian network. In particular, she knows the conditional distribution of the signals that the expert receives given the state; she knows the expert’s prior and that the expert uses Bayes’ rule for belief updating; she also knows the conditional distribution of the recommendations given the expert’s belief, either because the expert is benevolent, or through the equilibrium strategy of the expert. From the knowledge above, the decision maker is able to calculate the conditional distribution of the recommendations given the state, which enables her to update her belief using Bayes’ rule.

In contrast, in our model, the decision maker knows that the BB is benevolent when making recommendations, but she knows little about how the BB’s belief depends on the unobserved state. Take a complex machine learning recommendation algorithm as an example. Either directly or indirectly, the recommendation algorithm essentially wants to form a belief about the decision maker’s taste type, but the algorithm is so complex that not even the programmers themselves understand how the

[6] In a Bayesian network, each node represents a random variable and each directed edge represents the conditional dependency of the head node on the tail node.

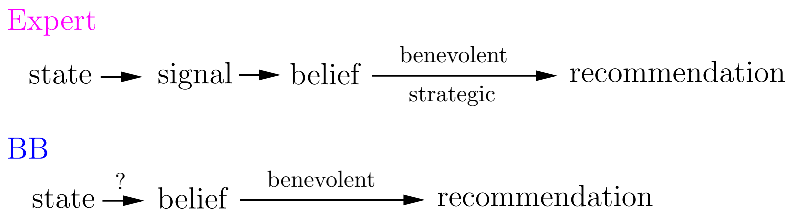


FIGURE 1: Bayesian Expert vs BB

belief is estimated from data, not to mention the conditional distribution of the BB's belief given each state.

We take a model-free approach and introduce an axiomatic theory of learning from a BB on a general primitive, an *updating rule*. An updating rule associates, with each prior belief and each recommendation, a posterior belief. Formally, let $\mathcal{I} = \{(a, A) | a \in A \in \mathcal{M}\}$, which is the set of all recommendations the decision maker may receive. An updating rule is a function $\pi : \Delta(\Theta) \times \mathcal{I} \to \Delta(\Theta)$. When there is no risk of confusion, we write $p_I$ instead of $\pi(p, I)$ for any $p \in \Delta(\Theta)$ and $I \in \mathcal{I}$.

In some results, we need the following technical assumption. We say that $\mathcal{A}$ is *rich* if it contains any continuous function $f : \Theta \to [-1, 1]$. This assumption can be easily satisfied, for example, in standard Anscombe–Aumann type of settings.

# 3 Behavioral Postulates

In this section, we introduce the postulates that will be imposed on the updating rule. The first one is confidence monotonicity. Its main idea can be easily seen in a binary-state example. Suppose $\Theta = \{0, 1\}$. In this case, the decision maker's belief is fully described by a number between 0 and 1, indicating her belief that the state $\theta$ is 1. Confidence monotonicity assumes that, fixing any recommendation $I$, the decision maker's posterior belief $p_I$ is increasing in her prior belief $p$:[7]

$$\text{For any } p, q \in \Delta(\Theta) \text{ and } I \in \mathcal{I},\ p \geq q \text{ implies } p_I \geq q_I.$$

In this example, the space of probability measures is linearly ordered so that we may state inequalities such as $p \geq q$. The space of probability measures is not necessarily linearly ordered in general. Therefore, we introduce a preorder (reflexive and transitive binary relation) on $\Delta(\Theta)$.

**Definition 1.** *For any acts $a, b \in \mathcal{A}$ such that $\{a, b\} \in \mathcal{M}$, we say that $p \in \Delta(\Theta)$ is less confident than $q \in \Delta(\Theta)$ about $a$ being better than $b$ if $U(a, \gamma p + (1-\gamma)r) \geq U(b, \gamma p + (1-\gamma)r)$ implies $U(a, \gamma q + (1-\gamma)r) \geq U(b, \gamma q + (1-\gamma)r)$ for any $\gamma \in [0,1]$ and $r \in \Delta(\Theta)$. We denote this by $p \sqsubset_b^a q$.*

To understand this definition, consider two cases with beliefs $p$ and $q$ respectively, and suppose that the decision maker entertains a third belief $r$ for both cases and forms new beliefs $\gamma p + (1-\gamma)r$ and $\gamma q + (1-\gamma)r$. If whenever $a$ is better than $b$ under the new belief $\gamma p + (1-\gamma)r$, so must $a$ be better than $b$ under the new belief $\gamma q + (1-\gamma)r$, then we say that the original belief $p$ is less confident than $q$ about $a$ being better than $b$. See Figure 2 for an example when $|\Theta| = 3$.

The following lemma shows that $\sqsubset_b^a$ has a simple cardinal representation.

**Lemma 1.** *For any $p, q \in \Delta(\Theta)$ and $\{a, b\} \in \mathcal{M}$,*

$$p \sqsubset_b^a q \iff U(a, p) - U(b, p) \leq U(a, q) - U(b, q).$$

Below are two notions of confidence monotonicity that extend the one from the

[7] Under Bayes' rule, fixing the conditional distribution of signals given each state, the decision maker's posterior belief about $\theta = 1$ will be increasing in her prior belief upon receiving the same signal.

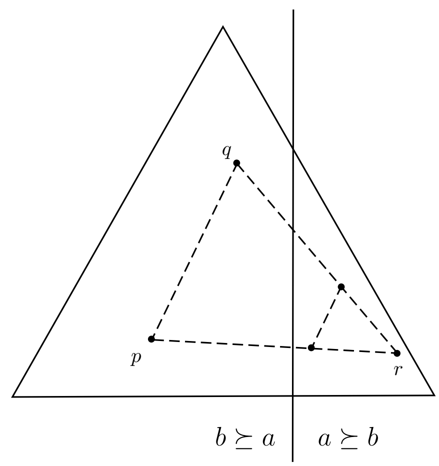


FIGURE 2: Whenever $\gamma p + (1-\gamma) r$ is in the region where $a$ is better, so must $\gamma q + (1-\gamma) r$. The vertical line in the graph indicates the beliefs under which $a$ and $b$ are equally good. The expression $b \succeq a$ in the picture indicates that under any beliefs to the left of the line, $b$ is better than $a$. Similarly, the expression $a \succeq b$ indicates that under any beliefs to the right of the line, $a$ is better than $b$.

binary-state example. As will be seen in Section 4.1, confidence monotonicity's implications will be similar to the independence axiom from expected utility theory. Therefore, it can be viewed as the analogue of the independence axiom in our setting.

**Postulate 1** (Weak Confidence Monotonicity)**.** *For any* $p, q \in \Delta(\Theta)$ *and* $\{a, b\} \in \mathcal{M}$, $p \sqsubset_b^a q$ *implies* $p_{(a,\{a,b\})} \sqsubset_b^a q_{(a,\{a,b\})}$.

**Postulate 2** (Strong Confidence Monotonicity)**.** *For any* $p, q \in \Delta(\Theta)$, $\{a, b\} \in \mathcal{M}$, *and* $I \in \mathcal{I}$, $p \sqsubset_b^a q$ *implies* $p_I \sqsubset_b^a q_I$.

The idea behind confidence monotonicity is simple. If a belief is more confident than another about $a$ being better than $b$, upon receiving the same recommendation, the former belief should continue to be more confident than the latter about $a$ being better than $b$. This statement applies to all recommendations under strong confidence monotonicity, and only applies to the recommendation that indicates that $a$ is better than $b$ under weak confidence monotonicity, in which $a, b$ are the acts involved in the confidence preorder.

Clearly, if the BBs' recommendations are uninformative whatsoever, no learning should take place. The rest of our postulates describe the decision maker's natural behavior when she believes that the recommendations are (at least minimally) informative. We do not specify the notion of informativeness in the model, but our representations will reveal the decision maker's perception about, for example, the truthworthiness of the recommendations.

To state the postulates, we first introduce a definition. For any belief $p \in \Delta(\Theta)$ and menu $A \in \mathcal{M}$, let

$$\alpha(p, A) = \{(a, A) | U(a, p) \geq U(b, p) \text{ for any } b \in A\}$$

be the set of benevolent recommendations from $A$ assuming that the belief is $p$.

**Postulate 3** (Partial Compliance)**.** *For any $I = (a, A) \in \mathcal{I}$, there exists $p \in \Delta(\Theta)$ such that $I \notin \alpha(p, A)$ but $I \in \alpha(p_I, A)$.*

Partial compliance says that for any recommendation $(a, A)$, there are always some cases in which the decision maker disagrees with the recommendation before receiving it, but becomes convinced by the recommendation upon receiving it. Presumably, this may happen when the decision maker's prior is close to the set of beliefs that make $a$ optimal in $A$.

Our last postulate describes the decision maker's long-run learning behavior in a hypothetical stationary setting, which will impose reasonable restrictions on the decision maker's short-run learning behavior back in our original setting. Consider a situation in which the decision maker consults multiple BBs sequentially on the same menu, and, more importantly, the BBs make independent recommendations of similar quality. Throughout this learning process, the decision maker does not make any choices. For example, imagine that the decision maker browses through several apps that make independent recommendations of similar quality, and the decision maker keeps learning from the recommendations about the state (how well each product matches with her) without choosing or consuming anything yet.

In this hypothetical setting, we assume that the decision maker adopts a stationary updating rule throughout the learning process: Each time she receives a recommendation, she updates her current belief using the same updating rule $\pi$.

To see why we make this assumption, it is useful to understand the Bayesian analogue of this hypothetical setting. Consider a Bayesian decision maker who has a belief over the state space. In every period, there is a (new) Bayesian expert who receives a private signal that is informative about the state and recommends an act benevolently from a fixed menu. Signals are i.i.d. given the state. No expert observes any other expert's signal or recommendation. As a result, each expert's belief and recommendation are i.i.d. given the state. Past recommendations are uninformative about the conditional distribution of the new expert's belief given the state. The decision maker's belief over the state space becomes a sufficient statistic of past recommendations. Therefore, the Bayesian decision maker uses the same function that maps her current belief and the current recommendation to a posterior belief to learn throughout the process.

Analogously, when the BBs' beliefs, and hence recommendations, are i.i.d. given the state, there is no reason for our decision maker to adopt different updating rules at different points in time. Hereafter, if we have a sequence of BBs whose beliefs are i.i.d. given each state, we call them i.i.d. BBs.

To simplify notation, recursively, we define $p_{I_1 I_2 \dots I_n} = \pi(p_{I_1 I_2 \dots I_{n-1}}, I_n)$. This notation, $p_{I_1 I_2 \dots I_n}$, is used to denote the decision maker's posterior after learning $I_1, I_2, \dots, I_n$ sequentially from i.i.d. BBs. For any $I \in \mathcal{I}$ and $n \in \mathbb{N}$, let $I^n$ denote a string of $n$ consecutive $I$'s. Then, for example, $p_{I^2} = p_{II}$, $p_{I^3} = p_{III}$.

To state our last postulate, we define a special sequence of recommendations generated by i.i.d. BBs. First, we say that two recommendations $(a, A), (b, A) \in \mathcal{I}$ are *compatible* if there exists some $p \in \Delta(\Theta)$ such that $(a, A), (b, A) \in \alpha(p, A)$. In other words, $(a, A), (b, A) \in \mathcal{I}$ are compatible if under some belief, $a, b$ are both the best from menu $A$.

**Definition 2.** *A sequence of recommendations $\{I_n\}_{n=1}^{\infty}$ (generated by i.i.d. BBs) reveals the indifference between $a \in \mathcal{A}$ and $b \in \mathcal{A}$ in $A \in \mathcal{M}$ if $I_{2k-1} = (a, A)$ and $I_{2k} = (b, A)$ for any $k \geq 1$ and $(a, A), (b, A) \in \mathcal{I}$ are compatible.*

The last postulate concerns the asymptotic beliefs of the decision maker, which also has two versions.

**Postulate 4** (Weak Long-run Compliance)**.** *For any $p \in \Delta(\Theta)$ and $I = (a, A) \in \mathcal{I}$, each accumulation point $q$ of $\{p_{I^n}\}_{n=1}^{\infty}$ satisfies $I \in \alpha(q, A)$.*

**Postulate 5** (Strong Long-run Compliance)**.** *For any $p \in \Delta(\Theta)$, if $\{I_n\}_{n=1}^{\infty}$ reveals the indifference between $a$ and $b$ in $A \in \mathcal{M}$, then each accumulation point $q$ of $\{p_{I_1 I_2 \ldots I_n}\}_{n=1}^{\infty}$ satisfies $(a, A), (b, A) \in \alpha(q, A)$.*

Weak long-run compliance says that if a sequence of i.i.d. BBs recommend $a$ from $A$ repeatedly, regardless of the decision maker's initial belief, she should eventually at least be almost convinced by the BBs. Strong long-run compliance says that if a sequence of i.i.d. BBs recommend acts $a$ and $b$ from $A$ in an alternating fashion, then the decision maker's belief should eventually be arbitrarily close to beliefs under which $a$ and $b$ are both the best from $A$. Of course, in both cases, it is also possible that after a finite number of periods, the decision maker is already convinced.

Below, we analyze the implications of these postulates in different combinations.

# 4 The Contraction Rule

We first introduce a useful definition. For each recommendation $I = (a, A) \in \mathcal{I}$, we define

$$P(a, A) = \{p \in \Delta(\Theta) | U(a, p) \geq U(b, p) \text{ for any } b \in A\}.$$

In other words, $P(a, A)$, also written as $P(I)$, is the set of beliefs that are consistent with the BB's recommendation $(a, A)$.

We introduce a representation of the updating rule that is similar to DeGroot (1974) in spirit (see more discussion in the Introduction). An updating rule $\pi : \Delta(\Theta) \times \mathcal{I} \to \Delta(\Theta)$ is a *contraction rule* if there exists a mapping $\rho : \mathcal{I} \to \Delta(\Theta)$ such that $\rho(I) \in P(I)$ for each $I \in \mathcal{I}$, and a functional $\varepsilon : \mathcal{I} \to [0, 1)$, such that

$$\pi(p, I) = \varepsilon(I) \cdot p + (1 - \varepsilon(I)) \cdot \rho(I)$$

for any $p \in \Delta(\Theta)$ and $I \in \mathcal{I}$. In this case, we say that the decision maker has a contraction rule $(\varepsilon, \rho)$.

By definition, $\rho(I) \in P(I)$ if and only if $I \in \alpha(\rho(I), A)$, for any $I = (a, A) \in \mathcal{I}$. Hence, a contraction rule says that from the decision maker's point of view, each recommendation $I$ can be reduced to one belief that is consistent with the recommendation $I$, $\rho(I)$, and a measure of the trustworthiness of this particular recommendation, $\varepsilon(I)$. The decision maker's posterior is formed by mixing her prior with $\rho(I)$ with probability $\varepsilon(I)$. In this sense, the representation reveals the decision maker's perception of the BBs, and how she wants to learn from the BBs.

The next result characterizes the contraction rule as a representation of the updating rule.

**Theorem 1.** *Suppose* $|\Theta| \geq 3$ *and* $\mathcal{A}$ *is rich. The following statements are true:*

1. *An updating rule satisfies strong confidence monotonicity and weak long-run compliance if and only if it is a contraction rule.*

2. *A contraction rule satisfies partial compliance if* $\rho(I) \in \mathrm{int}(P(I))$ *for all* $I \in \mathcal{I}$.

3. *Two contraction rules* $(\varepsilon, \rho)$ *and* $(\tilde{\varepsilon}, \tilde{\rho})$ *are identical if and only if* $\varepsilon = \tilde{\varepsilon}$ *and* $\rho = \tilde{\rho}$.

The contraction rule imposes little informational demand on the decision maker. Upon receiving a recommendation, a Bayesian decision maker needs to know, or least theorize about, the probability of receiving this recommendation in each state in order to calculate the posterior likelihoods among the states. In contrast, the contraction rule allows the decision maker to know nothing beyond the fact that the recommender is benevolent. Since no model of how the recommendations are generated is needed, whether the recommendation received is expected or a surprise does not matter.

The $\varepsilon$ and $\rho$ functions allow us to model the decision maker's learning rule of thumb in a flexible way. One interesting possibility is the following. The contraction rule allows two recommendations $I, J$ to have the same set of beliefs that are consistent with the recommendations respectively (i.e., $P(I) = P(J)$), but have different $\rho(I)$ and $\rho(J)$ and different $\varepsilon(I)$ and $\varepsilon(J)$.[8] For example, we may find an act $c$ that is not always dominated by either $a$ or $b$, and have $P(a, \{a, b\}) = P(a, \{a, b, c\})$. Compared with the recommendation $(a, \{a, b\})$, the recommendation $(a, \{a, b, c\})$ helps the decision maker rule out more acts, and hence may seem more trustworthy to her. This can be captured if $\varepsilon(a, \{a, b, c\}) < \varepsilon(a, \{a, b\})$.

Next, we provide the sketch of the proof. It will be seen that the role of strong confidence monotonicity in our theory is analogous to that of the independence axiom in expected utility theory.

[8] This feature can be removed if we require that $P(I) = P(J) \Rightarrow p_I = p_J$ for any $p \in \Delta(\Theta)$.

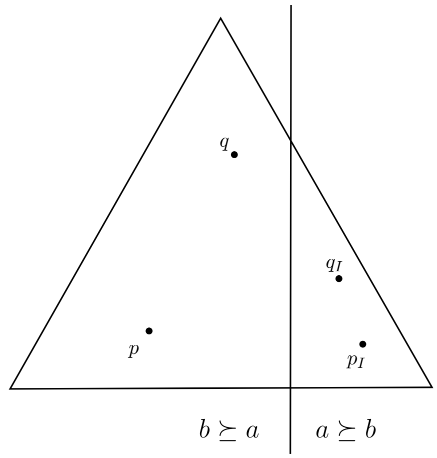


FIGURE 3: If $p_I - q_I \neq \varepsilon(p-q)$, we can find $a, b$ such that $p \sqsubset_b^a q$ but $p_I \not\sqsubset_b^a q_I$.

## 4.1 The Sketch of the Proof

The key step in the proof is to exploit strong confidence monotonicity to show that the difference between the posteriors, as a finite signed measure, has to be a scalar multiple of the difference between the priors. In other words, in the vector space of finite signed measures (denoted as $\Delta^*$), for any two priors $p$ and $q$, the line connecting them must be parallel to the line connecting their posteriors $p_I$ and $q_I$.

To do this, first, we show that for any beliefs $p, q$ and any recommendation $I$, there exists an $\varepsilon$ such that $p_I - q_I = \varepsilon(p-q)$. Suppose this is not true. We can find two alternatives $a, b$ such that $p$ is less confident than $q$ about $a$ being better than $b$, but $p_I$ is more confident than $q_I$ about $a$ being better than $b$, as shown in Figure 3. This is because by richness, for any half-space of $\Delta(\Theta)$, there exist alternatives $a, b$ such that $P(a, \{a, b\})$ is identical to that half-space (see Lemma 3).

So far $\varepsilon$ may depend on the choice of $p, q$ and $I$. Next, we show that $\varepsilon$ must only depend on $I$. Suppose we have $p, q, p', q'$ and $p_I - q_I = \varepsilon(p-q)$, $p'_I - q'_I = \varepsilon'(p'-q')$. We need to show that $\varepsilon = \varepsilon'$. For simplicity, suppose that $p, q, p'$ and $p, p', q'$ form two triangles in $\Delta(\Theta)$, called $pqp'$ and $pp'q'$.[9] They share one edge, the line segment

[9]The fact that we can construct such triangles requires that the number of the states be no less

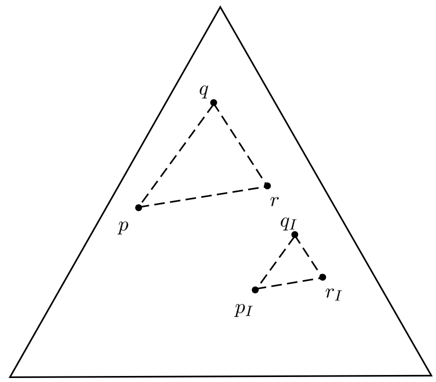


FIGURE 4: Take an arbitrary triangle formed by beliefs $p, q, r$. We know that $p_I - q_I = \varepsilon(p-q)$, $q_I - r_I = \varepsilon'(q-r)$, and, $p_I - r_I = \varepsilon''(p-r)$ some some $\varepsilon, \varepsilon'$ and $\varepsilon''$. It follows that triangle $p_I q_I r_I$ must be similar to $pqr$, which implies that $\varepsilon = \varepsilon' = \varepsilon''$.

between $q$ and $p'$. Focus on $pqp'$. Our previous argument implies that this triangle must be similar to the triangle formed by $p_I, q_I, p'_I$, because these two triangles have parallel edges (see Figure 4). Since $p_I - q_I = \varepsilon(p-q)$, $q_I - p'_I = \varepsilon_1(q-p')$, and $p'_I - p_I = \varepsilon_2(p'-p)$, it must be true that $\varepsilon = \varepsilon_1 = \varepsilon_2$. The same applies to the other triangle $pp'q'$. Since $pqp'$ and $pp'q'$ share one edge, it must be true then that $\varepsilon = \varepsilon'$. Therefore, for any $I \in \mathcal{I}$, there exist $\varepsilon(I)$ such that for any beliefs $p, q$, $p_I - q_I = \varepsilon(I)(p-q)$.

Finally, for each recommendation $I$, the parallel property implied by strong confidence monotonicity implies that the updating rule is weakly continuous. This, together with the compactness of $\Theta$ and properties of $\Delta^*$, allows us to apply the Schauder–Tychonoff fixed point theorem and show that $\pi(\cdot, I)$ has a fixed point. Denote that fixed point by $\rho(I)$. We must have

$$p_I - \rho(I)_I = \varepsilon(p - \rho(I)),$$

which implies that $p_I = \varepsilon p + (1-\varepsilon)\rho(I)$ because $\rho(I)_I = \rho(I)$. By weak long-run

than 3.

compliance, it can be shown that $\varepsilon(I) \in [0, 1)$ and $\rho(I) \in P(I)$.

# 5 Conservatism and a Negative Result

We consider another combination of the postulates. We say that an updating rule is *conservative* if for any $\{a, b\} \in \mathcal{M}$, $U(a, p) \leq U(b, p)$ implies $U(a, p_{(a,\{a,b\})}) \leq U(b, p_{(a,\{a,b\})})$. To understand this, suppose that the decision maker believes that alternative $b$ is weakly better than alternative $a$, but is recommended $a$ between $a$ and $b$. A conservative updating rule says that the decision maker holds on to her initial belief and is never convinced that $a$ is strictly better than $b$. Applying this definition inductively, it remains true even if we have a sequence of i.i.d. BBs who make the same recommendation $(a, \{a, b\})$.

**Theorem 2.** *The following statements are true:*

1. *If an updating rule satisfies weak confidence monotonicity and strong long-run compliance, then it is conservative.*

2. *If an updating rule satisfies weak confidence monotonicity, partial compliance, and strong long-run compliance, then for any $\{a, b\} \in \mathcal{M}$, $U(a, p) \leq U(b, p)$ implies that there exists $N \in \mathbb{N}$ such that $U(a, p_{(a,\{a,b\})^n}) = U(b, p_{(a,\{a,b\})^n})$ for any $n \geq N$.*

Hence, any updating rule that satisfies weak confidence monotonicity and strong long-run compliance must feature a form of conservatism. Conservatism remains if, in addition, the updating rule satisfies partial compliance, which is a fairly desirable property: As the decision maker receives $(a, \{a, b\})$, there will be a period after which the decision maker will always conclude that $a$ and $b$ are equally good.

Note that the theorem does not imply that under those postulates, if the decision maker initially believes that $b$ is weakly better than $a$, she will never change her mind ($a$ is strictly better than $b$). Again, using our binary-state example, suppose for some $\{a, b, c\} \in \mathcal{M}$, we have $P(a, \{a, b\}) = [0, 1/2]$, $P(b, \{a, b\}) = [1/2, 1]$, and $P(a, \{a, b, c\}) = [0, 1/3]$. If the decision maker's initial prior belief is in $[1/2, 1]$, but keeps receiving the recommendation $(a, \{a, b, c\})$, weak long-run compliance (implied by strong long-run compliance) requires that her posterior will eventually be strictly less than $1/2$. This is because $(a, \{a, b, c\})$ favors $a$ more strongly than $(a, \{a, b\})$ does.

Conservative updating rules that satisfy the three postulates also appear in Zhao (2021) and Dominiak et al. (2021), in which the decision maker always fully trusts the source of information and chooses her posterior by minimizing a certain measure of divergence from her prior.[10] If the divergence measure is well behaved, such a minimization procedure necessitates that the decision maker sets two actions to be equally as good if she believes one is weakly better but receives an opposite recommendation.

Specifically, Zhao (2021) considers the Kullback–Leibler divergence and shows that if the decision maker receives a non-contradicting set of qualitative probabilistic statements cyclically, her limiting beliefs must be consistent with all statements. More broadly, in the convex optimization literature, the method of cyclical projections is widely used to find a feasible solution satisfying some convex constraints.[11] In particular, it is shown that projecting cyclically onto a finite collection of convex and closed sets yields a point in their intersection asymptotically.[12] We view Theorem 2

[10] The updating rules used in those papers are slightly different from ours. In their papers, information that the decision maker uses to update her belief takes the form of a set of beliefs, rather than a pair of a recommended act and a menu as in our case. Ignoring technical differences, their updating rule replaces the recommendation $I$ in our updating rule with $P(I)$.

[11] See Bauschke, Borwein, and Lewis (1997) for a survey.

[12] See von Neumann (1950), Theorem 13.7; Halperin (1962); Bregman (1966).

as a partial converse of the aforementioned results, since it implies that conservatism is, to a certain extent, necessary for the asymptotic properties in the spirit of long-run compliance to hold.

One may wonder what happens if one puts strong confidence monotonicity, partial compliance, and strong long-run compliance together. It turns out that two of them are incompatible with each other. We have to make a choice between them.

**Theorem 3.** *Suppose $|\Theta| \geq 3$ and $\mathcal{A}$ is rich. There does not exist any updating rule that satisfies strong confidence monotonicity and strong long-run compliance.*

To see the intuition behind the impossibility result, recall that weak confidence monotonicity and strong long-run compliance implies conservativeness. When weak confidence monotonicity is replaced by strong confidence monotonicity, a stronger form of conservativeness will result: For any $a, b \in A \in \mathcal{M}$ such that $(a, A)$ and $(b, A)$ are compatible, $U(a, p) \leq U(b, p)$ implies $U(a, p_{(a,A)}) \leq U(b, p_{(a,A)})$. In other words, even recommendations that favor $a$ more strongly than $(a, \{a, b\})$ cannot change the decision maker's belief that $b$ is weakly better than $a$. This strong form of conservatism is in fact contradictory to weak long-run compliance. See Figure 5 for the argument when $|\Theta| = 3$.

However, it is possible that a contraction rule is conservative, as illustrated in the result below. The proof is trivial and thus omitted.

**Theorem 4.** *A contraction rule $(\varepsilon, \rho)$ is conservative if and only if $U(a, \rho(a, \{a, b\})) = U(b, \rho(a, \{a, b\}))$ for any $a, b \in \mathcal{A}$ such that $\{a, b\} \in \mathcal{M}$.*

Therefore, given a contraction rule, conservatism further requires that whenever the menu is binary, the recommended beliefs selected by $\rho$ must be the ones that imply that the two acts in the menu are indifferent to each other.

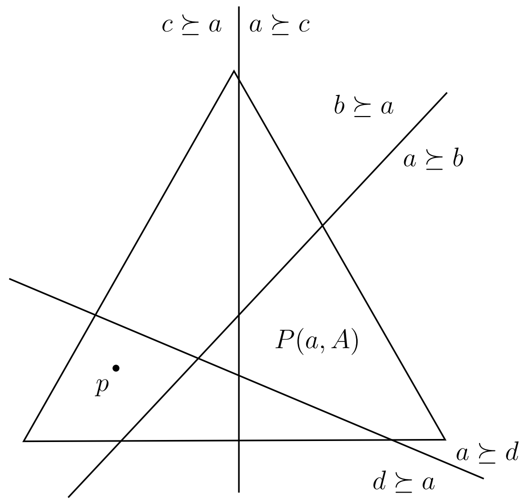


FIGURE 5: The strong form of conservatism requires that starting from belief $p$, upon receiving $(a, A)$ in which $A = \{a, b, c, d\}$ repeatedly, the decision maker never cross the line that indicates $a$ and $b$ are equally good, or the line that indicates $a$ and $d$ are equally good. Thus, the decision maker's belief can never get into the $P(a, A)$ region, violating weak long-run compliance.

## 5.1 A Subjectively Bayesian Perspective

Can the decision maker's updating rule be represented by Bayes' rule with a subjective joint prior over the state and the recommendation? As explained in the Introduction, this is not the direction that we think our theory should be going, but our primitive does not rule out this possibility. Theorem 3, however, implies that our postulates rule it out.

To see why Bayes' rule is ruled out, consider the following example. Let $\Theta = \{0, 1\}$, $a_L = \mathbf{1}_{\theta=1}, a_R = \mathbf{1}_{\theta=0}$, $a_M = 0.6$, and $I = (a_M, \{a_L, a_M, a_R\})$.[13] It is clear that $\{a_L, a_M, a_R\} \in \mathcal{M}$ and $P(I) = [0.4, 0.6]$. Suppose the decision maker has a subjective model of how recommendations are generated in each state, i.e., she posits some conditional distribution of recommendations for each state, denoted as $\mu(\cdot|\theta = 1)$ and $\mu(\cdot|\theta = 0)$. Under Bayes' rule, if the decision maker observes recommendation $I$

[13] Although Theorem 3 requires $|\Theta| \geq 3$, Bayes' rule is ruled out when $|\Theta| \geq 2$. Alternatively, $\Theta$ can be viewed as a binary partition of a higher dimensional state space.

repeatedly from i.i.d. sources, the posterior likelihood ratio after $n$ iterations will be given by

$$\frac{p_{I^n}(1)}{p_{I^n}(0)} = \frac{p(1)}{p(0)} \cdot \left(\frac{\mu(I|\theta=1)}{\mu(I|\theta=0)}\right)^n . \tag{1}$$

As $n \to \infty$, the posterior likelihood in equation (1) can only converge to 0, $\infty$, or, in the case that $I$ is uninformative, her prior likelihood ratio $p(1)/p(0)$. In other words, if $I$ is informative, Bayes' rule implies that the decision maker's long-run belief will concentrate on either state 0 or state 1. Since neither state is consistent with $I$, under weak long-run compliance, the decision maker will not be convinced by $I$ in the limit even though she receives $I$ repeatedly. Thus, weak long-run compliance is violated.

More broadly, under Bayes' rule, as the decision maker keeps receiving signals from i.i.d. sources, the subjective uncertainty will vanish in the long run. While this may be a desirable feature in standard models, in our context, due to the nature of the BBs, it is unlikely that the decision maker will be confident enough to rely on her subjective model and eliminate all the uncertainty regarding how the recommendations are generated. When the decision maker is not fully confident about her subjective model, it seems a reasonable and natural rule of thumb to follow the recommendation and settle on some beliefs that do not eliminate all the uncertainty.

Strong confidence monotonicity also has some implication that differs from Bayes' rule. Under Bayes' rule, when information (i.e., an event) arrives, the decision maker's interpretation of the information (i.e., her conditional belief given the event) depends on her prior belief. Under the contraction rule, the decision maker's prior belief also influences her posterior, but it does so differently—the decision maker's posterior is formed by mixing her prior belief with some recommended belief. The recommended belief itself depends on the recommendation, but not on the prior belief. Roughly speaking, this means that how the decision maker interprets the received information is independent of her prior. This is different from Bayes' rule, and is an implication from strong confidence monotonicity. This feature is shared with the DeGroot (1974) social learning model, and one of the main reasons that the DeGroot model is a much more tractable rule of thumb than Bayes' rule.

# 6 Belief Divergence

We use simple applications below to illustrate some implications of the contraction rule. We show that in many situations, the beliefs of the decision makers who adopt the contraction rule will diverge in the long run as they receive alternating recommendations. The divergence of beliefs in these situations is not because the decision makers are stuck in echo chambers, as in many existing theories, but because the contraction rule may violate strong long-run compliance as the decision makers receive alternating recommendations.

We will use the following setup throughout the section. Consider a decision maker who has a contraction rule $(\varepsilon, \rho)$, and whose initial belief is $p \in \Delta(\Theta)$. There is an i.i.d. sequence of BBs. In each period, a BB recommends an act from $\{a_L, a_R\} \in \mathcal{M}$ to the decision maker. For any $j \in \{L, R\}$, we use $-j$ to denote the other element in $\{L, R\}$ that is not $j$. Let $I_j = (a_j, \{a_j, a_{-j}\})$ for any $j \in \{L, R\}$. Let $\varepsilon_j = \varepsilon(I_j)$ and $\rho_j = \rho(I_j)$ for any $j \in \{L, R\}$.

Fixing any $j \in \{L, R\}$, consider $\{I_n\}_{n=1}^{\infty}$ such that $I_{2k-1} = I_j$ and $I_{2k} = I_{-j}$ for any $k \geq 1$. Such a sequence will be called the $RL$ sequence if $a_R$ is recommended in all odd periods. Otherwise, it will be called the $LR$ sequence. Let $p_n = p_{I_1 I_2 \dots I_n}$. Standard calculations imply that

$$\lim_{k\to\infty} p_{2k-1} = \frac{(1-\varepsilon_j)\rho_j + \varepsilon_j(1-\varepsilon_{-j})\rho_{-j}}{1-\varepsilon_j\varepsilon_{-j}}$$

and

$$\lim_{k\to\infty} p_{2k} = \frac{(1-\varepsilon_j)\varepsilon_{-j}\rho_j + (1-\varepsilon_{-j})\rho_{-j}}{1-\varepsilon_j\varepsilon_{-j}}.$$

It can be seen that in the long run, the decision maker's posterior oscillates between the two limits above.

Sometimes to distinguish between the $RL$ sequence and the $LR$ sequence, we use $\{p_n\}_{n=1}^{\infty}$ to denote the sequence of the decision maker's posteriors under the $RL$ sequence, and $\{q_n\}_{n=1}^{\infty}$ to denote the sequence of the decision maker's posteriors under the $LR$ sequence.

## 6.1 Divergence Due to Persistent Recency Bias

Recency bias describes the situation in which the decision maker favors recent information over past information. Given a sequence of recommendations from i.i.d. sources, under Bayes' rule, the order in which recommendations are presented to the decision maker does not matter. This is not true under the contraction rule.

To see this in a simple example, consider a binary state space and suppose that $\varepsilon_L = \varepsilon_R = \varepsilon \in (0,1)$, $\rho_L = 0$, and $\rho_R = 1$. The recency bias can be seen from the first two recommendations of the $RL$ sequence and the $LR$ sequence. For the $RL$ sequence, we have

$$p_1 = \varepsilon p + (1-\varepsilon) \text{ and } p_2 = \varepsilon p_1 = \varepsilon^2 p + \varepsilon(1-\varepsilon).$$

For the $LR$ sequence, we have

$$q_1 = \varepsilon p \text{ and } q_2 = \varepsilon q_1 + (1-\varepsilon) = \varepsilon^2 p + (1-\varepsilon).$$

In both cases, the decision maker receives $I_L$ once and $I_R$ once. However, the order in which $I_L$ and $I_R$ are received clearly matters. Moreover, the decision maker's learning behavior exhibits recency bias. For example, in the second period, we have $p_2 < q_2$. Intuitively, this is because under the $RL$ sequence, $I_L$ is more recent and hence the decision maker's posterior favors $I_L$ more, but under the $LR$ sequence, it is the opposite.

The recency bias is persistent and may be reinforced over time. Consider the full $RL$ sequence. For any $k \geq 0$,

$$p_{2k+1} = \varepsilon p_{2k} + (1-\varepsilon) \text{ and } p_{2k+2} = \varepsilon p_{2k+1} = \varepsilon^2 p_{2k} + \varepsilon(1-\varepsilon),$$

in which $p_0 = p$. For the full $LR$ sequence, for any $k \geq 0$,

$$q_{2k+1} = \varepsilon q_{2k} \text{ and } q_{2k+2} = \varepsilon q_{2k+1} + (1-\varepsilon) = \varepsilon^2 q_{2k} + (1-\varepsilon),$$

in which $q_0 = p$. It can be verified that the difference between $p_{2k}$ and $q_{2k}$ is increasing in $k$, and

$$\lim_{k\to\infty} q_{2k} - p_{2k} = \lim_{k\to\infty} p_{2k+1} - q_{2k+1} = \frac{1-\varepsilon}{1+\varepsilon}.$$

The observation above is true in more general settings when both of the decision makers share the same initial belief, adopt the same contraction rule, but receive the $RL$ sequence and the $LR$ sequence respectively. The proof of the next result is straightforward and hence omitted. Let $\phi = \frac{1}{1-\varepsilon_R\varepsilon_L}(1-\varepsilon_R)(1-\varepsilon_L)(\rho_R - \rho_L)$ and $\psi = p_1 - q_1 = (\varepsilon_R - \varepsilon_L)p + (1-\varepsilon_R)\rho_R - (1-\varepsilon_L)\rho_L$. Note that $\phi$ and $\psi$ are vectors in the linear space spanned by $\Delta(\Theta)$.

**Proposition 1.** *Suppose the decision maker's updating rule is a contraction rule. For any $p \in \Delta(\Theta)$, $\{a_L, a_R\} \in \mathcal{M}$, and positive integer $k$, $q_{2k} - p_{2k} = \phi(1 - \varepsilon_R^k \varepsilon_L^k)$ and $p_{2k-1} - q_{2k-1} = \psi \varepsilon_R^{k-1}\varepsilon_L^{k-1} - \phi(1 - \varepsilon_R^{k-1}\varepsilon_L^{k-1})$.*

In the general case, the recency bias increases over time in even periods as long as $\rho_R \neq \rho_L$. In other words, after each iteration of the repeated information, the belief difference between the two sequences of recommendations increases. In odd periods, the belief difference between the two sequences of recommendations may not increase over time, but will converge to $-\phi$, which again is not zero as long as $\rho_R \neq \rho_L$.

## 6.2 Divergence Due to Differences in Updating Rules

In this section, we consider two decision makers, labeled $L$ and $R$ respectively, who share the same initial belief and will both learn from the $RL$ sequence, but adopt different contraction rules.

In addition to the assumptions made at the beginning of this section, we further assume the following. The state space is binary. For both decision makers, $P(a_L, \{a_L, a_R\}) = [0, 0.5]$ and $P(a_R, \{a_L, a_R\}) = [0.5, 1]$. Decision maker $j$ uses the contraction rule $(\varepsilon^j, \rho^j)$, $j \in \{L, R\}$, such that for some $\epsilon, \gamma \in [0, 0.5]$, $\varepsilon^L_L = \varepsilon^R_R = \epsilon$, $\varepsilon^L_R = \varepsilon^R_L = 1 - \epsilon$, $\rho^L_L = 0$, $\rho^L_R = 1 - \gamma$, $\rho^R_L = \gamma$, and $\rho^R_R = 1$, in which the superscript denotes the decision maker's identity and the subscript denotes the recommendation. Note that this setting is symmetric, with decision maker $j$'s updating rule favoring state $j$ for $j \in \{L, R\}$. To be more specific, decision maker $j$ thinks that $I_j$ is more trustworthy than $I_{-j}$, and represents $I_{-j}$ with a less extreme belief compared with $I_j$.

Under these assumptions, with small $\epsilon$ and $1 - \gamma$, decision maker $j$'s contraction rule favors $I_j$ over $I_{-j}$. Simple calculations show that decision maker $R$'s posterior belief will oscillate between the following two limits

$$\lim_{k \to \infty} p^R_{2k-1} = \frac{(1 - \epsilon) + \epsilon^2 \gamma}{1 - \epsilon(1 - \epsilon)} > \lim_{k \to \infty} p^R_{2k} = \frac{(1 - \epsilon)^2 + \epsilon \gamma}{1 - \epsilon(1 - \epsilon)}.$$

The superscript of the posterior beliefs denotes the decision maker's identity. Similarly, decision maker $L$'s posterior belief will oscillate between the following two limits

$$\lim_{k\to\infty} p^L_{2k-1} = \frac{\epsilon(1-\gamma)}{1-\epsilon(1-\epsilon)} > \lim_{k\to\infty} p^L_{2k} = \frac{\epsilon^2(1-\gamma)}{1-\epsilon(1-\epsilon)}.$$

We have the following observations. The proof is straightforward and hence omitted.

**Proposition 2.** *For any $p \in \Delta(\Theta)$, the following statements are true:*

1. *For any $\varepsilon \in [0, 0.5]$, if $\gamma > \max\{\frac{1-\varepsilon(1-\varepsilon)}{2\varepsilon} - \frac{(1-\varepsilon)^2}{\varepsilon}, 1 - \frac{1-\varepsilon(1-\varepsilon)}{2\varepsilon}\}$, there exists some positive integer $K$ such that for any $k > K$, $p^R_k > 0.5$ and $p^L_k < 0.5$.*

2. *The sequence $\{p^R_{2k} - p^L_{2k}\}_k$ is increasing. The sequence $\{p^R_{2k-1} - p^L_{2k-1}\}_k$ is increasing if and only if $p \geq \frac{\varepsilon(1-\gamma)}{1-\varepsilon(1-\varepsilon)}$.*

The first statement shows that with some degree of bias embedded in the contraction rule, belief divergence occurs in the long run, regardless of the decision makers' common initial belief. The result can be equivalently stated by fixing $\gamma$ and identifying a bound for $\varepsilon$. For example, suppose $\gamma = 0$, i.e., there is no bias toward $L$ or $R$ from the $\rho$ function. We will observe belief divergence in the limit if (and only if) $\epsilon < (3-\sqrt{5})/2 \approx 0.38$, which does not seem to require a large amount of bias—the difference between the measures of trustworthiness for $I_L$ and $I_R$ can be close to 0.24.

Also notice that recommendations that are not favored by a decision maker still influence the decision maker's belief significantly, even when belief divergence happens. This can be seen from the fact that both decision makers' beliefs oscillate perpetually and diverge.

The result also says that the difference between the two decision makers' beliefs in even periods increases over time. Hence, belief divergence exacerbates over time. The difference between their beliefs in odd periods also increases over time if some additional condition is satisfied. This asymmetry comes from the fact that although both decision makers have the same initial belief and observe the same sequence of recommendations, one decision maker is first exposed to her more favored recommendation while the other is first exposed to her less favored recommendation. If the decision makers' initial beliefs are both equal to 0.5, this additional condition holds.

To summarize, under the contraction rule, even if we start with two decision makers with unbiased common initial beliefs ($p = 0.5$), and the i.i.d. BBs make recommendations that do not seem to favor any act over the other to both decision makers, the learning process itself will manifest and gradually amplify the bias embedded in the decision makers' updating rules.

# 7 Concluding Remarks

In this paper, we study a decision maker's learning behavior when she learns from a BB. A BB may be a complicated machine learning algorithm using high-dimensional datasets, or an expert whose process for generating recommendations is not understood by the decision maker.

We introduce several reasonable behavioral postulates on the decision maker's updating rule. Some of them lead to what we call the contraction rule. In the contraction rule, the decision maker reduces each recommendation from the BB to a single belief that is consistent with the recommendation, and assesses the trustworthiness of each recommendation. When she receives a recommendation, which induces a recommended belief and a measure of trustworthiness, she forms a posterior by mixing her prior with the recommended belief, weighted by the measure of trustworthiness. We apply the contraction rule to simple examples to show that it may generate persistent recency bias and belief divergence.

We study the consequence of some other combinations of the behavioral postulates, through which we identify sufficient conditions for the updating rule to exhibit a form of conservatism. We provide a negative result to show that not all of our behavioral postulates can be satisfied simultaneously.

In one of the postulates, we discuss how the decision maker learns from a sequence of i.i.d. BBs making recommendations from a fixed menu. In that case, if the decision maker has a contraction rule, the decision maker will use the same contraction rule inductively to learn from the BBs' recommendations. Our paper does not specify how the decision maker learns in a more general dynamic setting. A simple way to apply our theory to the more general dynamic setting is to again adopt the same contraction rule inductively when there is a sequence of recommendations from (not necessarily i.i.d.) BBs and the menus may change over time. Clearly, there may be more sophisticated ways to do this. There may be reasons for the decision maker to adopt different contraction rules over time. For example, if the decision maker knows that the BBs are getting better over time, she may increase the measure of trustworthiness over time. If the decision maker can to some extent verify the BBs' recommendations, she may revise the measure of trustworthiness accordingly. We leave these possibilities for future research.

# References

Baliga, S., E. Hanany, and P. Klibanoff (2013). Polarization and Ambiguity. *American Economic Review 103*(7), 3071–83.

Bansal, R. and I. Shaliastovich (2010). Confidence Risk and Asset Prices. *American Economic Review 100*(2), 537–41.

Barocas, S. and A. D. Selbst (2016). Big Data's Disparate Impact. *California Law*

*Review 104*(3), 671–732.

Bathaee, Y. (2018). The Artificial Intelligence Black Box and the Failure of Intent and Causation. *Harvard Journal of Law & Technology 31*, 889.

Bauschke, H. H., J. M. Borwein, and A. S. Lewis (1997). The Method of Cyclic Projections for Closed Convex Sets in Hilbert Space. *Contemporary Mathematics 204*, 1–38.

Bourbaki, N. (1987). *Elements of Mathematics: Topological Vector Spaces*. Springer-Verlag Berlin.

Bowen, R., D. Dmitriev, and S. Galperti (2022). Learning from Shared News. Working paper.

Bregman, L. M. (1966). The Relaxation Method of Finding the Common Point of Convex Sets and Its Application to the Solution of Problems in Convex Programming. *USSR Computational Mathematics and Mathematical Physics 7*, 200–217.

Caticha, A. (2004). Relative Entropy and Inductive Inference. In G. Erickson and Y. Zhai (Eds.), *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, Volume 707, pp. 75–96.

Chambers, C. P. and T. Hayashi (2010). Bayesian Consistent Belief Selection. *Journal of Economic Theory 145*(1), 432–439.

Chen, J. (2022). Sequential Learning under Informational Ambiguity. Working paper.

Cheng, X. (2021). Relative Maximum Likelihood Updating of Ambiguous Beliefs. *Journal of Mathematical Economics* (forthcoming).

Cheng, X. (2022). Robust Data-Driven Decisions under Model Uncertainty. Working paper.

Cobzas, S. (2006). Fixed Point Theorems in Locally Convex Spaces - the Schauder Mapping Method. *Fixed Point Theory and Applications 2006*.

Damiano, E. (2006). Choice under Limited Uncertainty. *The B.E. Journal of Theo-*

*retical Economics 6*(1), 1–35.

DeGroot, M. H. (1974). Reaching a Consensus. *Journal of the American Statistical Association 69*(345), 118–121.

DeMarzo, P. M., D. Vayanos, and J. Zwiebel (2003). Persuasion Bias, Social Influence, and Unidimensional Opinions. *Quarterly Journal of Economics 118*(3), 909–968.

Dominiak, A., M. Kovach, and G. Tserenjigmid (2021). Minimum Distance Belief Updating with General Information. Working paper.

Doshi-Velez, F. and B. Kim (2017). Towards a Rigorous Science of Interpretable Machine Learning. *arXiv: Machine Learning*.

Eguia, J. and T.-W. Hu (2022). Voter Polarization and Extremism. Working paper.

Epstein, L. G. (2006). An Axiomatic Model of Non-Bayesian Updating. *Review of Economic Studies 73*(2), 413–436.

Fudenberg, D., D. K. Levine, et al. (2014). Learning with Recency Bias. *Proceedings of the National Academy of Sciences 111*, 10826–10829.

Gilboa, I. and D. Schmeidler (1989). Maxmin Expected Utility with Non-unique Prior. *Journal of Mathematical Economics 18*(2), 141–153.

Golub, B. and M. O. Jackson (2010). Naïve Learning in Social Networks and the Wisdom of Crowds. *American Economic Journal: Microeconomics 2*(1), 112–49.

Golub, B. and M. O. Jackson (2012). How Homophily Affects the Speed of Learning and Best-response Dynamics. *Quarterly Journal of Economics 127*(3), 1287–1338.

Guidotti, R., A. Monreale, S. Ruggieri, F. Turini, F. Giannotti, and D. Pedreschi (2018). A Survey of Methods for Explaining Black Box Models. *ACM Computing Surveys 51*(5), 1–42.

Halperin, I. (1962). The Product of Projection Operators. *Acta Sci. Math.(Szeged) 23*(1), 96–99.

Jackson, M. O. (2011). An Overview of Social Networks and Economic Applications.

*Handbook of Social Economics 1*, 511–585.

Jackson, M. O. and Y. Zenou (2015). Games on Networks. In *Handbook of Game Theory with Economic Applications*, Volume 4, pp. 95–163. Elsevier.

Ke, S., C. Zhao, Z. Wang, and S.-L. Hsieh (2022). Behavioral Neural Networks. Working paper.

Kingston, J. K. (2016). Artificial Intelligence and Legal Liability. In *International Conference on Innovative Techniques and Applications of Artificial Intelligence*, pp. 269–279. Springer.

Kovach, M. (2021). Conservative Updating. Working paper.

Nunes, A., B. Reimer, and J. Coughlin (2018, 04). People Must Retain Control of Autonomous Vehicles. *Nature 556*, 169–171.

Ortoleva, P. (2012). Modeling the Change of Paradigm: Non-Bayesian Reactions to Unexpected News. *American Economic Review 102*(6), 2410–2436.

Park, H. and J. P. Tayawa (2022). Anchored Belief Updating from Recommendations. Working paper.

Pedreshi, D., S. Ruggieri, and F. Turini (2008). Discrimination-Aware Data Mining. In *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, New York, pp. 560–568. Association for Computing Machinery.

Perego, J. and S. Yuksel (2022). Media Competition and Social Disagreement. *Econometrica 90*(1), 223–265.

Pires, C. P. (2002). A Rule for Updating Ambiguous Beliefs. *Theory and Decision 53*(2), 137–152.

Ribeiro, M. T., S. Singh, and C. Guestrin (2016). "Why Should I Trust You?": Explaining the Predictions of Any Classifier. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD ’16, New York, pp. 1135–1144. Association for Computing Machinery.

Shore, J. E. and R. W. Johnson (1980). Axiomatic Derivation of the Principle of Maximum Entropy and the Principle of Minimum Cross-Entropy. *IEEE Transactions of Information Theory IT-26*, 26–27.

Skilling, J. (1988). The Axioms of Maximum Entropy. *Maximum-Entropy and Bayesian Methods in Science and Engineering 1*, 173–187.

Suleymanov, E. (2021). Robust Maximum Likelihood Updating. Working paper.

Varadarajan, V. S. (1958). Weak Convergence of Measures on Separable Metric Spaces. *Sankhyā: The Indian Journal of Statistics (1933-1960) 19*(1/2), 15–22.

von Neumann, J. (1950). Functional Operators. Volume II. The Geometry of Orthogonal Spaces. Reprint of 1933 lecture notes.

Wexler, R. (2017). When a Computer Program Keeps You in Jail: How Computers Are Harming Criminal Justice. *New York Times*.

Zhao, C. (2020). Representativeness and Similarity. Working paper.

Zhao, C. (2021). Pseudo-Bayesian Updating. *Theoretical Economics*, forthcoming.

# Appendix

PROOF OF LEMMA 1

*Proof.* Let $W(r) := U(a, r) - U(b, r) = \int_\Theta (a(\theta) - b(\theta))r(d\theta)$ for any $r \in \Delta(\Theta)$. Clearly $W$ is linear; i.e., $W(\alpha r + (1-\alpha)r') = \alpha W(r) + (1-\alpha)W(r')$ for any $r, r' \in \Delta(\Theta)$ and $\alpha \in [0, 1]$.

We prove the “if” part first. Suppose $W(p) \leq W(q)$ and $W(\gamma p + (1-\gamma)r) \geq 0$

for some $\gamma \in [0,1]$ and $r \in \Delta(\Theta)$. Thus,

$$\begin{aligned} W(\gamma q + (1-\gamma) r) &= \gamma W(q) + (1-\gamma) W(r) \\ &\geq \gamma W(p) + (1-\gamma) W(r) \\ &= W(\gamma p + (1-\gamma) r) \geq 0. \end{aligned}$$

Therefore, $W(\gamma q + (1-\gamma) r) \geq 0$.

Now we show the "only if" part. Suppose that for any $\gamma \in [0,1]$ and $r \in \Delta(\Theta)$, $W(\gamma p+(1-\gamma)r) \geq 0$ implies $W(\gamma q+(1-\gamma)r) \geq 0$. By way of contradiction, assume that $W(p) > W(q)$.

Since $\{a,b\} \in \mathcal{M}$, there exists $r, r' \in \Delta(\Theta)$ such that $W(r) < 0 < W(r')$. Suppose $W(p) = 0$. Then $W(q) < 0$ and we obtain a contradiction to $p \sqsubset_b^a q$. Suppose $W(p) < 0$; then linearity of $W$ implies that there exists $\gamma \in (0,1)$ such that

$$W(\gamma p + (1-\gamma) r') > 0 > W(\gamma q + (1-\gamma) r'),$$

which contradicts $p \sqsubset_b^a q$. Suppose $W(p) > 0$; then linearity of $W$ implies that there exists $\gamma \in (0,1)$ such that

$$W(\gamma p + (1-\gamma) r) > 0 > W(\gamma q + (1-\gamma) r),$$

which again contradicts $p \sqsubset_b^a q$. Hence, $W(p) \leq W(q)$. □

PROOF OF THEOREM 1

*Proof.* The second statement and the third statement are trivial. We will only prove the first statement. We present the "only if" part first. The proof is broken into a series of lemmas. Let the set of finite signed measures on $(\Theta, \Sigma)$ be $\Delta^*$. It is clear that $\Delta^*$ is a real vector space.

The following lemma regarding $\Delta^*$ will be useful.

**Lemma 2.** *For any $p, q \in \Delta^*$, if $p(S) = q(S)$ for any open subset $S \in \Sigma$, then $p = q$.*

*Proof.* Clearly the collection of all open sets, denoted as $\mathcal{O}$, is a $\pi$-system—it is closed under finite intersections. Let $\mathcal{E} := \{S \in \Sigma | p(S) = q(S)\}$. We know that $\mathcal{O} \subseteq \mathcal{E}$. Now we show that $\mathcal{E}$ is a $\lambda$-system. First, it is clear that $\Omega \in \mathcal{O} \subseteq \mathcal{E}$. Second, if $S \in \mathcal{E}$, then since $\Omega \in \mathcal{E}$, and $p, q$ are finite and additive, $\Omega \backslash S \in \mathcal{E}$. By the countable additivity of $p, q$, if $S_n \in \mathcal{E}$ for each $n$ and $S_i \cap S_j = \emptyset$, then

$$p\left(\bigcup_{n=1}^{\infty} S_n\right) = \sum_{n=1}^{\infty} p(S_n) = \sum_{n=1}^{\infty} q(S_n) = q\left(\bigcup_{n=1}^{\infty} S_n\right).$$

Thus, $\bigcup_{n=1}^{\infty} S_n \in \mathcal{E}$. By Dynkin's $\pi$-$\lambda$ theorem, the $\sigma$-algebra generated by $\mathcal{O}$ is a subset of $\mathcal{E}$, i.e., $\Sigma \subseteq \mathcal{E}$. Thus, $p(S) = q(S)$ for any $S \in \Sigma$. □

**Lemma 3.** *If $\mathcal{A}$ is rich, for any continuous function $\tilde{u} : \Theta \to \mathbb{R}$, there exist $a, b \in \mathcal{A}$ and $\lambda > 0$ such that $a - b = \lambda \tilde{u}$.*

*Proof.* Suppose $|\tilde{u}(\theta)| < L$ for any $\theta \in \Theta$. Let $\lambda = \frac{1}{L}$. Then $|\lambda \tilde{u}(\theta)| < 1$ for any $\theta \in \Theta$. Thus, by richness, there exists $a \in \mathcal{A}$ such that $a = \lambda \tilde{u}$. Let $b = 0$ and we are done. □

**Lemma 4.** *Suppose $\mathcal{A}$ is rich and strong confidence monotonicity holds. Then for any continuous function $\tilde{u} : \Theta \to \mathbb{R}$, $p, q \in \Delta(\Theta)$, and $I \in \mathcal{I}$, $\int_\Theta \tilde{u} dp \leq \int_\Theta \tilde{u} dq$ implies $\int_\Theta \tilde{u} dp_I \leq \int_\Theta \tilde{u} dq_I$.*

*Proof.* If $\tilde{u}$ is constant, then the condition in the lemma holds trivially. Suppose there exists $\theta, \theta' \in \Theta$ such that $\tilde{u}(\theta) > \tilde{u}(\theta')$. Then clearly there exists $\delta \in \mathbb{R}$ such that $\tilde{u}(\theta) + \delta > 0 > \tilde{u}(\theta) + \delta$. By Lemma 3, there exist $a, b \in \mathcal{A}$ and $\lambda > 0$ such that $a - b = \lambda(\tilde{u} + \delta)$. Hence, $a(\theta) - b(\theta) > 0 > a(\theta') - b(\theta')$. Clearly, there exists $p_1, p_2 \in \Delta(\Theta)$ such that $U(a, p_1) > U(b, p_1)$ and $U(a, p_2) < U(b, p_2)$ (taking the corresponding Dirac measures will suffice). Hence, $\{a, b\} \in \mathcal{M}$. By Lemma 1 and strong confidence monotonicity, $U(a, p) - U(b, p) \leq U(a, q) - U(b, q)$ implies $U(a, p_I) - U(b, p_I) \leq U(a, q_I) - U(b, q_I)$. It follows that $\int_\Theta \tilde{u}dp \leq \int_\Theta \tilde{u}dq$ implies $\int_\Theta \tilde{u}dp_I \leq \int_\Theta \tilde{u}dq_I$. $\square$

A function $\hat{u} : \Theta \to \mathbb{R}$ is a *simple function* if there exists $k \in \mathbb{N}$, $S_1, S_2, \ldots, S_k \in \Sigma$, and $\gamma_1, \gamma_2, \ldots, \gamma_k \in \mathbb{R}$ such that $\hat{u} = \sum_{i=1}^k \gamma_i \mathbf{1}_{S_i}$. A function $\hat{u} : \Theta \to \mathbb{R}$ is a *step function* if there exists $k \in \mathbb{N}$, open subsets $S_1, S_2, \ldots, S_k \in \Sigma$, and $\gamma_1, \gamma_2, \ldots, \gamma_k \in \mathbb{R}$ such that $\hat{u} = \sum_{i=1}^k \gamma_i \mathbf{1}_{S_i}$.

**Lemma 5.** *Suppose $\mathcal{A}$ is rich and strong confidence monotonicity holds. Then for any step function $\hat{u} : \Theta \to \mathbb{R}$, $p, q \in \Delta(\Theta)$, and $I \in \mathcal{I}$, $\left(\int_\Theta \hat{u}dp - \int_\Theta \hat{u}dq\right)\left(\int_\Theta \hat{u}dp_I - \int_\Theta \hat{u}dq_I\right) \geq 0$.*

*Proof.* First, note that $\Theta$ is compact metric space. Therefore, we may, without loss of generality, assume that $d(\theta, \theta') \leq 1$ for any $\theta, \theta' \in \Theta$. For any $\theta \in \Theta$ and any nonempty subset $S \subseteq \Theta$, let $d(\theta, S) = \inf_{\theta' \in S} d(\theta, \theta')$ and $d(\theta, \emptyset) = 1$. The first step is to show that $d(\cdot, S)$ is continuous for any $S \subseteq \Theta$. If $S = \emptyset$ there is nothing to prove. If $S \neq \emptyset$, for any $\theta_0 \in S$ and $\theta, \theta' \in \Theta$, we have

$$d(\theta, S) \leq d(\theta, \theta_0) \leq d(\theta, \theta') + d(\theta', \theta_0),$$

which implies that $d(\theta, S) \leq d(\theta, \theta') + d(\theta', S)$. It follows that $|d(\theta, S) - d(\theta', S)| \leq d(\theta, \theta')$, and thus $d(\cdot, S)$ is continuous.

For any open subset $S \subseteq \Theta$, let $S^n := \{\theta \in \Theta | d(\theta, \Theta \backslash S) \geq \frac{1}{n}\}$. It is clear that $S^n \subseteq S^{n+1}$ for all $n$, and $S = \bigcup_{n=1}^{\infty} S^n$. Define $u_n : \Theta \to \mathbb{R}$ as follows:

$$u_n(\theta) = \frac{d(\theta, \Theta \backslash S)}{d(\theta, \Theta \backslash S) + d(\theta, S^n)}.$$

Clearly $u_n$ is continuous. In particular, $u_n(\theta) = 0$ if $\theta \notin S$; $u_n(\theta) = 1$ if $\theta \in S^n$; $u_n(\theta) \in [0, 1]$ if $\theta \in S \backslash S^n$. Furthermore, $u_n$ converges pointwise to $\mathbf{1}_S$ and $|u_n(\theta)| \leq \mathbf{1}_S(\theta)$ for any $\theta \in \Theta$.

Consider any step function $\hat{u} = \sum_{i=1}^{k} \gamma_i \mathbf{1}_{S_i}$. Approximate each $\mathbf{1}_{S_i}$ with $u_{i,n}$ as above. For each $n$, $\sum_{i=1}^{k} \gamma_i u_{i,n}$ is continuous. Thus, by Lemma 4, for each $n$,

$$\left( \int_{\Theta} \sum_{i=1}^{k} \gamma_i u_{i,n} dp - \int_{\Theta} \sum_{i=1}^{k} \gamma_i u_{i,n} dq \right) \left( \int_{\Theta} \sum_{i=1}^{k} \gamma_i u_{i,n} dp_I - \int_{\Theta} \sum_{i=1}^{k} \gamma_i u_{i,n} dq_I \right) \geq 0.$$

Then Lebesgue's dominated convergence theorem completes the proof. □

**Lemma 6.** *Suppose $\mathcal{A}$ is rich and strong confidence monotonicity holds. Then for any $p, q \in \Delta(\Theta) \subseteq \Delta^*$, and $I \in \mathcal{I}$, there exists $\varepsilon \geq 0$ such that $p_I - q_I = \varepsilon(p - q)$.*

*Proof.* If $p = q$ there is nothing to prove. Suppose $p \neq q$. Then by Lemma 2 there exists an open subset $S$ such that $p(S) \neq q(S)$. Let $\varepsilon := \frac{p_I(S) - q_I(S)}{p(S) - q(S)}$. Again by Lemma 2, to show that $p_I - q_I = \varepsilon(p - q)$, since $p_I - q_I$ and $\varepsilon(p - q)$ are both elements of $\Delta^*$, it suffices to show that $p_I(T) - q_I(T) = \varepsilon(p(T) - q(T))$ for any open subset $T \in \Sigma$.

Consider any step function of the form $\hat{u} = \gamma \mathbf{1}_S + \mathbf{1}_T$. By Lemma 5, for any $\gamma \in \mathbb{R}$, we have

$$(\gamma(p(S) - q(S)) + p(T) - q(T))(\gamma(p_I(S) - q_I(S)) + p_I(T) - q_I(T)) \geq 0. \quad (2)$$

Suppose for some $\gamma \in \mathbb{R}$, $\gamma(p(S) - q(S)) + p(T) - q(T) = 0$ but $\gamma(p_I(S) - q_I(S)) +$

$p_I(T) - q_I(T) \neq 0$. Then, since $p(S) \neq q(S)$, there always exists $\gamma'$ which is close to $\gamma$ such that

$$(\gamma'(p(S) - q(S)) + p(T) - q(T))(\gamma'(p_I(S) - q_I(S)) + p_I(T) - q_I(T)) < 0,$$

which is a contradiction. Hence, for any $\gamma \in \mathbb{R}$, if $\gamma(p(S) - q(S)) + p(T) - q(T) = 0$, then $\gamma(p_I(S) - q_I(S)) + p_I(T) - q_I(T) = 0$.

Since $p(S) \neq q(S)$, $\gamma(p(S) - q(S)) + p(T) - q(T) = 0$ if and only if

$$\gamma = -\frac{p(T) - q(T)}{p(S) - q(S)}.$$

Thus, it must be the case that

$$-(p(T) - q(T))\frac{p_I(S) - q_I(S)}{p(S) - q(S)} + p_I(T) - q_I(T) = 0$$

which implies

$$p_I(T) - q_I(T) = \varepsilon(p(T) - q(T)). \tag{3}$$

Furthermore, equations (2) and (3) imply that

$$\varepsilon(\gamma(p(S) - q(S)) + p(T) - q(T))^2 \geq 0$$

for any $\gamma \in \mathbb{R}$. Since $p(S) \neq q(S)$, it is clear that $\varepsilon \geq 0$. □

**Lemma 7.** *Suppose $\mathcal{A}$ is rich, $|\Theta| \geq 3$, and strong confidence monotonicity holds. Then for any $I \in \mathcal{I}$, there exists $\varepsilon \geq 0$ such that $p_I - q_I = \varepsilon(p - q)$ for any $p, q \in \Delta(\Theta) \subseteq \Delta^*$.*

*Proof.* Consider any $p, q, r \in \Delta(\Theta) \subseteq \Delta^*$ that are linearly independent. Since $|\Theta| \geq 3$, such $p, q, r$ exist. By the previous lemma, suppose that $p_I - q_I = \varepsilon_1(p - q)$, $q_I - r_I = \varepsilon_2(q - r)$, and $r_I - p_I = \varepsilon_3(r - p)$. It follows that

$$0 = \varepsilon_1(p - q) + \varepsilon_2(q - r) + \varepsilon_3(r - p),$$

which, since $p, q, r$ are linearly independent, implies that $\varepsilon_1 = \varepsilon_2 = \varepsilon_3 := \varepsilon$.

Now let $p', q' \in \Delta(\Theta) \subseteq \Delta^*$ with $p' \neq q'$. By the previous lemma, there exists $\varepsilon' \geq 0$ such that $p'_I - q'_I = \varepsilon'(p' - q')$. We now show that $\varepsilon' = \varepsilon$.

First, we show that there must exist $\tilde{p} \in \{p, q, r\}$ such that $\tilde{p}, p', q'$ are linearly independent. By way of contradiction, suppose that $\tilde{p}, p', q'$ are linearly dependent for each $\tilde{p} \in \{p, q, r\}$. In other words, for each $\tilde{p} \in \{p, q, r\}$, there exists $\alpha, \beta, \gamma \in \mathbb{R}$ such that at least one of them is nonzero, and that

$$\alpha\tilde{p} + \beta p' + \gamma q' = 0.$$

Note that $p', q' \in \Delta(\Theta)$ and $p' \neq q'$ implies that $p', q'$ are linearly independent. Since $p', q'$ are linearly independent, $\alpha \neq 0$. Thus, for each $\tilde{p} \in \{p, q, r\}$, there exists $\tilde{\alpha}, \tilde{\beta} \in \mathbb{R}$ such that $\tilde{p} = \tilde{\alpha}p' + \tilde{\beta}q'$, which contradicts the fact that $p, q, r$ are linearly independent.

Without loss of generality, assume that $\tilde{p} = p$, and thus $p, p', q'$ are linearly independent. By the same argument as in the first paragraph, it follows that $p_I - p'_I = \varepsilon'(p - p')$. The last step is to show that there must exist $\tilde{q} \in \{q, r\}$ such that $p, \tilde{q}, p'$ are linearly independent. Suppose not. Since $p, p'$ are linearly independent, there exist $\alpha_1, \alpha_2, \beta_1, \beta_2 \in \mathbb{R}$ such that

$$q = \alpha_1 p + \beta_1 p'$$
$$r = \alpha_2 p + \beta_2 p',$$

which contradicts the fact that $p, q, r$ are linearly independent.

Without loss of generality, assume that $\tilde{q} = q$, and thus $p, q, p'$ are linearly independent. Again by the same argument as in the first paragraph, $p_I - q_I = \varepsilon'(p - q)$. Then $p \neq q$ implies that $\varepsilon = \varepsilon'$, which establishes the lemma. □

**Lemma 8.** *Suppose $\mathcal{A}$ is rich, $|\Theta| \geq 3$, and strong confidence monotonicity holds. Then for any $I \in \mathcal{I}$, $p_n$ converges to $p$ weakly implies that $(p_n)_I$ converges to $p_I$ weakly.*

*Proof.* By the previous lemma we know that there exists $\varepsilon \geq 0$ such that, for any $n$,

$$(p_n)_I - p_I = \varepsilon(p_n - p).$$

To establish the lemma, it suffices to show that for any continuous function $f$,

$$\int_\Theta f d(p_n)_I - \int_\Theta f dp_I = \varepsilon \left( \int_\Theta f dp_n - \int_\Theta f dp \right). \tag{4}$$

By the definition of Lebesgue integral, we only need to show (4) if $f$ is nonnegative and continuous.

Consider any simple function

$$g = \sum_{i=1}^{n} \gamma_i \mathbf{1}_{S_i}$$

in which $\gamma_i \geq 0$, $S_i \in \Sigma$, and $S_i \cap S_j = \emptyset$ for all $i, j$. It is clear that

$$\begin{aligned}\int_\Theta g d(p_n)_I - \int_\Theta g dp_I &= \sum_{i=1}^n \gamma_i[(p_n)_I(S_i) - p_I(S_i)] \\ &= \varepsilon \sum_{i=1}^n \gamma_i[p_n(S_i) - p(S_i)] \\ &= \varepsilon \left( \int_\Theta g dp_n - \int_\Theta g dp \right).\end{aligned}$$

Let $g_k$ be a sequence of simple functions such that $g_k$ converges pointwise to $f$, and $0 \leq g_k(\theta) \leq g_{k+1}(\theta)$ for all $k$ and $\theta \in \Theta$. Then

$$\int_\Theta g_k d(p_n)_I - \int_\Theta g_k dp_I = \varepsilon \left( \int_\Theta g_k dp_n - \int_\Theta g_k dp \right).$$

By Beppo Levi's monotone convergence theorem, letting $k \to \infty$, we obtain

$$\int_\Theta f d(p_n)_I - \int_\Theta f dp_I = \varepsilon \left( \int_\Theta f dp_n - \int_\Theta f dp \right)$$

and the lemma is established. $\square$

**Lemma 9.** *Suppose $\mathcal{A}$ is rich, $|\Theta| \geq 3$, and strong confidence monotonicity holds. Then for any $I \in \mathcal{I}$, there exists $q \in \Delta(\Theta)$ such that $q_I = q$. Thus, for any $I \in \mathcal{I}$, there exists $q \in \Delta(\Theta)$ and $\varepsilon \in [0, 1]$ such that*

$$p_I = \varepsilon p + (1 - \varepsilon) q$$

*for any $p \in \Delta(\Theta)$.*

*Proof.* Let $C(\Theta)$ be the set of continuous real-valued functions defined on $\Theta$. Since $\Theta$ is compact, each $f \in C(\Theta)$ is also bounded and uniformly continuous. We equip

$\Delta^*$ with the weak topology $\sigma(\Delta^*, C(\Theta))$, i.e., the topology of weak convergence. It is well known that (i) $\Delta(\Theta)$ is compact (see Varadarajan (1958), Theorem 3.4); (ii) $\sigma(\Delta^*, C(\Theta))$ is locally convex (see Bourbaki (1987), page II.40-42); (iii) $\sigma(\Delta^*, C(\Theta))$ is Hausdorff (see Bourbaki (1987), page II.41, Proposition 1; page II.43, Proposition 2; and Varadarajan (1958), Lemma 2.3).

Thus, $\Delta^*$ is a Hausdorff locally convex topological vector space, and $\Delta(\Theta)$ is a convex and compact subset of $\Delta^*$. By the previous lemma, for each $I \in \mathcal{I}$, the mapping $\pi(\cdot, I) : \Delta(\Theta) \to \Delta(\Theta)$ is continuous. By the Schauder–Tychonoff fixed point theorem (see Cobzas (2006), Theorem 2.3 for the exact version of the theorem and a proof), there exists $q \in \Delta(\Theta)$ such that $q_I = q$. Then by Lemma 7, for any $I \in \mathcal{I}$, there exists $q \in \Delta(\Theta)$ and $\varepsilon \geq 0$ such that

$$p_I = \varepsilon p + (1 - \varepsilon) q$$

for any $p \in \Delta(\Theta)$. It follows that

$$p_{I^n} = \varepsilon^n p + (1 - \varepsilon^n) q$$

for any $n \in \mathbb{N}$ and $p \in \Delta(\Theta)$.

Suppose $\varepsilon > 1$. Pick any $p \neq q$. Then there exists $S \in \Sigma$ such that $p(S) - q(S) < 0$. It follows that there exists $n$ large enough such that $p_{I^n}(S) = q(S) + \varepsilon^n(p(S) - q(S)) < 0$, which contradicts the definition of an updating rule. Hence, we conclude that $\varepsilon \in [0, 1]$. $\square$

**Lemma 10.** *Suppose $\mathcal{A}$ is rich, $|\Theta| \geq 3$, and, strong confidence monotonicity and weak long-run compliance hold. Then for any $I \in \mathcal{I}$, there exists $q \in P(I)$ and*

$\varepsilon \in [0, 1)$ *such that*

$$p_I = \varepsilon p + (1 - \varepsilon)q$$

*for any* $p \in \Delta(\Theta)$.

*Proof.* Fix $I \in \mathcal{I}$. By the previous lemma, there exists $q \in \Delta(\Theta)$ such that $q_I = q$. If $q \notin P(I)$, then the sequence $q_{I^n}$ does not have any accumulation point in $P(I)$, which contradicts weak long-run compliance. Hence, any fixed point of the mapping $\pi(\cdot, I)$ must be in $P(I)$. Hence, there exist $q \in P(I)$ and $\varepsilon \in [0, 1]$ such that

$$p_I = \varepsilon p + (1 - \varepsilon)q$$

for any $p \in \Delta(\Theta)$.

Let $I = (a, A)$ in which $a \in A \in \mathcal{M}$. By $A \in \mathcal{M}$, there exist $b \neq a$ and $p \in \Delta(\Theta)$ such that $b \in A$ and $U(b, p) > U(a, p)$. Hence $(a, A) \notin \alpha(p, A)$. If $\varepsilon = 0$, $p_{(a,A)^n} = p$ for any $n$, which contradicts weak long-run compliance. Thus, $\varepsilon \in [0, 1)$ and we are done. $\square$

Now we show the "if" part. It is easy to see that any contraction rule will satisfy weak long-run compliance.

To show strong confidence monotonicity, we first show that for any continuous function $f$ and $I \in \mathcal{I}$,

$$\int_\Theta f dp \geq \int_\Theta f dq \quad \Rightarrow \quad \int_\Theta f dp_I \geq \int_\Theta f dq_I. \tag{5}$$

Note that any contraction rule satisfies

$$p_I - q_I = \varepsilon(I)(p - q)$$

for any $p, q \in \Delta(\Theta)$. Using the same argument as the one for (4), we have

$$\int_\Theta f dp_I - \int_\Theta f dq_I = \varepsilon(I) \left( \int_\Theta f dp - \int_\Theta f dq \right),$$

which implies (5).

Now suppose $p \sqsubset_b^a q$. If $\{a, b\} \in \mathcal{M}$, by Lemma 1, we have $\int_\Theta (a(\theta) - b(\theta)) dp \leq \int_\Theta (a(\theta) - b(\theta)) dq$. It follows from (5) that $\int_\Theta (a(\theta) - b(\theta)) dp_I \leq \int_\Theta (a(\theta) - b(\theta)) dq_I$. Applying Lemma 1 again yields $p_I \sqsubset_b^a q_I$. Thus, strong confidence monotonicity holds. $\square$

Proof of Theorem 2

*Proof.* We first establish a useful lemma.

**Lemma 11.** *If each accumulation point $q$ of $\{p_n\}_{n=1}^\infty \subseteq \Delta(\Theta)$ satisfies $(a, A), (b, A) \in \alpha(q, A)$, then $U(a, p_n) - U(b, p_n)$ converges to 0.*

*Proof.* Suppose there exists $\varepsilon > 0$ and a subsequence $\{p_{n_j}\}_{j=1}^\infty$ such that $|U(a, p_{n_j}) - U(b, p_{n_j})| \geq \varepsilon$ for any $j$. Clearly $K := \{p \in \Delta(\Theta) \mid \left|\int_\Theta (a(\theta) - b(\theta)) p(d\theta)\right| \geq \varepsilon\}$ is closed, since $a - b$ is bounded and continuous. Furthermore, since $\Theta$ is compact, $\Delta(\Theta)$ is also compact (see Varadarajan (1958), Theorem 3.4). Hence $K$ is also compact (and sequentially compact, since $\Delta(\Theta)$ with the topology of weak convergence is metrizable). Thus, $\{p_{n_j}\}_{j=1}^\infty$ has a subsequence that converges weakly to some point in $K$. It follows that $\{p_n\}_{n=1}^\infty$ has an accumulation point in $K$, denoted as $q$. Clearly, $(a, A), (b, A)$ cannot both be in $\alpha(q, A)$, a contradiction. $\square$

Now we show the first statement. Suppose the updating satisfies weak confidence monotonicity and strong long-run compliance. Let $\{a, b\} \in \mathcal{M}$, $I = (a, \{a, b\})$, and $J = (b, \{a, b\})$. Define $W(q) = U(a, q) - U(b, q)$ for all $q \in \Delta(\Theta)$. We want to show that $W(p) \leq 0$ implies $W(p_I) \leq 0$. We prove by contradiction. Suppose $W(p) \leq 0$ but $W(p_I) > 0$. Consider $p_{IJ}$. There are two cases: $W(p_{IJ}) \geq W(p)$ or $W(p_{IJ}) < W(p)$.

Suppose $W(p_{IJ}) \geq W(p)$. We show that $W(p_{(IJ)^n}) \geq W(p)$ and $W(p_{(IJ)^n I}) \geq W(p_I) > 0$ for any $n \geq 1$. By weak confidence monotonicity and Lemma 1, we have $W(p_{IJI}) \geq W(p_I)$, and thus the claim holds for $n = 1$. Suppose $W(p_{(IJ)^k}) \geq W(p)$ and $W(p_{(IJ)^k I}) \geq W(p_I)$ for some $k \geq 1$. Then $W(p_{(IJ)^k I}) \geq W(p_I)$ and weak confidence monotonicity with respect to $J$ together yields $W(p_{(IJ)^{k+1}}) \geq W(p_{IJ}) \geq W(p)$. Then, applying weak confidence monotonicity with respect to $I$ yields $W(p_{(IJ)^{k+1}I}) \geq W(p_{IJI}) \geq W(p_I)$, which establishes the claim. Since $W(p_{(IJ)^n I}) \geq W(p_I) > 0$ for any $n \geq 1$, $W(p_{(IJ)^n I})$ does not converge to 0 as $n \to \infty$. By Lemma 11, this is a contradiction to strong long-run compliance.

Now suppose $W(p_{IJ}) < W(p)$. Since $W(p) \leq 0$, we have $W(p_{IJ}) < 0$. We show that $W(p_{(IJ)^n}) \leq W(p_{IJ})$ and $W(p_{(IJ)^n I}) \leq W(p_I)$ for any $n \geq 1$. It is clear that by weak confidence monotonicity and Lemma 1, $W(p_{IJI}) \leq W(p_I)$, and thus the claim holds for $n = 1$. Suppose $W(p_{(IJ)^k}) \leq W(p_{IJ}) < 0$ and $W(p_{(IJ)^k I}) \leq W(p_I)$ for some $k \geq 1$. Then, $W(p_{(IJ)^k I}) \leq W(p_I)$ and weak confidence monotonicity with respect to $J$ together yield $W(p_{(IJ)^{k+1}}) \leq W(p_{IJ})$. Then, applying weak confidence monotonicity with respect to $I$ yields $W(p_{(IJ)^{k+1}I}) \leq W(p_{IJI}) \leq W(p_I)$, which establishes the claim. Since $W(p_{(IJ)^n}) \leq W(p_{IJ}) < 0$ for any $n \geq 1$, $W(p_{(IJ)^n})$ does not converge to 0 as $n \to \infty$. By Lemma 11, this is a contradiction to strong long-run compliance.

Next, we show the second statement. Let $\{a, b\} \in \mathcal{M}$. By partial compliance, there exists $q \in \Delta(\Theta)$ such that $W(q) < 0$ but $W(q_I) \geq 0$. By the first statement of the theorem, it must be the case that $W(q_I) = 0$. Now consider any $p \in \Delta(\Theta)$ such that $W(p) \leq 0$. By weak long-run compliance, which is implied by strong long-run compliance, there exists $M \in \mathbb{N}$ such that $W(p_{I^M}) \geq W(q)$. Since $W(p) \leq 0$, inductively applying the first statement of the theorem yields $W(p_{I^n}) \leq 0$ for any $n \in \mathbb{N}$. On the other hand, by weak confidence monotonicity and Lemma 1, $W(p_{I^{M+1}}) \geq W(q_I) = 0$. It follows that $W(p_{I^{M+1}}) = 0$. Finally, inductively applying weak confidence monotonicity yields $0 \geq W(p_{I^n}) \geq W(q_{I^{n-M}}) \geq W(q_{I^{n-M-1}}) \geq \cdots \geq W(q_I) = 0$ for any $n \geq M+1$. □

## Proof of Theorem 3

*Proof.* First, we show a useful lemma. We say that an updating rule is *strongly conservative* if for any $a, b \in A \in \mathcal{M}$ such that $(a, A)$ and $(b, A)$ are compatible, $U(a, p) \leq U(b, p)$ implies $U(a, p_{(a,A)}) \leq U(b, p_{(a,A)})$.

**Lemma 12.** *If an updating rule satisfies strong confidence monotonicity and strong long-run compliance, then it is strongly conservative.*

*Proof.* The argument is similar to the one for the first statement of Theorem 2. One only needs to set $I = (a, A)$ and $J = (b, A)$, and replace weak confidence monotonicity with strong confidence monotonicity in the argument. □

Now we proceed to construct the counterexample. First, pick distinct states $\theta_1, \theta_2, \theta_3 \in \Theta$ and let $\Delta = \{p \in \Delta(\Theta) | \text{supp}(p) \subseteq \{\theta_1, \theta_2, \theta_3\}\}$. Identify each $p \in \Delta$ with the 3-dimensional vector $(p(\theta_1), p(\theta_2), p(\theta_3))$. Clearly, there exist $x, y, z \in [-\frac{1}{4}, \frac{1}{4}]^3$ and $p_x, p_y, p_z, q \in \Delta$ such that

$$\begin{aligned} p_x \cdot x > p_x \cdot y = p_x \cdot z, \\ p_y \cdot y > p_y \cdot x = p_y \cdot z, \\ p_z \cdot z > p_z \cdot x = p_z \cdot y, \\ q \cdot x = q \cdot y = q \cdot z. \end{aligned}$$

By the Tietze extension theorem, $x, y, z$ have continuous extensions to $\Theta$, denoted respectively as $a_x, a_y, a_z$, such that the range of each of them is in $[-\frac{1}{4}, \frac{1}{4}]$. Since $\mathcal{A}$ is rich, $a_x, a_y, a_z \in \mathcal{A}$. Furthermore, by construction,

$$\begin{aligned} U(a_x, p_x) > U(a_y, p_x) &= U(a_z, p_x), \\ U(a_y, p_y) > U(a_x, p_y) &= U(a_z, p_y), \\ U(a_z, p_z) > U(a_x, p_z) &= U(a_y, p_z), \\ U(a_x, q) = U(a_y, q) &= U(a_z, q). \end{aligned}$$

Let $A = \{a_x, a_y, a_z\}$. Clearly $A \in \mathcal{M}$. Pick any arbitrary $p \in \Delta(\Theta)$ and assume without loss generality that $U(a_x, p) \geq U(a_y, p) \geq U(a_z, p)$.

Now consider a class of acts $b_\varepsilon = \frac{1}{3} \sum_{a \in A} a + \varepsilon$, in which

$$0 < \varepsilon < \frac{2}{3} \min\{U(a_x, p_x) - U(a_y, p_x), U(a_y, p_y) - U(a_z, p_y), U(a_z, p_z) - U(a_x, p_z)\}.$$

Clearly, $\varepsilon < \frac{1}{3}$. Thus, by richness, each $b_\varepsilon$ is in $\mathcal{A}$. Next, we show that $A_\varepsilon = A \cup \{b_\varepsilon\} \in \mathcal{M}$. Observe that

$$U(a_x, p_x) = \frac{1}{3} \sum_{a \in A} U(a, p_x) + \frac{2}{3}(U(a_x, p_x) - U(a_y, p_x)) > U(b_\varepsilon, p_x).$$

Similarly, we have $U(a_y, p_y) > U(b_\varepsilon, p_y)$ and $U(a_z, p_z) > U(b_\varepsilon, p_z)$. In addition, since $U(a_x, q) = U(a_y, q) = U(a_z, q)$, for any $a \in A$,

$$U(a, q) < \frac{1}{3} \sum_{\tilde{a} \in A} U(\tilde{a}, q) + \varepsilon = U(b_\varepsilon, q)$$

as desired.

Suppose the decision maker starts with $p$ and receives $I = (a_z, A_\varepsilon)$ repeatedly. Recall that $U(a_x, p) \geq U(a_y, p) \geq U(a_z, p)$. To apply the previous lemma, we need to check that (i) $(a_x, A_\varepsilon)$ and $(a_z, A_\varepsilon)$ are compatible, and (ii) $(a_y, A_\varepsilon)$ and $(a_z, A_\varepsilon)$ are compatible. We will only check (i) since the argument for (ii) is symmetric. Clearly, there exists $\eta \in (0,1)$ such that $U(a_x, \eta p_x + (1-\eta)p_z) = U(a_z, \eta p_x + (1-\eta)p_z) > U(a_y, \eta p_x + (1-\eta)p_z)$. Note that

$$U(a_x, \eta p_x + (1-\eta)p_z) > \frac{1}{3}\sum_{a \in A} U(a, \eta p_x + (1-\eta)p_z) = U(\frac{1}{3}\sum_{a \in A} a, \eta p_x + (1-\eta)p_z).$$

Thus, we can always pick $\varepsilon$ small enough such that

$$U(a_x, \eta p_x + (1-\eta)p_z) > U(\frac{1}{3}\sum_{a \in A} a + \varepsilon, \eta p_x + (1-\eta)p_z) = U(b_\varepsilon, \eta p_x + (1-\eta)p_z),$$

which implies that $(a_x, A_\varepsilon)$ and $(a_z, A_\varepsilon)$ are compatible.

By the previous lemma, the updating rule is strongly conservative. Thus, we have

$$\min\{U(a_x, p_{I^n}), U(a_y, p_{I^n})\} \geq U(a_z, p_{I^n})$$

for any $n$. Hence any accumulation point $r$ of $\{p_{I^n}\}_{n=1}^{\infty}$ satisfies $\min\{U(a_x, r), U(a_y, r)\} \geq U(a_z, r)$. It follows that

$$U(b_\varepsilon, r) = \frac{1}{3}\sum_{a \in A} U(a, r) + \varepsilon \geq U(a_z, r) + \varepsilon > U(a_z, r),$$

a contradiction to weak long-run compliance. $\square$