Non-Bayesian Updating: a Theoretical Framework

Epstein, Larry G., Jawwad Noor, and Alvaro Sandroni

Working Paper No. 518
July 2005

NON-BAYESIAN UPDATING: A THEORETICAL FRAMEWORK*

Larry G. Epstein   Jawwad Noor   Alvaro Sandroni

July 4, 2005

Abstract

This paper models an agent in a multi-period setting who does not update according to Bayes' Rule, and who is self-aware and anticipates her updating behavior when formulating plans. Choice-theoretic axiomatic foundations are provided. Then the model is specialized axiomatically to capture updating biases that reflect excessive weight given to (i) prior beliefs, or alternatively, (ii) the realized sample. Finally, the paper describes a counterpart of the exchangeable Bayesian model, where the agent tries to learn about parameters, and some answers are provided to the question "what does a non-Bayesian updater learn?"

1. INTRODUCTION

This paper models an agent in a multi-period setting who does not update according to Bayes' Rule, and who is self-aware and anticipates her updating behavior when formulating plans. Three central questions are addressed.

Are there axiomatic foundations for such a model? We provide such foundations in the form of a representation theorem for suitably defined preferences.

*Epstein and Noor are at the Department of Economics, University of Rochester, Rochester, NY 14627; lepn@ troi.cc.rochester.edu, and jwdb@troi.cc.rochester.edu. Sandroni is at Kellogg School of Management, MEDS, Northwestern University, Evanston, IL 60208, sandroni@kellogg.northwestern.edu. We have benefitted from conversations with Joseph Perktold, Werner Ploberger and especially Igor Kopylov.

A dynamic version of the (Savage or) Anscombe-Aumann theorem provides the foundation for reliance on a probability measure representing prior beliefs and for subsequent Bayesian updating of the prior belief as information arrives. We generalize this Anscombe-Aumann theorem so that *both* the prior *and* the way in which it is updated are subjective, that is, are derived from preference. The model is dynamic: consumption processes are the ultimate source of utility, and dynamic choice behavior is derived from preferences at time 0. Thus, though the model is not normative, the agent is rational in the sense of maximizing a stable, transitive and complete preference relation.

What updating rules are permitted? Our general framework is rich: just as the Savage and Anscombe-Aumann theorems provide foundations for subjective expected utility theory without restricting beliefs, the present framework imposes a specific structure for preferences without unduly restricting the nature of updating. Richness is demonstrated by axiomatic specializations that capture excessive weight given at the updating stage to (i) prior beliefs, or alternatively, (ii) the realized sample. A counterpart of the exchangeable Bayesian model, where the agent tries to learn about parameters, is also described.

To illustrate the scope of our framework, consider an agent who is trying to learn the true parameter in a set Θ. Updating of beliefs in response to observations $s_1, ..., s_t$, leads to the process of posteriors $\{\mu_t\}$ where each μ_t is a probability measure on Θ. Bayesian updating leads to the process

$$\mu_{t+1}(\cdot) = BU(\mu_t; s_{t+1})(\cdot),$$

where $BU(\mu_t; s_{t+1})(\cdot)$ denotes the Bayesian update of μ_t, that is,

$$d\left[BU(\mu_t; s_{t+1})\right](\theta) = \frac{\ell(s_{t+1} \mid \theta)\ d\mu_t(\theta)}{\int \ell(s_{t+1} \mid \theta')\ d\mu_t(\theta')}, \tag{1.1}$$

for a given likelihood function ℓ. One alternative consistent with our model is the process

$$\mu_{t+1} = (1 - \kappa_{t+1})\, BU(\mu_t; s_{t+1}) + \kappa_{t+1}\mu_t,$$

where $\kappa_{t+1} \leq 1$. If κ_{t+1} does not depend on the latest observation s_{t+1} and if $\kappa_{t+1} \geq 0$, then the updating rule can be interpreted as attaching too much weight to prior beliefs μ_t and hence underreacting to observations. Another alternative has the form

$$\mu_{t+1} = (1 - \kappa_{t+1})\, BU(\mu_t; s_{t+1}) + \kappa_{t+1}\psi_{t+1},$$

where ψ_0 is a suitable noninformative prior and subsequent ψ_t's are obtained via Bayesian updating. This updating rule for the posteriors μ_t can be interpreted (under the assumptions for κ_{t+1} stated above) as attaching too much weight to the sample. (See Section 5 for elaboration of these examples.)

It will be evident that there are many other kinds of updating biases that can be accommodated, including biases similar to some that have been observed in experimental psychology; see Tversky and Kahneman [20] and the surveys by Camerer [2] and Rabin [16], for example. Our model does not address the experimental evidence directly, however, because the latter deals with the updating of objective probabilities, while our model is more intuitive if, as we assume, probabilities are subjective.[1]

What do non-Bayesian updaters learn? A central focus of the literature on Bayesian learning is on what is learned asymptotically and how an agent forecasts as more and more observations are available. Bayesian forecasts are eventually correct with probability 1 under the truth given suitable conditions, the key condition being absolute continuity of the true measure with respect to initial beliefs. Hence, multiple repetitions of Bayes' Rule transforms the historical record into a near perfect guide for the future. We investigate the corresponding question for non-Bayesian updaters who face a statistical inference problem and conform to one of the above noted biases. We describe simple non-Bayesian updating rules that, if repeated multiple times, will also uncover the true data generating process. However, our richer hypothesis about updating behavior permits a broader range of possibilities for what is learned in the long run. In one of our results, we show that convergence to correct forecasts holds for an agent who underreacts to observations when updating. If she overreacts then her forecasts are eventually correct with positive probability - an example shows that with positive probability she may become certain that a false parameter is true and thus converge to precise but false forecasts.

The issue of foundations for non-Bayesian updating is taken up in [5] in a three-period framework, where the agent updates once and consumption occurs

[1] See [17] and [14], for example, for models of updating for objective probabilities that address the experimental evidence. Though the associated models of preference are not made explicit, to the best of our understanding these authors assume implicitly that the agent is an expected utility maximizer who is naive in the sense of not anticipating future deviations from Bayesian updating nor the fact that today's plans may not be implemented. In contrast, our agent is sophisticated and dynamically consistent.

only at the terminal time. The model is extended here to any finite horizon. We take as the benchmark the standard specification of utility in dynamic modeling, whereby utility at time t is given by

$$U_t(c) = E_t\left[\Sigma_{\tau=t}^{T}\delta^{\tau-t}u(c_\tau)\right], \quad t = 0, 1, ..., T, \tag{1.2}$$

where $c = (c_\tau)$ is a consumption process, δ and u have the familiar interpretations and E_t denotes the expectation operator associated with a subjective prior that is updated by Bayes' Rule. Our model generalizes (1.2) to which it reduces when updating conforms to Bayes' Rule.

In common with [5], the present paper adapts the Gul and Pesendorfer [7, 8] model of temptation and self-control.[2] While these authors (henceforth GP) focus on behavior associated with non-geometric discounting, we adapt their approach to model non-Bayesian updating. The connection drawn here between temptation and updating is as follows: at period t, the agent has a prior view of the relationship between the next observation s_{t+1} and the future uncertainty $(s_{t+2}, s_{t+3}, ...)$ that she considers 'correct'. But after observing a particular realization s_{t+1}, she changes her view on the noted relationship. For example, she may respond exuberantly to a good (or bad) signal after it is realized and decide that it is an even better (or worse) signal about future states than she had thought ex ante. She tries to resist the temptation to behave in accordance with the new view rather than in accordance with the view she considers correct. Temptation might be resisted but at a cost. Thus she acts as though forming a compromise posterior belief - it differs from what would be implied by Bayesian updating of the original prior and in that sense reflects non-Bayesian updating. The exuberant agent described above would appear to an outside observer as someone who overreacts to data.

GP show that temptation and self-control are revealed through preference over menus - for instance, preferring smaller menus from which to choose ex-post is a way of dealing with (and thus may be regarded as an expression of) self-control problems . For this reason, menus play a central role in [5] and in this model as well - the agent's ranking of *contingent menus* are shown to reveal behavior that is consistent with non-Bayesian updating rules being used when choosing out of menus. However, the current model differs in one important respect from GP and [5]: while the main concern of these models is how the agent makes choices out of menus (that is, how she chooses or updates when experiencing temptation), these

[2] At a technical level, we rely heavily on generalizations of the Gul-Pesendorfer model proven by Kopylov [9].

choices are not part of their primitives. With an ex-ante preference ordering over menus/contingent menus as their only primitive, the story about ex-post preferences they adopt is only *suggested* by the representation of this ex-ante preference. As a result, it is not clear what dynamic behavior constitutes a refutation of the model. This gap is closed in the current paper by taking preferences in every period as the primitive. In this respect, our model is related to Noor [15], who provides a revealed preference characterization of stationary GP-type models by taking as a primitive a choice correspondence that describes choice out of menus in every period.

The paper proceeds as follows: Section 2 defines the formal domain of choice, the space of contingent menus, and then the functional form for utility. Section 3 provides axiomatic foundations. Section 4 illustrates the nature and scope of the model by describing axiomatic specializations that capture specific updating biases.[3] Section 5 specializes further to capture an agent who is trying to learn about parameters as in the Bayesian model with an exchangeable prior. Some results are provided concerning what is learned in the long run. Section 6 concludes. Proofs are collected in appendices.

2. UTILITY

2.1. Primitives

The model's primitives include:

- time $t = 0, 1, 2, \ldots, T + 1$
- (finite) period state space S
  full state space $\Pi_{\mathbf{t}=1}^{T+1} S_t$, where $S_t = S$ for all t
- period consumption space $C_t = C$
  compact metric mixture space[4]

[3] Readers who are more interested in the functional forms implied by our model than in their axiomatic foundations may wish to skip Sections 3 and 4 and proceed directly to Section 5; the latter is in large part self-contained.

[4] We use this term to include the property that the mixture operation $(c, c', \alpha) \longmapsto \alpha c + (1 - \alpha) c'$ is continuous with respect to the obvious product metric on $C \times C \times [0, 1]$.

Though we often refer to c_t in C_t as period t consumption, it is more accurately thought of as a lottery over period t consumption. Thus we adopt an Anscombe-Aumann style domain where outcomes are lotteries.

Information available at t is given by the history $s_1^t = (s_1, ..., s_t)$. Thus time t consumption, conditional beliefs, conditional preferences and so on, are taken to be suitably measurable, though dependence on s_1^t is often suppressed in the notation.

For any compact metric space X, the set of acts from S into X is X^S; it is endowed with the product topology. A closed (hence compact) subset of $C \times X^S$ is called a *menu* (of pairs (c, F), where $c \in C$ and $F \in X^S$). Denote by $\mathcal{M}(X)$ the set of all compact subsets of X, endowed with the Hausdorff metric. Analogously, $\mathcal{M}\left(C \times X^S\right)$ is the set of menus of pairs (c, F) as above; it inherits the compact metric property [1, Section 3.16].

At each time $t = 0, 1, ..., T$, the agent chooses both current consumption and a physical action. Actions taken at the penultimate period T are modeled by acts as in the standard model, that is, by elements of

$$\mathcal{C}_T = (C_{T+1})^{S_{T+1}} . \tag{2.1}$$

Each element of $\mathcal{C}_T$ describes random consumption at $T+1$ as a function of the realized state in S_{T+1}. Consider next actions taken at time $t < T$, where consumption at t has already been determined. The consequence of that action is a menu, contingent on the state s_{t+1}, of alternatives for $t+1$, where these alternatives include both choices to be made at $t+1$ - namely, the choice of both consumption and also another action. This motivates identifying each physical action at time t with a mapping F_t, called a *contingent menu*, where

$$F_t : S_{t+1} \longrightarrow \mathcal{M}\left(C_{t+1} \times \mathcal{C}_{t+1}\right). \tag{2.2}$$

Here $\mathcal{C}_t = \left(\mathcal{M}\left(C_{t+1} \times \mathcal{C}_{t+1}\right)\right)^{S_{t+1}}$ denotes the set of all (time t) contingent menus and these sets are defined recursively by (2.2) for $t = 0, 1, ..., T-1$, and by the terminal condition (2.1).

Each $\mathcal{C}_t$ is compact metric. It also admits a mixing operation: given any space X where mixtures $\lambda x + (1-\lambda)y$ are well defined, mix elements of $\mathcal{M}(X)$ by

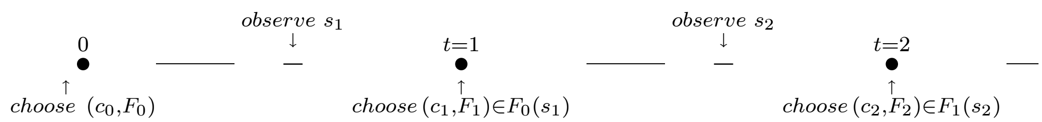
$$\lambda M + (1-\lambda) N = \{\lambda x + (1-\lambda) y : x \in M, y \in N\}.$$

Define mixtures of contingent menus recursively by: $\left(\lambda F_t' + (1-\lambda) F_t\right)(s_{t+1}) =$

$$\{\left(\lambda c_{t+1}' + (1-\lambda) c_{t+1}, \lambda H_{t+1}' + (1-\lambda) H_{t+1}\right) : \left(c_{t+1}', H_{t+1}'\right) \in F_t'(s_{t+1}), (c_{t+1}, H_{t+1}) \in F_t(s_{t+1})\},$$

where (Anscombe-Aumann) acts, elements of $\mathcal{C}_{T+1} = (C_{T+1})^{S_{T+2}}$, are mixed in the usual way. Finally, mixtures are defined componentwise on $C_t \times \mathcal{C}_t$.

The final primitive is a process of preference relations $(\succeq_t)_{t=0}^{T}$, one for each time $t \leq T$ and history s_1^t, where the domain of $\succeq_t$ is $C_t \times \mathcal{C}_t$. At time 0, the agent uses $\succeq_0$ to choose (c_0, F_0) in $C_0 \times \mathcal{C}_0$. She does this as though anticipating the following: at 1^-, a signal s_1 is realized, and this determines a menu $F_0(s_1) \subset C_1 \times \mathcal{C}_1$; at time 1, she updates and uses the order $\succeq_1$ (which corresponds to the history s_1)



to choose some (c_1, F_1) from $F_0(s_1)$. She consumes c_1 and her (contingent) options for the future are described by F_1. Continuing in this way, and given some previous choice of contingent menu F_t, she observes a signal s_{t+1}, updates and then uses the order $\succeq_{t+1}$ (corresponding to the history $(s_1, s_2, ...s_{t+1})$) to choose some (c_{t+1}, F_{t+1}) from $F_t(s_{t+1})$. (See the time line.) Her choice (c_T, F_T) from $F_{T-1}(s_T)$ at time T yields consumption c_T and a standard act $F_T : S_{T+1} \rightarrow C_{T+1}$, which describes random consumption in the terminal period $T+1$.

Our model subsumes the standard model where choice at every time t is guided by a single preference $\succeq_t$ and the evaluation of contingent menus involves valuing menus according to their best alternative. We differ in permitting the agent to exhibit a *preference for commitment* - at t she may strictly prefer a menu M_{t+1} (conditionally on some s_{t+1}) to a larger menu M'_{t+1} even though it renders infeasible the $\succeq_{t+1}$-best alternative in the latter. The idea is that $\succeq_{t+1}$ may represent choice that is affected by temptation and thus ex-ante, at t, the agent may not value menus according to their $\succeq_{t+1}$-best alternatives.

2.2. Functional Form

We describe the representation of $(\succeq_t)$. Components of the functional form include: a discount factor $0 < \delta < 1$, $u : C \longrightarrow \mathbb{R}^1$ linear, continuous and nonconstant, a probability measure p_0 on S_1 with full support, and an adapted process $(p_t,\, q_t, \alpha_t)_{t=1}^{T}$, where,[5]

[5] $\Delta(S)$ is the set of probability measures on the finite set S. A stochastic process (X_t) on $\Pi_1^\infty S_\tau$ is adapted if X_t is measurable with respect to the σ-algebra $\mathcal{S}_t$ that is generated by all sets of the form $\{s_1\} \times ... \times \{s_t\} \times \Pi_{t+1}^\infty S_\tau$. Below we often write $p_t(\cdot)$ rather than $p_t(\cdot \mid s_1^t)$.

$$\alpha_t \in (0,1]\,, \quad p_t,\, q_t \in \Delta\left(S_{t+1}\right), \text{ and each } p_t \text{ has full support.}$$

For each $(c_t, F_t) \in C_t \times \mathcal{C}_t$, define

$$\mathcal{U}_t\left(c_t, F_t\right) = u(c_t) + \delta \int_{S_{t+1}} U_{t+1}\left(F_t\left(s_{t+1}\right), s_{t+1}\right) dp_t, \text{ for } t \geq 0, \tag{2.3}$$

$$\mathcal{V}_t\left(c_t, F_t\right) = u(c_t) + \delta \int_{S_{t+1}} U_{t+1}\left(F_t\left(s_{t+1}\right), s_{t+1}\right) dq_t, \text{ for } t > 0, \tag{2.4}$$

where $U_{t+1}\left(\cdot, s_{t+1}\right) : \mathcal{M}\left(C_{t+1} \times \mathcal{C}_{t+1}\right) \longrightarrow \mathbb{R}^1$ is given by

$$U_{t+1}\left(M, s_{t+1}\right) = \max_{(c_{t+1}, F_{t+1}) \in M} \Big\{ \mathcal{U}_{t+1}\left(c_{t+1}, F_{t+1}\right) + \frac{1-\alpha_{t+1}}{\alpha_{t+1}} \left(\mathcal{V}_{t+1}\left(c_{t+1}, F_{t+1}\right) - \max_{\left(c'_{t+1}, F'_{t+1}\right) \in M} \mathcal{V}_{t+1}\left(c'_{t+1}, F'_{t+1}\right) \right) \Big\}, \tag{2.5}$$

and the boundary condition

$$U_{T+1}\left(c, s_{T+1}\right) = u\left(c\right).$$

Then $\succeq_0$ is represented by $\mathcal{U}_0(\cdot)$ and for each $t > 0$, $\succeq_t$ is represented by $\alpha_t \mathcal{U}_t(\cdot) + (1-\alpha_t)\mathcal{V}_t(\cdot)$.

The Bayesian intertemporal utility model (1.2) is specified by u, δ and a process (p_t) of one-step-ahead conditionals, which determines a unique prior on the full state space $\Pi_1^{T+1} S_t$. It is obtained as the special case where $(1-\alpha_t)(q_t - p_t) \equiv 0$ for all t. Then (2.5) reduces to

$$U_{t+1}\left(M, s_{t+1}\right) = \max_{(c_{t+1}, F_{t+1}) \in M} \mathcal{U}_{t+1}\left(c_{t+1}, F_{t+1}\right),$$

and $\succeq_t$ is represented by

$$\mathcal{U}_t\left(c_t, F_t\right) = u(c) + \delta \int_{S_{t+1}} \left(\max_{(c_{t+1}, F_{t+1}) \in F_{t+1}(s_{t+1})} \mathcal{U}_{t+1}\left(c_{t+1}, F_{t+1}\right) \right) dp_t, \qquad (c_t, F_t) \in C_t \times \mathcal{C}_t.$$

When we want to emphasize dependence on the last observation s_t, we write $p_t\left(\cdot \mid s_t\right)$. Similarly, history is suppressed notationally below when we write $\mathcal{U}_t\left(c_t, F_t\right)$ and $\mathcal{V}_t\left(c_t, F_t\right)$.

This is the standard model in the sense that it extends the model of utility over consumption processes given by (1.2) to contingent menus by assuming that menus are valued according to the best alternative they contain (a property termed *strategic rationality* by Kreps [10]). In particular, time t conditional beliefs about the future are obtained by applying Bayes' Rule to the prior on $\Pi_1^{T+1} S_t$ that is induced by the one-step-ahead conditionals (p_t).

More generally, *two* processes of one-step-ahead conditionals, p_t's and q_t's, must be specified, as well as the process of α_t's. The way in which these deliver non-Bayesian updating is explained below along with further discussion and interpretation. Sections 4 and 5 provide several examples. See also [5] for discussion in the context of a three-period model.

2.3. Interpretation

To facilitate interpretation, and also for later purposes, consider some subclasses of $\mathcal{C}_t$. The contingent menu F_t provides commitment for the next period if $F_t(s_{t+1})$ is a singleton for each s_{t+1}. Proceeding recursively in the obvious way, we can define the set of contingent menus that provide commitment for *all* future periods - denote it by $\mathcal{C}_t^c \subset \mathcal{C}_t$. Each F_t in $\mathcal{C}_t^c$ determines a unique (random variable) consumption process $c^{F_t} = \left(c_\tau^{F_t}\right)_{\tau \geq t}$. If each $c_\tau^{F_t}$ is measurable with respect to information at time $t+1$, then all uncertainty is resolved next period - the set of all such contingent menus is $\mathcal{C}_t^{c,+1} \subset \mathcal{C}_t^c$. An example is a (one-step-ahead) *bet on the event* $G \subset S_{t+1}$, which pays off with a good deterministic consumption stream if the state next period lies in G and with a poor one otherwise.

Compute that for any c_t and contingent menu F_t that provides commitment $(F_t \in \mathcal{C}_t^c)$,

$$\mathcal{U}_t(c, F) = u(c_t) + \delta \int_{S_{t+1}} U_{t+1}(F_t(s_{t+1}), s_{t+1}) \; dp_t(s_{t+1}).$$

It follows that if $F \in \mathcal{C}_0^c$, then

$$\mathcal{U}_0(c, F) = \int_{S_1 \times S_2 \times \ldots} \left(\Sigma_1^{T+1} \delta^{t-1} u\left(c_t^F\right)\right) dP_0(\cdot),$$

where c^F is the consumption process induced by F as just explained, and where $P_0(\cdot)$ is the unique measure on $\Pi_1^{T+1} S_t$ satisfying

$$P_0(s_1, s_2 \ldots, s_{T+1}) = p_0(s_1) \times \ \ldots \times \ p_t(s_{t+1} \mid s_1^t) \times \ \ldots \times p_T(s_{T+1} \mid s_1^T). \quad (2.6)$$

Thus $\succeq_0$ restricted to $\mathcal{C}_0^c$ conforms to subjective expected (intertemporally additive) utility with prior P_0. To interpret P_0 further, consider its one-step-ahead conditionals p_t for $t \geq 1$. Because these conditional beliefs are formed with the detachment and objectivity afforded by an ex ante stage ($t = 0$), the agent views them as 'correct'.[6] She will continue to view them as correct as time passes. If she were not subject to other influences, her posterior at t would be

$$P_t\left(s_{t+1}, s_{t+2}..., s_{T+1} \mid s_1^t\right) = p_t(s_{t+1} \mid s_1^t) \times \; ... \times p_T(s_{T+1} \mid s_1^T). \tag{2.7}$$

the Bayesian update of P_0. However, as explained shortly, she may update differently and be led to different posteriors.

Her actual updating underlies the preference $\succeq_t$ prevailing after an arbitrary history s_1^t. By assumption, $\succeq_t$ is represented by $\alpha_t \mathcal{U}_t(\cdot) + (1 - \alpha_t)\mathcal{V}_t(\cdot)$. To proceed, define the one-step-ahead conditional measure m_t by

$$m_t(s_{t+1}) = m_t(s_{t+1} \mid s_1^t) = \alpha_t p_t(s_{t+1} \mid s_1^t) + (1 - \alpha_t) q_t(s_{t+1} \mid s_1^t).$$

Next compute that for any c_t and contingent menu $F_t \in \mathcal{C}_t^c$,

$$\alpha_t \mathcal{U}_t(c_t, F_t) + (1 - \alpha_t)\mathcal{V}_t(c_t, F_t) =$$

$$u(c_t) + \delta \int_{S_{t+1}} U_{t+1}\left(F_t(s_{t+1}), s_{t+1}\right) \, dm_t(s_{t+1}) =$$

$$\int_{S_{t+1} \times S_{t+2} \times ...} \left(\Sigma_{t+1}^{\infty} \delta^{\tau - t - 1} u\left(c_\tau^{F_t}\right)\right) dQ_t\left(\cdot \mid s_1^t\right),$$

where $Q_t(\cdot \mid s_1^t)$ is the unique measure on $\Pi_{t+1}^{T+1} S_\tau$ satisfying

$$Q_t\left(s_{t+1}, s_{t+2}..., s_{T+1} | s_1^t\right) = m_t(s_{t+1} \mid s_1^t) \times p_{t+1}(s_{t+2} \mid s_1^{t+1}) \times ... \times p_T(s_{T+1} \mid s_1^T). \tag{2.8}$$

Evidently, at t the agent's behavior (at least within $\mathcal{C}_t^c$) corresponds to the posterior $Q_t(\cdot \mid s_1^t)$, and this differs from the period 0 perspective $P_t(\cdot \mid s_1^t)$. Note that Q_t is not the Bayesian update of P_0, nor is it the Bayesian update of Q_{t-1}. The difference between P_t and Q_t lies in the way that one-step-ahead beliefs over S_{t+1} are formulated - the conditional one-step-ahead belief actually adopted at t is $m_t(\cdot)$, whereas the one that seems appropriate from the perspective of the initial period is $p_t(\cdot)$.

[6] Since p_0 is not relevant to the subsequent response to signals, its interpretation is less important here. See the comments at the end of the section.

The behavioral meaning of m_t is sharper if we restrict attention to contingent menus in $\mathcal{C}_t^{c,+1}$ (providing perfect commitment and such that all uncertainty resolves at $t+1$). Then beliefs about states in $S_{t+2} \times ... \times S_{T+1}$ are irrelevant - conclude that m_t guides the ranking of such contingent menus, for example, it guides the ranking of bets on S_{t+1}. Because the ranking of one-step-ahead bets, and more specifically the way in which it depends on past observations, is a common and natural way to understand updating behavior, we refer to m_t frequently below when considering more specific models.

The story underlying the noted difference between P_t and Q_t is as follows: consider the evaluation of a pair (c_t, F_t) in $C_t \times \mathcal{C}_t$ after having observed the history s_1^t. The functions $\mathcal{U}_t$ and $\mathcal{V}_t$ describe two ways that (c_t, F_t) may be evaluated. Both evaluate immediate consumption c_t in the same way, and they discount the expected utility of the contingent menu F_t in the same way as well. However, they disagree on how to compute the expected utility of F_t: $\mathcal{U}_t$ uses p_t and $\mathcal{V}_t$ uses q_t. The former is the 'correct' one-step-ahead conditional. But in our model, after having observed s_1^t, the agent changes her view of the world to the one-step-ahead conditional q_t. For instance, if s_1^t represents a run of bad signals, she may believe that the likelihood of another bad state is higher than her ex-ante assessment. Alternatively, she may feel that a good signal 'is due' and assign it a higher conditional probability than she did when anticipating possibilities with the cool-headedness afforded by temporal distance. Thus there are conflicting incentives impinging on the agent at t. The period 0 perspective calls for maximizing $\mathcal{U}_t$, but having seen the sample history s_1^t and having changed her view of the world, she is tempted to maximize $\mathcal{V}_t$. Resisting temptation is costly and she recognizes that the time 0 perspective is 'correct'. She is led to compromise and to maximize $\alpha_t \mathcal{U}_t(\cdot) + (1-\alpha_t)\mathcal{V}_t(\cdot)$, the utility function representing $\succeq_t$. Note that corresponding behavior is as though she used the *compromise* one-step-ahead conditional $\alpha_t p_t + (1-\alpha_t) q_t$, which is just m_t. The parameter α_t captures her ability to resist temptation.

The cost of self-control incurred when compromising between $\mathcal{U}_t(\cdot)$ and $\mathcal{V}_t(\cdot)$ is reflected not in the representation of $\succeq_t$, but rather in that of $\succeq_{t-1}$, specifically in the utility of a menu $M_t \in \mathcal{M}(C_t \times \mathcal{C}_t)$ given by the function $U_t(M_t, s_t)$. The nonpositive term

$$\frac{1-\alpha_t}{\alpha_t}\left[\mathcal{V}_t(c_t, F_t) - \max_{(c_t', F_t') \in M_t} \mathcal{V}_t(c_t', F_t')\right] \leq 0,$$

appearing in (2.5) can be interpreted as the utility cost of self-control. Thus (2.5) states that for any menu M_t received after the history s_1^t, $U_t(M_t, s_t)$ is the maximum over M_t of $\mathcal{U}_t(\cdot)$ net of self-control costs. Observe that this maximization is equivalent to

$$\max_{(c_t,F_t)\in M_t} \left\{ \mathcal{U}_t(\cdot) + \frac{1-\alpha_t}{\alpha_t} \mathcal{V}_t(\cdot) \right\},$$

and that $\mathcal{U}_t(\cdot) + \frac{1-\alpha_t}{\alpha_t} \mathcal{V}_t(\cdot)$ represents $\succeq_t$. Thus (2.5) suggests that choosing the $\succeq_t$-best element in M_t involves incurring a utility cost of self-control.

Our agent is self-aware and forward looking - she anticipates at period 0 that she will later adopt conditional beliefs different from those that seem correct now; similarly, at any time t she anticipates that she will later deviate from her current view of conditional likelihoods. Thus she may value commitment: a smaller menu may be strictly preferable because it could reduce self-control costs. In spite of the value of commitment, the above constitutes a coherent model of dynamic choice. Unlike the case in the modeling approach growing out of Strotz [18], there is no need to add assumptions about how the agent resolves her intertemporal inconsistencies. If you like, these resolutions are already embedded in her utility function defined on contingent menus. This aspect of the model uses the insight of GP.

However, there is an important difference from GP in terms of the primitives of the model. The primitive adopted by GP, and also by [5], is a single preference ordering that describes choices at one point in time. Yet these are meant to be models of dynamic choice - a story about choices in subsequent periods is "suggested" by the primitive preference, and in particular, its representation. But a characterization of a model in terms of behavior in a single period does not constitute foundations for the story about dynamic choice; it begs the question how one can use temporal data to refute the story. Such a question does not arise in our model since our primitive consists of in principle observable preferences in each period. Consequently, by characterizing our agent in terms of choices in each period (see the next section), we are able to provide 'complete' foundations for a model of dynamic choice.

Finally, a comment on the asymmetry in the representations of $\succeq_0$ and $\succeq_t$ for $t > 0$ is in order. The utility function $\alpha_t \mathcal{U}_t(\cdot) + (1-\alpha_t)\mathcal{V}_t(\cdot)$ for $t > 0$, makes explicit the conflict experienced by the agent in forming the belief m_t over S_{t+1}. The representation $\mathcal{U}_0(\cdot)$ for $\succeq_0$ is neutral in this regard: it says nothing beyond the fact that at 0 the agent has some belief p_0 over S_1, which may or may not have been formed after resolving some conflict. The reason for this stems from the fact that, as in GP, we take a preference for commitment as the behavioral manifestation of a conflict - the decomposition of the belief m_t into its correct 'p_t' and temptation 'q_t' components is based on preferences, in particular on attitudes towards commitment opportunities, prevailing at time $t-1$. A similar decomposition of p_0 would involve preferences in (unmodeled) periods prior to time 0. Hence the asymmetry in the representation for $t=0$ and $t>0$. The reader should note, however, that p_0 is not relevant for understanding updating behavior, and consequently, its decomposition is of little interest for our purposes.

3. AXIOMATIC FOUNDATIONS

In what follows, states s vary over S, consumption c varies over C, and unless otherwise specified, time t varies over $0, 1, ..., T$. A generic element of $C_t \times \mathcal{C}_t$ is $f_t = (c_t, F_t)$; t-subscripts will be dropped where there is no risk of confusion. Denote by $\left[G_{-s_{t+1}}, M\right]$ the contingent menu in $\mathcal{C}_t$ that yields $G\left(s'_{t+1}\right)$ if $s'_{t+1} \neq s_{t+1}$ and M otherwise. The menu M is identified with the constant contingent menu that delivers M in all states.

The first two axioms are standard.

Axiom 1 (Order). $\succeq_t$ *is complete and transitive.*

Axiom 2 (Continuity). *Both* $\{f : f \succeq_t g\}$ *and* $\{f : g \succeq_t f\}$ *are closed.*

In Section 2.1, we described a way to mix any two elements in $C_t \times \mathcal{C}_t$. Thus we can state the Independence axiom appropriate for our setting.

Axiom 3 (Independence). *For every* $0 < \lambda \leq 1$, *and all* f *and* g *in* $C_t \times \mathcal{C}_t$,

$$f \succeq_t g \Longleftrightarrow \lambda f + (1-\lambda) h \succeq_t \lambda g + (1-\lambda) h.$$

Intuition for Independence is similar to that provided in [5] for a three-period setting, and thus we do not elaborate here.

Given two contingent menus F and G in $\mathcal{C}_t$, define their union statewise, that is,

$$(F \cup G)(s) = F(s) \cup G(s).$$

The counterpart of GP's central axiom is:

Axiom 4 (Set-Betweenness). *For all* $t < T$, *states* s_{t+1}, *consumption* $c \in C_t$ *and all* F *and* G *in* $\mathcal{C}_t$ *such that* $G\left(s'_{t+1}\right) = F\left(s'_{t+1}\right)$ *for all* $s'_{t+1} \neq s_{t+1}$,

$$(c, F) \succeq_t (c, G) \implies (c, F) \succeq_t (c,\ F \cup G) \succeq_t (c, G). \tag{3.1}$$

Since immediate consumption and the outcome of states other than s_{t+1} is the same in all the above rankings, the axiom is essentially a statement about how the agent feels about receiving the menus $F(s_{t+1}), G(s_{t+1})$ or $F(s_{t+1}) \cup G(s_{t+1})$ conditional on s_{t+1}. As a statement about the ranking of menus, Set-Betweenness may be understood as the behavioral manifestation of temptation and self-control - GP show this in their setting and [5] adapts their interpretation to the domain of (three-period) contingent menus. The ranking of $(c,\ F)$ and $(c,\ F \cup G)$ reveals anticipation of temptation: the strict preference

$$(c, F) \succ_t (c, F \cup G), \tag{3.2}$$

suggests that the decision-maker prefers that some elements of $G(s_{t+1})$ not be available as an option conditional on s_{t+1}, and presumably, this preference for commitment reveals that she anticipates being tempted by some element of $G(s_{t+1})$ when choosing from the menu $F(s_{t+1}) \cup G(s_{t+1})$ conditional on s_{t+1}. For perspective, note that temptations do not exist for a standard decision-maker who evaluates a menu by its best element. In particular, she does not exhibit a preference for commitment and satisfies the stronger axiom:

$$F \succeq_t G \implies F \sim_t F \cup G$$

for all F and G that agree in all but one state s. Following Kreps [10, Ch. 13], we call this axiom *strategic rationality*.

Set-Betweenness allows us also to infer the agent's anticipated time $t+1$ choices from menus (note that in conjunction with (3.2) this allows us to infer whether she expects to succumb to temptation or to exert self-control). To illustrate, suppose that $F = \left[H_{-s_{t+1}},\ \{f\}\right]$ and $G = \left[H_{-s_{t+1}},\ \{g\}\right]$ and also that the decision-maker exhibits the preference

$$(c, \left[H_{-s_{t+1}},\ \{f\}\right]) \succ_t (c, \left[H_{-s_{t+1}},\ \{g\}\right]). \tag{3.3}$$

This ranking suggests that from the ex-ante perspective of period t, she prefers to end up with f rather than with g conditional on s_{t+1}, and in particular, that she prefers f to be chosen from $\{f, g\}$ conditional on s_{t+1}. Whether she anticipates f actually being chosen from $\{f, g\}$ is then revealed by her ranking of $(c, \left[H_{-s_{t+1}},\ \{f, g\}\right])$ and $(c, \left[H_{-s_{t+1}},\ \{g\}\right])$. For instance, if

$$(c, \left[H_{-s_{t+1}},\ \{f, g\}\right]) \succ_t (c, \left[H_{-s_{t+1}},\ \{g\}\right]), \tag{3.4}$$

then she has a strict preference for f being available ex-post, which reveals that she anticipates choosing f from $\{f, g\}$ at $t+1$. On the other hand, if she is indifferent to f being available ex-post, that is,

$$(c, [H_{-s_{t+1}},\ \{f, g\}]) \sim_t (c, [H_{-s_{t+1}},\ \{g\}]), \tag{3.5}$$

then she anticipates a weak preference at $t+1$ for choosing g from $\{f, g\}$. To see this, observe that given (3.3), (3.5) implies (3.2), which in turn implies that g is tempting. Thus, the indifference in (3.5) implies that she expects either to submit to g, or to be indifferent between submitting to g and resisting it. That is, she anticipates a weak preference for g at $t+1$.

Discussion of (3.4)-(3.5) revolved around what the decision-maker *anticipates* at time t about her choices at time $t+1$. The next axiom connects her time t expectations regarding future behavior and her actual future behavior.[7]

Axiom 5 (Sophistication). *Let* $(c, [G_{-s_{t+1}},\ \{f\}]) \succ_t (c, [G_{-s_{t+1}},\ \{g\}])$, *where* $t < T$. *Then*

$$(c, [G_{-s_{t+1}},\ \{f, g\}]) \succ_t (c, [G_{-s_{t+1}},\ \{g\}]) \Longleftrightarrow f \succ_{t+1} g.$$

The axiom states that she is sophisticated in that her expectations are correct (at least for anticipated choices out of binary menus $\{f, g\}$). To see this, start by taking f, g such that in period t she would prefer to commit to f rather than g conditionally on s_{t+1} (as in the hypothesis). As in the earlier discussion, this relationship between f and g allows us to deduce her expected $t+1$ choice out of $\{f, g\}$ from her $\succeq_t$-ranking of $(c, [G_{-s_{t+1}},\ \{f, g\}])$ and $(c, [G_{-s_{t+1}},\ \{g\}])$. Her actual choice out of $\{f, g\}$ is given by her $\succeq_{t+1}$-ranking of f and g. The axiom states that the decision-maker expects to choose f at $t+1$ if and only if she in fact chooses f at $t+1$.

Some axioms below involve the evaluation of streams of lotteries (or lottery streams), and it is convenient to introduce relevant notation at this point. Any risky consumption stream for the time period $[t+1, T+1]$, that is, where a unique (independent of states) consumption level c_τ is prescribed for each $t+1 \leq \tau \leq T+1$, may be identified with an element of $C_{t+1} \times \mathcal{C}_{t+1}$; if $t = T$, then a stream may be identified with an element of C_{T+1}. Denote by $\mathcal{L}_{t+1}$ the subset of all such risky consumption streams; a generic element is $\ell = (\ell_\tau)_{\tau=t+1}^{T+1}$.

[7] In the axiom, $\succeq_t$ and $\succeq_{t+1}$ are the preferences corresponding to histories $(s_1, ..., s_t)$ and $(s_1, ..., s_t, s_{t+1})$ respectively.

In order to obtain meaningful probabilities, a form of state independence is needed.

Axiom 6 (State Independence). *For all s_{t+1}, contingent menus F in $\mathcal{C}_{t+1}$ and $\ell', \ell \in \mathcal{L}_{t+1}$,*

$$(c,\ \{\ell'\}) \succeq_t (c,\ \{\ell\}) \Longleftrightarrow$$
$$(c, \left[F_{-s_{t+1}},\ \{\ell'\}\right]) \succeq_t (c, \left[F_{-s_{t+1}},\ \{\ell\}\right]).$$

The axiom states that the ranking of the lottery streams ℓ' and ℓ received unconditionally does not change if they are received conditionally on any specific s_{t+1} obtaining. Thus, time preferences and risk attitudes are not state-dependent.

In our model, temptation arises only because of a change in beliefs. This is reflected in the next axiom.[8]

Axiom 7 (Restricted Strategic Rationality (RSR)). *For all $t < T$, states s_{t+1}, s_{t+2}, consumption $c, c' \in C$, and contingent menus $F \in \mathcal{C}_t$ and $H, H' \in \mathcal{C}_{t+1}$ such that $H(s'_{t+2}) = H'(s'_{t+2})$ for all $s'_{t+2} \neq s_{t+2}$, if*

$$(c', \left[F_{-s_{t+1}},\ \{(c, H')\}\right]) \succeq_t (c', \left[F_{-s_{t+1}},\ \{\ (c, H)\}\right]), \tag{3.6}$$

then

$$(c', \left[F_{-s_{t+1}},\ \{\ (c, H')\}\right]) \sim_t (c', \left[F_{-s_{t+1}},\ \{(c, H'), (c, H)\}\right]) \tag{3.7}$$
$$\text{and}\ \ (c, H') \succeq_{t+1} (c, H)\,. \tag{3.8}$$

Suppose that, on observing s_{t+1}, the agent at $t+1$ has to choose from the menu $\{(c, H'), (c, H)\}$ where $H'(s'_{t+2}) = H(s'_{t+2})$ for all $s'_{t+2} \neq s_{t+2}$ for some s_{t+2}. Since H' and H differ only in the single state s_{t+2}, their ranking does not depend on beliefs over S_{t+2} - there are no trade-offs across states that must be made. Consequently, there is no temptation when choosing out of $\{(c, H'), (c, H)\}$, and, therefore, conditional on any s_{t+1}, the agent never exhibits a preference for commitment. In particular, her preference $\succeq_t$ satisfies a form of strategic rationality. This is the content of the implication '(3.6) $\Longrightarrow$ (3.7)'. The implication '(3.6) $\Longrightarrow$ (3.8)' is another expression of the absence of temptation: if the $t+1$ choice between the prospects (c, H') and (c, H) is not subject to temptation, then

[8] As in Sophistication, the preferences $\succeq_t$ and $\succeq_{t+1}$ correspond to histories $(s_1, ..., s_t)$ and $(s_1, ..., s_t, s_{t+1})$ respectively.

there is no reason for her $t+1$ perspective to deviate from her prior, time t perspective regarding the two prospects. The latter perspective is revealed by (3.6), the agent's time t preference for committing to (c, H') versus (c, H) conditionally on s_{t+1}.

The final axiom places structure on the agent's preferences over lottery streams.

Axiom 8 (Risk Preference). *There exist* $0 < \delta < 1$ *and* $u : C \longrightarrow \mathbb{R}^1$ *nonconstant, linear and continuous, such that, for each* ℓ' *and* ℓ *in* $\mathcal{L}_{t+1}$,

$$\ell' \succeq_t \ell \quad \Longleftrightarrow$$
$$\Sigma_{t+1}^{T+1}\ \delta^{\tau-(t+1)} u\left(\ell'_\tau\right) \geq \Sigma_{t+1}^{T+1}\ \delta^{\tau-(t+1)} u\left(\ell_\tau\right). \tag{3.9}$$

The axiomatic characterization of the utility function over streams of lotteries appearing in (3.9) is well known (see [4], for example). Because time and risk preferences are not our primary focus, we content ourselves with the statement of the above unorthodox 'axiom.'

3.1. Representation Result

Say that $\left(\delta, u, p_0, (\alpha_t, p_t, q_t)_{1\leq t\leq T}\right)$ *represents* $(\succeq_t)$ if $\succeq_0$ is represented by $\mathcal{U}_0(\cdot)$ and for each $t > 0$, $\succeq_t$ is represented by $\alpha_t \mathcal{U}_t(\cdot) + (1-\alpha_t)\mathcal{V}_t(\cdot)$, where these functions are defined in (2.3)-(2.5) and where u, δ, p_0 and $(\alpha_t, p_t,\ q_t)_{1\leq t\leq T}$ satisfy the properties stated there. For any $c \in C_{t+1}$ and $M \subset \mathcal{C}_{t+1}$, write (c, M) instead of $\{c\} \times M \in \mathcal{M}(C_{t+1} \times \mathcal{C}_{t+1})$.

Theorem 3.1. *The process of preferences* $(\succeq_t)$ *satisfies Axioms 1-8 if and only if there exists some* $(\delta, u, p_0, (\alpha_t, p_t, q_t)_{1\leq t\leq T})$ *representing* $(\succeq_t)$. *Moreover, if* $(\delta', u', p'_0, (\alpha'_t, p'_t, q'_t)_{1\leq t\leq T})$ *also represents* $(\succeq_t)$, *then* $\delta' = \delta$, $u' = au + b$ *for some* $a > 0$, *and*

$$p'_0 = p_0,\ \ \alpha'_t p'_t + (1-\alpha'_t) q'_t = \alpha_t p_t + (1-\alpha_t) q_t \quad \text{for } 0 < t \leq T. \tag{3.10}$$

Finally, if t *and* s_{t+1} *are such that*

$$\left(F_{-s_{t+1}}, (c, M')\right) \quad \succ_t \quad \left(F_{-s_{t+1}}, (c, M' \cup M)\right) \tag{3.11}$$

for some $c \in C_{t+1}$ *and* $M',\ M\ \subset \mathcal{C}_{t+1}$, *then*

$$\left(\alpha'_{t+1}\left(s_{t+1}\right), q'_{t+1}\left(\cdot \mid s_{t+1}\right)\right) = \left(\alpha_{t+1}\left(s_{t+1}\right), q_{t+1}\left(\cdot \mid s_{t+1}\right)\right). \tag{3.12}$$

Absolute uniqueness of all components is not to be expected. For example, if $\alpha_{t+1}(s_{t+1}) = 0$, then every measure $q_{t+1}(\cdot \mid s_{t+1})$ leads to the same s_{t+1}-conditional preference; similarly, if $q_{t+1}(\cdot \mid s_{t+1}) = p_{t+1}(\cdot \mid s_{t+1})$, then $\alpha_{t+1}(s_{t+1})$ is of no consequence and hence indeterminate. These degenerate cases constitute precisely the circumstances under which s_{t+1}-conditional preference is strategically rational, which is what is excluded by condition (3.11). Once strategic rationality is excluded, the strong uniqueness property in (3.12) obtains.

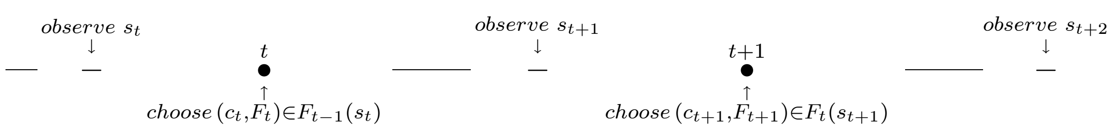
4. SOME SPECIFIC UPDATING BIASES

The framework described in Theorem 3.1 is rich. One way to see this is to focus on one-step-ahead beliefs at any time $t+1$. As pointed out in Section 2.3, these are represented by $m_{t+1} = \alpha_{t+1}p_{t+1} + (1-\alpha_{t+1})\,q_{t+1}$, while Bayesian updating of time t beliefs would lead to beliefs described by p_{t+1}. Thus, speaking roughly, updating deviates from Bayes' Rule in a direction given by $q_{t+1} - p_{t+1}$ and to a degree determined by α_{t+1}, neither of which is constrained by our framework. Consequently, the modeler is free to specify the nature and degree of the updating bias, including how these vary with history, in much the same way that a modeler who works within the Savage or Anscombe-Aumann framework of subjective expected utility theory is free to specify beliefs as she sees fit.

In this section, we go further and describe axiomatic specializations of the model that impose structure on updating. Two alternatives are explored, whereby excess weight at the updating stage is given to either (i) prior beliefs, or (ii) the sample frequency. The axioms imply restrictions on the relation between q_{t+1} and p_{t+1}, but not on α_{t+1}. Thus they limit the direction but not the magnitude of the updating bias.

4.1. Prior-Bias

At time $t > 0$, after some history s_1^t has been realized, the agent holds some view about $\Pi_{t+1}^{T+1} S_\tau$. On observing the further realization s_{t+1}, the agent at $t+1$ forms beliefs about S_{t+2} by updating this view.



Consider the possibility that at this time, she attaches inordinate weight to prior (time t) beliefs over S_{t+2}. To express this, denote by $\{(c_{t+1}, G)\}$ the contingent menu in $\mathcal{C}_t$ that assigns the singleton $\{(c_{t+1}, G)\}$ to every s_{t+1}.

Axiom 9 (Prior-Bias). *For all $t < T$, s_{t+1} and $c', c \in C$, and all $F' \in \mathcal{C}_t$, and F and G in $\mathcal{C}_{t+1}$: if*

$$(c', \ [F'_{-s_{t+1}}, \ \{(c, F)\}]) \succ_t (c', \ [F'_{-s_{t+1}}, \ \{(c, G)\}]) \text{ and} \tag{4.1}$$

$$(c', \ \{(c, F)\}) \sim_t (c', \ \{(c, G)\}), \tag{4.2}$$

then

$$(c', \ [F'_{-s_{t+1}}, \ \{(c, F)\}]) \sim_t (c', \ [F'_{-s_{t+1}}, \ \{(c, F), (c, G)\}]). \tag{4.3}$$

To interpret the axiom, we suppress the fixed consumption c' and c (and do the same for interpretations in the sequel). Condition (4.1) states that at time t, the agent strictly prefers to commit to F rather than to G conditionally on s_{t+1}. According to (4.2), she is indifferent between them when they are received unconditionally. The absence of a preference for commitment (4.3) implies that under these circumstances, she is not tempted by G conditionally on s_{t+1} at time $t+1$. Thus the absence of temptation conditionally on s_{t+1} depends not only on how F and G are ranked conditionally at time t, but also on how attractive they were at t, prior to the realization of s_{t+1}. This indicates excessive influence of prior, time t beliefs when updating at time $t+1$.

Prior-Bias begs the question what happens to temptation if the indifference in (4.2) is not satisfied. We consider two alternative strengthenings of the axiom that provide different answers.

Label by **Positive Prior-Bias** the axiom obtained when (4.2) is replaced by

$$(c', \ \{(c, F)\}) \succeq_t (c', \ \{(c, G)\}). \tag{4.4}$$

This implies that G is tempting conditionally on s_{t+1} only if it was more attractive according to (time t) prior beliefs about S_{t+2}. An alternative, labeled **Negative Prior-Bias**, is the axiom obtained when (4.2) is replaced by

$$(c', \ \{(c, F)\}) \preceq_t (c', \ \{(c, G)\}). \tag{4.5}$$

In this case, G is preferred at time t, but the signal s_{t+1} reverses the ranking in favor of F. Thus s_{t+1} is a strong positive signal for F. The agent is greatly influenced by signals. Thus she is not tempted by G after seeing s_{t+1}.

Corollary 4.1. *Suppose that* $(\succeq_t)$ *has a representation* $(\delta, u, p_0, (\alpha_t, p_t, q_t)_{1 \leq t \leq T})$. *Then* $(\succeq_t)$ *satisfies Prior-Bias if and only if*

$$q_{t+1}(\cdot \mid s_{t+1}) = (1 - \lambda_{t+1}) \, p_{t+1}(\cdot \mid s_{t+1}) + \lambda_{t+1} \left[\Sigma_{s'_{t+1}} m_t \left(s'_{t+1} \right) p_{t+1} \left(\cdot \mid s'_{t+1} \right) \right], \tag{4.6}$$

for some adapted process (λ_t) *with* $\lambda_{t+1} \leq 1$.[9]

Further, $(\succeq_t)$ *satisfies (i) Positive Prior-Bias or (ii) Negative Prior-Bias if and only if (4.6) is satisfied with respectively (i)* $0 \leq \lambda_{t+1} \leq 1$ *and (ii)* $\lambda_{t+1} \leq 0$.

To understand the result, recall from Section 2.3 that at t, after the history s_1^t, the agent's beliefs about future uncertainty are captured by the measure

$$Q_t \left(s_{t+1}, s_{t+2}, ..., s_{T+1} \mid s_1^t \right) = m_t \left(s_{t+1} \right) \, p_{t+1}(s_{t+2} \mid s_1^{t+1}) \times ... \times p_T(s_{T+1} \mid s_1^T).$$

Refer to it as the agent's prior view at t. The measure $\Sigma_{s'_{t+1}} m_t \left(s'_{t+1} \right) p_{t+1} \left(\cdot \mid s'_{t+1} \right)$ represents beliefs about S_{t+2} held at t; refer to it as the prior view of S_{t+2} at t. The measure $p_{t+1}(\cdot \mid s_{t+1})$ over S_{t+2} is the Bayesian update of the prior view at t conditional on observing s_{t+1}. Corollary 4.1 establishes that Prior-Bias is characterized by q_{t+1} being expressible as a linear combination of the Bayesian update p_{t+1} of the prior view at t and the prior view of S_{t+2} at t. The weight on the former is non-negative but the weight λ_{t+1} on the latter could be negative. This is ruled out under Positive Prior-Bias but is compatible with Negative Prior-Bias.

These functional forms for q_{t+1} support our choice of terminology. When $q_{t+1} = p_{t+1}$, updating consists of responding to data by applying Bayes' Rule to the prior view. On the other hand, using the prior view of S_{t+2} (expressed by $\Sigma_{s'_{t+1}} m_t \left(s'_{t+1} \right) p_{t+1} \left(\cdot \mid s'_{t+1} \right)$) as the posterior would give all the weight to prior beliefs and none to data. Thus an agent who updates according to the average scheme in (4.6) exhibits a positive bias to the prior if $\lambda_{t+1} > 0$ and a negative one if $\lambda_{t+1} < 0$.

Though q_{t+1} leads to urges for making choices at $t+1$, the agent balances it with the view represented by p_{t+1} as described in Section 2.3, and acts as though she forms the compromise one-step-ahead posterior $m_{t+1} = \alpha_{t+1} p_{t+1} + (1 - \alpha_{t+1}) \, q_{t+1}$. The above noted bias of q_{t+1} extends to this mixture of p_{t+1} and q_{t+1}: substitute for q_{t+1} from (4.6) and deduce that

$$m_{t+1} = (1 - \lambda_{t+1}(1 - \alpha_{t+1})) \, p_{t+1} + \lambda_{t+1}(1 - \alpha_{t+1}) \left[\Sigma_{s'_{t+1}} m_t \left(s'_{t+1} \right) p_{t+1} \left(\cdot \mid s'_{t+1} \right) \right], \tag{4.7}$$

[9] When $\lambda_{t+1} < 0$ in (4.11), q_{t+1} is well-defined as a probability measure only under special conditions.

which admits an interpretation analogous to that described above.[10]

Note that (4.6) defines all q_t's inductively given the p_t's and λ_t's. Thus the corresponding model of utility is completely specified by δ, u, p_0 and the process $(p_t, \alpha_t, \lambda_t)_1^T$.

Further content can be introduced into the model described in (4.6) by imposing structure on the way in which λ_{t+1} depends on the history s_1^{t+1}. For example, it might depend not only on the empirical frequency of observations but also on their order due to sensitivity to streaks or other patterns. While each specialization we have described fixes a sign for λ_{t+1} that is constant across times and histories, one can imagine that an agent might react differently depending on the history. Formulating a theory of the λ_{t+1}'s is a subject for future research.

4.2. Sample-Bias

In the last section, temptation and hence also the updating bias, depended on prior beliefs. Here we describe an alternative specialization of the general model in which temptation and the updating bias depend instead on sample frequencies.

Denote by Ψ_{t+1} the empirical frequency measure on S given the history s_1^{t+1}; that is, $\Psi_{t+1}(s)$ is the relative frequency of s in the sample s_1^{t+1}. Let G lie in $\mathcal{C}_{t+1}$. Then $G(s_{t+2})$ is a subset of $C_{t+2} \times \mathcal{C}_{t+2}$ and so is the mixture $\int G\left(s'_{t+2}\right) d\Psi_{t+1}$. Consider the contingent menu in $\mathcal{C}_{t+1}$, denoted $\int G d\Psi_{t+1}$, that assigns $\int G\left(s'_{t+2}\right) d\Psi_{t+1}$ to every s_{t+2}. Then $\left(c_{t+1}, \int G d\Psi_{t+1}\right)$ denotes the obvious singleton menu.

The axioms to follow parallel the trio of axioms stated in the last section. One difference is that the contingent menus F and G appearing in these axioms are assumed, for reasons given below, to lie in $\mathcal{C}_{t+1}^{c,+1} \subset \mathcal{C}_{t+1}$. Thus F and G provide perfect commitment and are such that all relevant uncertainty is resolved by $t+2$.

Axiom 10 (Sample-Bias). *For all $t < T$, s_{t+1} and c', c, for all F' in $\mathcal{C}_t$, and for all F and G in $\mathcal{C}_{t+1}^{c,+1}$: if*

$$\left(c', \left[F'_{-s_{t+1}},\ \{(c, F)\}\right]\right) \succ_t \left(c', \left[F'_{-s_{t+1}},\ \{(c, G)\}\right]\right) \text{ and}$$

[10] We considered naming the above axioms Underreaction and Overreaction respectively, because attaching too much weight to the prior (as in Positive Prior-Bias) presumably means that in a sense too little weight is attached to data (and similarly for the other axiom). However, the term underreaction suggests low sensitivity of the posterior to the signal s_{t+1}, which need not be the case in (4.7) unless α_{t+1} and λ_{t+1} do not depend on s_{t+1}. See Section 5.1 for more on underreaction and overreaction.

$$(c', \left[F'_{-s_{t+1}},\ \{(c, \int Fd\Psi_{t+1})\}\right]) \sim_t (c', \left[F'_{-s_{t+1}},\ \{(c, \int Gd\Psi_{t+1})\}\right]), \tag{4.8}$$

then

$$(c',\ \left[F'_{-s_{t+1}},\ \{(c,F)\}\right]) \sim_t (c', \left[F'_{-s_{t+1}},\ \{(c,F),(c,G)\}\right]).$$

The next two axioms provide alternative strengthenings of Sample-Bias. Label by **Positive Sample-Bias** the axiom obtained if (4.8) is replaced by

$$(c', \left[F'_{-s_{t+1}},\ \{(c, \int Fd\Psi_{t+1})\}\right]) \succeq_t (c', \left[F'_{-s_{t+1}},\ \{(c, \int Gd\Psi_{t+1})\}\right]). \tag{4.9}$$

Similarly, 'define' **Negative Sample-Bias** by using the hypothesis

$$(c', \left[F'_{-s_{t+1}},\ \{(c, \int Fd\Psi_{t+1})\}\right]) \preceq_t (c', \left[F'_{-s_{t+1}},\ \{(c, \int Gd\Psi_{t+1})\}\right]). \tag{4.10}$$

Interpret Positive Sample-Bias; the other interpretations are similar. First, we interpret (4.9) as saying that *the sample* s_1^{t+1} *makes* F *look more attractive than* G: F delivers $F(s_{t+2})$ in state s_{t+2} and s_{t+2} appears with frequency $\Psi_{t+1}(s_{t+2})$ in the sample. Thus 'on average', F yields $\int Fd\Psi_{t+1}$. But the agent is indifferent between F and its average because she satisfies Independence. Thus (4.9) implies that the average for F is better than that of G. Now the axiom asserts that if commitment to F is preferred (conditionally on s_{t+1}) to commitment to G, *and if* the sample makes F look more attractive than G, then G is not tempting conditionally. The fact that the sample may influence temptation after realization of s_{t+1}, above and beyond its role in the conditional ranking, reveals the excessive influence of the sample at the updating stage. The influence is 'positive' because G can be tempting conditionally only if it was more attractive according to the sample history.

The preceding intuition, specifically the indifference between F and $\int Fd\Psi_{t+1}$, relies on F lying in $\mathcal{C}_{t+1}^{c,+1}$. That is because as s_{t+2} varies, not only does $F(s_{t+2})$ vary but so also does the information upon which the agent bases evaluation of the menu $F(s_{t+2})$. Independence implies indifference to the former variation but not to the latter. For F in $\mathcal{C}_{t+1}^{c,+1}$, however, information is irrelevant because all uncertainty is resolved once s_{t+2} is realized.

Corollary 4.2. *Suppose that* $(\succeq_t)$ *has a representation* $(\delta, u, p_0, (\alpha_t, p_t, q_t)_{1\leq t\leq T})$. *Then* $(\succeq_t)$ *satisfies Sample-Bias if and only if*

$$q_{t+1}(\cdot \mid s_{t+1}) = (1-\lambda_{t+1})\, p_{t+1}(\cdot \mid s_{t+1}) + \lambda_{t+1}\Psi_{t+1}(\cdot), \tag{4.11}$$

for some adapted process (λ_t) *with* $\lambda_{t+1} \leq 1$.[11]

Further, $(\succeq_t)$ *satisfies (i) Positive Sample-Bias or (ii) Negative Sample-Bias if and only if (4.11) is satisfied with respectively (i)* $0 \leq \lambda_{t+1} \leq 1$ *and (ii)* $\lambda_{t+1} \leq 0$.

The implications of the functional form (4.11) are best seen through the implied adjustment rule for one-step-ahead beliefs, which has the form

$$m_{t+1} = (1 - \lambda_{t+1}(1 - \alpha_{t+1}))\, p_{t+1} + \lambda_{t+1}(1 - \alpha_{t+1})\Psi_{t+1}.$$

Under Positive Sample-Bias ($\lambda_{t+1} \geq 0$), the Bayesian update $p_{t+1}(s_{t+2})$ is adjusted in the direction of the sample frequency $\Psi_{t+1}(s_{t+2})$, implying a bias akin to the *hot-hand fallacy* - the tendency to over-predict the continuation of recent observations. For Negative Bias,

$$m_{t+1} = p_{t+1} + (-\lambda_{t+1}(1 - \alpha_{t+1}))\,(p_{t+1} - \Psi_{t+1}),$$

and the adjustment is proportional to $(p_{t+1} - \Psi_{t+1})$, as though expecting the next realization to compensate for the discrepancy between p_{t+1} and the past empirical frequency. This is a form of negative correlation with past realizations as in the *gambler's fallacy.*

Because she uses the empirical frequency measure to summarize past observations, the temptation facing an agent satisfying any of the models in the above corollary depends equally on all past observations, although it might seem more plausible that more recent observations have a greater impact on temptation. This can be accommodated. For example, both the interpretations of the above axioms and the corollary remain intact if Ψ_{t+1} is a weighted empirical frequency measure

$$\Psi_{t+1}(\cdot) = \Sigma_1^{t+1} w_{\tau,t+1} \delta_{s_\tau}(\cdot).$$

Here $\delta_{s_\tau}(\cdot)$ is the Dirac measure on the observation at time τ and $w_{\tau,t+1} \geq 0$ are weights; the special case $w_{\tau,t+1} = \frac{1}{t+1}$ for all τ yields the earlier model. Thus the framework, including axiomatic foundations, permits a large variety of biases due to undue influence of the sample. For example, an agent who is influenced only by the most recent observation is captured by the law of motion

$$m_{t+1} = (1 - \lambda_{t+1}(1 - \alpha_{t+1}))\, p_{t+1} + \lambda_{t+1}(1 - \alpha_{t+1})\, \delta_{s_{t+1}}.$$

[11]When $\lambda_{t+1} < 0$ in (4.11), q_{t+1} is well-defined as a probability measure only under special conditions; for example, it suffices that $\frac{-\lambda_{t+1}}{1-\lambda_{t+1}} \leq \min_{s_{t+2}} p_{t+1}(s_{t+2} \mid s_{t+1})$.

If $\lambda_{t+1} < 0$, the resulting model admits interpretation (in terms of sampling without replacement from changing urns) analogous to that offered by Rabin [17] for his model of the law of small numbers.

5. LEARNING ABOUT PARAMETERS

This section describes an example of our model in which the data generating process is unknown up to a parameter $\theta \in \Theta$. In the benchmark Bayesian model, time t beliefs have the form

$$P_t(\cdot) = \int_\Theta \otimes_{t+1}^T \ell(\cdot \mid \theta)\ d\mu_t, \tag{5.1}$$

where: $\ell(\cdot \mid \theta)$ is a likelihood function (measure on S), μ_0 represents prior beliefs on Θ, and μ_t denotes Bayesian posterior beliefs about the parameter at time t and after observations s_1^t. The de Finetti Theorem shows that beliefs admit such a representation if and only if P_0 is exchangeable. We describe, without axiomatic foundations, a generalization of (5.1) that accommodates non-Bayesian updating.

Our specialization of the model in Section 2.2 to accommodate parameters is defined by a suitable specification for (p_t, q_t), taking (α_t), δ and u as given. We fix (Θ, ℓ, μ_0) and suppose for now that we are also given a process (ν_t), where each ν_t is a probability measure on Θ. (The σ-algebra associated with Θ is suppressed.) The prior μ_0 on Θ induces time 0 beliefs about S_1 given by

$$p_0(\cdot) = m_0(\cdot) = \int_\Theta \ell(\cdot \mid \theta)\, d\mu_0.$$

Proceed by induction: suppose that μ_t has been constructed and define μ_{t+1} by

$$\mu_{t+1} = \alpha_{t+1} BU(\mu_t; s_{t+1}) + (1 - \alpha_{t+1})\nu_{t+1}, \tag{5.2}$$

where $BU(\mu_t; s_{t+1})(\cdot)$ is the Bayesian update of μ_t (see (1.1)). This equation constitutes the *law of motion* for beliefs about parameters. Finally, define (p_{t+1}, q_{t+1}) by

$$p_{t+1}(\cdot) = \int_\Theta \ell(\cdot \mid \theta)\, d\left(BU(\mu_t; s_{t+1})\right) \quad \text{and} \tag{5.3}$$

$$q_{t+1}(\cdot) = \int_\Theta \ell(\cdot \mid \theta)\, d\nu_{t+1}. \tag{5.4}$$

This completes the specification of the model for any given process (ν_t).

Notice that

$$m_{t+1}(\cdot) = \alpha_{t+1}p_{t+1} + (1-\alpha_{t+1})q_{t+1} = \int_{\Theta} \ell(\cdot \mid \theta)\, d\mu_{t+1}. \tag{5.5}$$

In light of the discussion in Section 2.3, preferences at $t+1$ are based on the beliefs about parameters represented by μ_{t+1}. If $\alpha_{t+1} \equiv 0$, then (μ_t) is the process of Bayesian posteriors and the above collapses to the exchangeable model (5.1). More generally, differences from the Bayesian model depend on (ν_t), examples of which are given next.[12]

5.1. Prior-Bias with Parameters

Consider first the case where

$$\nu_{t+1} = (1-\lambda_{t+1})\, BU(\mu_t; s_{t+1}) + \lambda_{t+1}\mu_t, \tag{5.6}$$

where $\lambda_{t+1} \leq 1$. This is readily seen to imply (4.6) and hence Prior-Bias; the bias is positive or negative according to the sign of the λ's. Posterior beliefs about parameters satisfy the law of motion

$$\mu_{t+1} = (1-\lambda_{t+1}(1-\alpha_{t+1}))\;\; BU(\mu_t; s_{t+1}) + \lambda_{t+1}(1-\alpha_{t+1})\;\mu_t. \tag{5.7}$$

The latter equation reveals something of how the inferences of an agent with Prior-Bias differ from those of a Bayesian updater. Compute that (assuming $\alpha_{t+1} \neq 1$)

$$\tfrac{\mu_{t+1}(\theta)}{\mu_{t+1}(\theta')} < \tfrac{\ell(s_{t+1}|\theta)}{\ell(s_{t+1}|\theta')}\tfrac{\mu_t(\theta)}{\mu_t(\theta')} \quad \text{iff} \quad \lambda_{t+1}\ell(s_{t+1} \mid \theta') < \lambda_{t+1}\ell(s_{t+1} \mid \theta). \tag{5.8}$$

For a concrete example, consider coin tossing, with $S = \{H, T\}$, $\Theta \subset (0,1)$ and $\ell(H \mid \theta) = \theta$ and consider beliefs after a string of H's. If there is a Positive Prior-Bias (positive λ's), then repeated application of (5.8) establishes that the agent underinfers in the sense that

$$\tfrac{\mu_{t+1}(\theta)}{\mu_{t+1}(\theta')} < \tfrac{\mu^B_{t+1}(\theta)}{\mu^B_{t+1}(\theta')}, \quad \theta > \theta',$$

where μ^B_{t+1} is the posterior of a Bayesian who has the same prior at time 0. Similarly, Negative Prior-Bias leads to overinference.

[12] One general point is that, in contrast to the exchangeable Bayesian model, μ_{t+1} depends not only on the set of past observations, but also on the order in which they were realized.

Turn to the question of what is learned in the long run. (Here, for choice theoretic foundations we implicitly rely on an infinite-horizon extension of our model.) Learning may either signify learning the true parameter or learning to forecast future outcomes.[13] The latter kind of learning is more relevant to choice behavior and thus is our focus. Suppose that $\theta^* \in \Theta$ is the true parameter and thus that the i.i.d. measure $P^* = \otimes_{t=1}^{\infty} \ell(\cdot \mid \theta^*)$ is the probability law describing the process (s_t). Say that forecasts are *eventually correct on a path* s_1^∞ if, along that path,

$$m_t(\cdot) \longrightarrow \ell(\cdot \mid \theta^*) \quad \text{as } t \longrightarrow \infty.$$

Rewrite the law of motion for posteriors (5.7) in the form

$$\mu_{t+1} = \left(1 - \gamma_{t+1}\right) BU\left(\mu_t; s_{t+1}\right) + \gamma_{t+1}\mu_t, \tag{5.9}$$

where $\gamma_{t+1} = \lambda_{t+1}(1 - \alpha_{t+1}) \leq 1$. In general, γ_{t+1} is $\mathcal{S}_{t+1}$-measurable (γ_{t+1} may depend on the entire history s_1^{t+1}, including s_{t+1}), but we will be interested also in the special case where γ_{t+1} is $\mathcal{S}_t$-measurable. In that case, (5.9) can be interpreted not only in terms of Positive and Negative Prior-Bias as above, but also in terms of underreaction and overreaction to data. For example, let $\gamma_{t+1} \geq 0$ (corresponding to $\lambda_{t+1} \geq 0$). Then μ_{t+1} is a mixture, with weights that are independent of s_{t+1}, of two terms: (i) the Bayesian update $BU\left(\mu_t; s_{t+1}\right)$, which incorporates the 'correct' response to s_{t+1}, and (ii) the prior μ_t, which does not respond to s_{t+1} at all. In a natural sense, therefore, an agent with $\gamma_{t+1} \geq 0$ *underreacts* to data. Similarly, if $\gamma_{t+1} \leq 0$, then $BU\left(\mu_t; s_{t+1}\right)$ is a mixture of μ_{t+1} and μ_t, which suggests that μ_{t+1} reflects *overreaction*. Clearly, if $\gamma_{t+1} = 0$ then the model reduces to the Bayesian updating rule.

Theorem 5.1. *Let Θ be finite and $\mu_0(\theta^*) > 0$.*

(a) Suppose that γ_{t+1} is $\mathcal{S}_t$-measurable and that $\gamma_{t+1} \geq 0$. Then forecasts are eventually correct $P^* - a.s.$

(b) Suppose that γ_{t+1} is $\mathcal{S}_t$-measurable and that $\gamma_{t+1} \leq 1 - \epsilon$ for some $\epsilon > 0$. Then forecasts are eventually correct with P^*-strictly positive probability.

(c) If one drops either of the assumptions in (a), then there exist (S, Θ, ℓ, μ_0) and $\theta \neq \theta^*$ such that

$$m_t(\cdot) \longrightarrow \ell(\cdot \mid \theta) \quad \text{as } t \longrightarrow \infty,$$

with P^*-strictly positive probability.

[13] See [12] for the distinction between these two kinds of learning.

Assume that before any data are observed the prior belief puts positive weight on the true parameter, that is, assume that $\mu_0(\theta^*) > 0$. Then multiple repetition of Bayes' Rule leads to near correct forecasts. This result is central in the Bayesian literature because it shows that the mere repetition of Bayes' Rule eventually transforms the historical record into a near perfect guide for the future. Part (a) of the theorem generalizes the Bayesian result to the case of underreaction. This result shows that, if repeated sufficiently many times, all non-Bayesian updating rules in (5.9) with the additional proviso of a Positive Prior-Bias and the indicated added measurability assumption, eventually produce good forecasting. Hence, in the case of underreaction, agent's forecasts converge to rational expectations although the available information is not processed according to Bayesian laws of probability.

Part (b) shows that, with positive probability, forecasts are eventually correct provided that the Bayesian term on the right side of (5.9) receives weight that is bounded away from zero. This applies in the case of Negative Prior-Bias, corresponding to overreaction. In fact, the results holds even if the forecaster sometimes overreacts and sometimes underreacts to new information. However, part (c) shows that convergence to wrong forecasts may occur in the absence of either of the assumptions in (a). This is demonstrated by two examples. In the first example the weight γ_{t+1} is constant, but sufficiently negative, corresponding to a forecaster that sufficiently overreacts to new information. In the second example, the weight γ_{t+1} is positive corresponding to underreaction, but γ_{t+1} depends on the current signal and, therefore, γ_{t+1} is only $\mathcal{S}_{t+1}$-measurable. In both examples, forecasts may eventually converge to an incorrect limit. Moreover, wrong forecasts in the limit are at least as likely to occur as are correct forecasts.

The proof of Theorem 5.1 builds on classic arguments of the Bayesian literature. Consider the probability measure μ_t on the parameter space and let the random variable μ_t^* be the probability that μ_t assigns to the true parameter. It follows that the expected value (according to the true data generating process) of the Bayesian update of μ_t^* (given new information) is greater than μ_t^* itself. Hence, in the Bayesian case, the weight given to the true parameter tends to grow as new information is observed. This submartingale property ensures that Bayesian forecasts must converge to some value and cannot remain in endless random fluctuations. The submartingale property follows because under the Bayesian paradigm future changes in beliefs that can be predicted are incorporated in current beliefs. It is immediate from the linear structure in (5.9) that this basic submartingale property still holds in our model as long as the weight γ_{t+1} depends upon the history only up to period t. Hence, with this measurability assumption, forecasts in our model must also converge and, as in the Bayesian case, cannot remain in endless random fluctuations.[14] In addition, convergence to the truth holds in both the Bayesian paradigm and in the case of underreaction. However, given sufficiently strong overreaction, it is possible that forecasts will settle on an incorrect limit. This follows because the positive drift of the above mentioned submartingale property on μ_t^* may be compensated by sufficiently strong volatility which permits that, with positive probability, μ_t^* converges to zero.

5.2. Sample-Bias with Parameters

Sample-Bias can also be modeled when learning about parameters is taking place. Take as primitive a process (ψ_{t+1}) of probability measures on Θ that provides a representation for empirical frequency measures Ψ_{t+1} of the form

$$\Psi_{t+1} = \int \ell\left(\cdot \mid \theta\right) d\psi_{t+1}\left(\theta\right). \tag{5.10}$$

Let μ_0 be given and define μ_{t+1} and ν_{t+1} inductively for $t \geq 0$ by (5.2) and

$$\nu_{t+1} = \left(1 - \lambda_{t+1}\right)\ BU(\mu_t, s_{t+1}) + \lambda_{t+1}\psi_{t+1}, \tag{5.11}$$

for $\lambda_{t+1} \leq 1$. Then one obtains a special case of the Sample-Bias model of Corollary 4.2; the bias is positive or negative according to the sign of the λ's. The implied law of motion for posteriors is

$$\mu_{t+1} = \left(1 - \lambda_{t+1}(1 - \alpha_{t+1})\right) BU\left(\mu_t; s_{t+1}\right) + \lambda_{t+1}(1 - \alpha_{t+1})\ \psi_{t+1}. \tag{5.12}$$

To illustrate, suppose that $S = \{s^1, ..., s^K\}$ and that $\ell\left(s^k \mid \theta\right) = \theta_k$ for each $\theta = (\theta_1, ..., \theta_K)$ in Θ, the interior of the K-simplex. Then one can ensure (5.10) by taking ψ_0 to be a suitable noninformative prior; subsequently, Bayesian updating leads to the desired process (ψ_{t+1}). For example, the improper Dirichlet prior density

$$\frac{d\psi_0\left(\theta\right)}{\Pi_{k=1}^K d\theta_k} \propto \Pi_{k=1}^K \theta_k^{-1}$$

[14] We conjecture that beliefs μ_t may not converge in some examples when the weight γ_{t+1} is $\mathcal{S}_{t+1}$-measurable. In our example, it does converge, but to an incorrect limit.

yields the Dirichlet posterior with parameter vector $\left(n_t\left(s^1\right),...,n_t\left(s^K\right)\right)$, where $n_t\left(s^k\right)$ equals the number of realizations of s^k in the first t periods; that is,

$$\frac{d\psi_t(\theta)}{\Pi_{k=1}^K d\theta_k} \propto \Pi_{k=1}^K \theta_k^{n_t\left(s^k\right)-1}. \tag{5.13}$$

By the property of the Dirichlet distribution,

$$\int \ell\left(s^k \mid \theta\right) d\psi_t(\theta) = \int \theta_k \, d\psi_t(\theta) = \tfrac{n_k(t)}{t},$$

the empirical frequency of s^k, as required by (5.10).

Finally, compute from (5.12) and (5.13) that (assuming $\alpha_{t+1} \neq 0$)

$$\tfrac{\mu_{t+1}(\theta)}{\mu_{t+1}(\theta')} > \tfrac{\ell(s_{t+1}|\theta)}{\ell(s_{t+1}|\theta')}\tfrac{\mu_t(\theta)}{\mu_t(\theta')} \quad \text{iff} \quad \lambda_{t+1}\tfrac{\psi_t(\theta)}{\psi_t(\theta')} > \lambda_{t+1}\tfrac{\mu_t(\theta)}{\mu_t(\theta')}. \tag{5.14}$$

Suppose that all λ_{t+1}'s are negative (Negative Sample-Bias) and consider the coin-tossing example. As above, we denote by $\left(\mu_t^B\right)$ the Bayesian process of posteriors with initial prior $\mu_0^B = \mu_0$. Then it follows from repeated application of (5.13) and (5.14) that

$$\tfrac{\mu_{t+1}(\theta)}{\mu_{t+1}(\theta')} > \tfrac{\mu_{t+1}^B(\theta)}{\mu_{t+1}^B(\theta')},$$

if $s_1^{t+1} = (H,...,H)$, $\mid \theta - \frac{1}{2} \mid > \mid \theta' - \frac{1}{2} \mid$ and if the common initial prior μ_0 is uniform.[15] After seeing a string of H's the agent described herein exaggerates (relative to a Bayesian) the relative likelihoods of extremely biased coins. If instead we consider a point at which the history s_1^{t+1} has an equal number of realizations of T and H, then

$$\tfrac{\mu_{t+1}(\theta)}{\mu_{t+1}(1-\theta)} > \tfrac{\theta}{1-\theta}\tfrac{\mu_t(\theta)}{\mu_t(1-\theta)} = \tfrac{BU(\mu_t,H)(\theta)}{BU(\mu_t,H)(1-\theta)},$$

for any θ such that $\mu_t(\theta) > \mu_t(1-\theta)$. If there have been more realizations of H, then the preceding displayed inequality holds if

$$\left(\tfrac{\theta}{1-\theta}\right)^{n_{t+1}(H)-n_{t+1}(T)} < \tfrac{\mu_t(\theta)}{\mu_t(1-\theta)},$$

for example, if $\theta < \frac{1}{2}$ and $\mu_t(\theta) \geq \mu_t(1-\theta)$. Note that the bias in this case is towards coins that are less biased ($\theta < \frac{1}{2}$). The opposite biases occur in the case of Positive Sample-Bias.

[15] More generally, the latter two conditions can be replaced by $\frac{\theta'(1-\theta')}{\theta(1-\theta)} > \frac{\mu_0(\theta)}{\mu_0(\theta')}$.

We conclude with a result regarding learning in the long run. In order to avoid technical issues arising from Θ being a continuum as in the Dirichlet-based model, we consider the following variation: as before $S = \{s^1, ..., s^K\}$ and $\ell\left(s^k \mid \theta\right) = \theta_k$ for each k and θ. But now take Θ to be the set of points $\theta = (\theta_1, ..., \theta_K)$ in the interior of the K-simplex having rational co-ordinates. Define

$$\psi_{t+1}(\theta) = \begin{cases} 1 & \text{if the empirical frequency of } s^k \text{ is } \theta_k,\ 1 \leq k \leq K, \\ 0 & \text{otherwise.} \end{cases}$$

Then (5.10) is evident.[16] The law of motion can be written in the form

$$\mu_{t+1} = \left(1 - \gamma_{t+1}\right) BU\left(\mu_t; s_{t+1}\right) + \gamma_{t+1}\psi_{t+1}, \tag{5.15}$$

where $\gamma_{t+1} = \lambda_{t+1}(1 - \alpha_{t+1}) \leq 1$.

We have the following partial counterpart of Theorem 5.1.

Theorem 5.2. *Let S, (Θ, ℓ) and (ψ_t) be as just defined and suppose that posteriors (μ_t) evolve according to (5.15), where γ_{t+1} is $\mathcal{S}_t$-measurable and $0 < \underline{\gamma} \leq \gamma_{t+1} \leq 1$. Then forecasts are eventually correct $P^* - a.s.$*

The positive lower bound $\underline{\gamma}$ excludes the Bayesian case. The result does hold in the Bayesian case $\gamma_{t+1} = 0$. However, unlike the proof of Theorem 5.1, the proof of Theorem 5.2 is in some ways significantly different from the proof in the Bayesian case. We suspect that the differences in the approach make the lower bound assumption technically convenient but ultimately disposable. We also conjecture (but cannot yet prove) that just as in part (c) of Theorem 5.1, convergence to the truth fails in general if γ_{t+1} is only $\mathcal{S}_{t+1}$-measurable. The other case treated in the earlier theorem - γ_{t+1} is $\mathcal{S}_t$-measurable but possibly negative - (which in the context of that model corresponded to overreaction) is not relevant here because these conditions violate the requirement that each ν_{t+1} in (5.11) be a probability measure and hence non-negatively valued.

6. CONCLUDING REMARKS

Our main contribution is to provide a choice-theoretic model of updating. An important feature of the model is its richness - it can accommodate a range of

[16]If Θ were taken to be finite, then one could not assure (5.10) without admitting signed measures for ψ_{t+1} and hence also for μ_{t+1}. Bayesian updating is not well-defined for signed measures and even if that problem were overcome, the interpretation of such a model is not clear.

updating biases. We have illustrated this to a degree via the (axiomatic) specializations called Prior-Bias and Sample-Bias. However, much more might be done in this vein. For example, we characterized two alternative specializations of Sample-Bias that correspond roughly to the hot-hand fallacy (Positive Sample-Bias) and the gambler's fallacy (Negative Sample-Bias) respectively. However, while in each case the agent is assumed to suffer from the indicated fallacy at all times and histories, it is intuitive that she may move from one fallacy to another depending on the sample history. Thus one would like a theory that explains which fallacy applies at each history. Our framework gives this task a concrete form: in light of Corollary 4.2, one must 'only' explain how the weights λ_{t+1} vary with history. Similarly with regard to further specializations of Prior-Bias.

References

[1] C.D. Aliprantis and K. Border, *Infinite Dimensional Analysis*, Springer Verlag, 1994.

[2] C. Camerer, Individual decision-making, in *Handbook of Experimental Economics*, J. Kagel and A. Roth eds., Princeton U. Press, 1995.

[3] E. Dekel, B. Lipman and A. Rustichini, A unique subjective state space for unforeseen contingencies, *Econometrica* 69 (2001), 891-934.

[4] L.G. Epstein, Stationary cardinal utility and optimal growth under uncertainty, *J. Econ. Theory* 31 (1983), 133-152.

[5] L.G. Epstein, An axiomatic model of non-Bayesian updating, revised January 2005.

[6] D. Freedman, "On tail probabilities for martingales," *Annals of Probability* 3 (1975), 100-118.

[7] F. Gul and W. Pesendorfer, Temptation and self-control, *Econometrica* 69 (2001), 1403-35.

[8] F. Gul and W. Pesendorfer, Self-control and the theory of consumption, *Econometrica* 72 (2004), 119-158.

[9] I. Kopylov, Temptations in general settings, 2005.

[10] D. Kreps, *Notes on the Theory of Choice*, Westview, 1988.

[11] D. Kreps, Static choice in the presence of unforeseen contingencies, in *Essays in Honour of F. Hahn*, P. Dasgupta et al eds., MIT Press, 1992.

[12] E. Lehrer and R. Smorodinsky, Merging and learning, *Statistics, Probability and Game Theory, IMS Lecture Notes* 30 (1996), 147-168.

[13] O.L. Mangasarian, *Nonlinear Programming*, McGraw-Hill, New York: 1969.

[14] S. Mullainathan, Thinking through categories, MIT, 2000.

[15] J. Noor, Temptation, welfare and revealed preference, 2004.

[16] M. Rabin, Psychology and economics, *J. Ec. Lit.* 36 (1998), 11-46.

[17] M. Rabin, Inference by believers in the law of small numbers, *Quart. J. Econ.* 117 (2002), 775-816.

[18] R.H. Strotz, Myopia and inconsistency in dynamic utility maximization, *Rev. Ec. Studies* 23 (1955), 165-180.

[19] A. Tversky and D. Kahneman, Belief in the law of small numbers, *Psychology Bulletin* 76 (1971), 105-110.

[20] A. Tversky and D. Kahneman, Judgement under uncertainty: heuristics and biases, *Science* 185 (1974), 1124-31.

A. APPENDIX: Proof of Main Representation Result

A.1. Preliminaries

For any compact metric space D endowed with a continuous mixture operation, say that a preference $\succeq$ over $\mathcal{M}(D)$ has a (U,V) representation if the functions $U,V : D \to \mathbb{R}$ are continuous and linear, and if $\succeq$ is represented by $W_{U,V} : \mathcal{M}(D) \to \mathbb{R}$, where

$$W_{U,V}(M) = \max_{c \in M}\{U+V\} - \max_{c' \in M} V, \quad M \in \mathcal{M}(D).$$

Say that $\succeq$ is *strategically rational* if for all $M, M' \in \mathcal{M}(D)$,

$$M \succeq M' \Longrightarrow M \sim M \cup M'.$$

Lemma A.1. *If $\succeq$ has a (U,V) representation with U nonconstant, then:*

(a) $\succeq$ is strategically rational iff $V = aU+b$ for some $a \geq 0$. In particular, if V is nonconstant then $\succeq$ is strategically rational iff $U+V = aV+b$ for some $a > 1$.

(b) $\succeq$ is strategically rational iff for all $\bar{c}, \underline{c} \in D$,

$$\{\bar{c}\} \succeq \{\underline{c}\} \Longrightarrow \{\bar{c}\} \sim \{\bar{c}, \underline{c}\}. \tag{A.1}$$

Proof. (a) The argument is similar to [7, p. 1414].

(b) Sufficiency is clear. For necessity, suppose that $\succeq$ is not strategically rational so that, as in [7, p. 1414], U and V are nonconstant and U is not a positive affine transformation of V. Consequently, there exist $\bar{c}, \underline{c} \in D$ such that either [$U(c) > U(c')$ and $V(c) \leq V(c')$], or [$U(c) \geq U(c')$ and $V(c) < V(c')$]. Linearity and nonconstancy of U and V imply the existence of $\bar{c}$ and $\underline{c}$ close to c and c', respectively, such that all inequalities are strict. Then

$$\{\bar{c}\} \succ \{\underline{c}\} \text{ and } \{\bar{c}\} \succ \{\bar{c}, \underline{c}\},$$

which violates (A.1) and yields the result. ■

Lemma A.2. *Suppose that $\succeq$ has a (U,V) representation and that there exists $\bar{c}, \underline{c}$ such that $\{\bar{c}, \underline{c}\} \succ \{\underline{c}\}$. Then a preference $\succeq^*$ over D is represented by $U+V$ if and only if it satisfies the vNM axioms and the following restriction:*

$$\textit{if } \{c\} \succ \{d\}, \textit{ then } \{c,d\} \succ \{d\} \Longleftrightarrow c \succ^* d. \tag{A.2}$$

Proof. That the conditions on $\succeq^*$ are necessary for it to be represented by $U+V$ follows immediately from the latter's linearity and continuity, and by Step 1 below. Establish the converse.

Step 1: Show that if $\{c\} \succ \{d\}$, then

$$\{c,d\} \succ \{d\} \Longleftrightarrow U(c)+V(c) > U(d)+V(d).$$

$\Longrightarrow$: Suppose $\{c,d\} \succ \{d\}$, and that, by way of contradiction, $U(c)+V(c) \not> U(d)+V(d)$. Then $\{c,d\} \succ \{d\} \Longrightarrow$

$\max_{\{c,d\}} \{U+V\} - \max_{\{c,d\}} V > U(d) \Longrightarrow$

$U(d)+V(d) - \max_{\{c,d\}} V > U(d) \Longrightarrow$

$V(d) > \max_{\{c,d\}} V$, which contradicts $\{c,d\} \succ \{d\}$.

$\Longleftarrow$: Suppose $U(c)+V(c) > U(d)+V(d)$ and by way of contradiction that $\{c,d\} \not\succ \{d\}$. Then $\{c,d\} \not\succ \{d\} \Longrightarrow$

$\max_{\{c,d\}} \{U+V\} - \max_{\{c,d\}} V \leq U(d) \Longrightarrow$

$U(c)+V(c) - \max_{\{c,d\}} V \leq U(d) \Longrightarrow$

$U(c)+V(c) \leq U(d) + \max_{\{c,d\}} V$.

If $\max_{\{c,d\}} V = V(c)$, then $U(c) \leq U(d)$, contradicting the hypothesis that $\{c\} \succ \{d\}$. If $\max_{\{c,d\}} V = V(d)$, then $U(c)+V(c) \leq U(d)+V(d)$, contradicting the hypothesis that $U(c)+V(c) > U(d)+V(d)$.

Step 2: The result.

By hypothesis, $\{\bar{c}, \underline{c}\} \succ \{\underline{c}\}$, and by Set Betweenness, $\{\bar{c}\} \succ \{\underline{c}\}$. Thus Step 1 and (A.2) imply that $U(\bar{c})+V(\bar{c}) > U(\underline{c})+V(\underline{c})$ and $\bar{c} \succ^* \underline{c}$. These observations will be used to prove that $U+V$ represents $\succeq^*$.

Take any $c, d \in D$ such that $c \succ^* d$ and suppose by way of contradiction that $U(d)+V(d) \geq U(c)+V(c)$. Step 1 and (A.2) rule out $\{c\} \succ \{d\}$. If, on the other hand, $\{d\} \succeq \{c\}$, then by Independence and linearity of $U+V$,

$$\{d\alpha\bar{c}\} \succ \{c\alpha\underline{c}\} \text{ and } U(d\alpha\bar{c}) + V(d\alpha\bar{c}) > U(c\alpha\underline{c}) + V(c\alpha\underline{c}),$$

for all $\alpha \in (0,1)$, where $d\alpha\bar{c}$ is short-hand for the mixture $\alpha d + (1-\alpha)\,\bar{c}$ and so on. By Step 1, $\{d\alpha\bar{c}, c\alpha\underline{c}\} \succ \{c\alpha\underline{c}\}$ for all $\alpha \in (0,1)$, and by (A.2), $d\alpha\bar{c} \succ^* c\alpha\underline{c}$ for all $\alpha \in (0,1)$. Continuity of $\succeq^*$implies $d \succeq^* c$, a contradiction.

Next suppose $c \sim^* d$ and wlog $U(c)+V(c) > U(d)+V(d)$. If $\{c\} \succ \{d\}$ then by Step 1, $\{c,d\} \succ \{d\}$, and by (A.2), $c \not\sim^* d$, a contradiction. If, on the other hand, $\{d\} \succeq \{c\}$, then since $\succeq$ and $\succeq^*$ satisfy the Independence axiom, it follows that

$$\{d\alpha\bar{c}\} \succ \{c\alpha\underline{c}\} \text{ and } d\alpha\bar{c} \succ^* c\alpha\underline{c},$$

for all $\alpha \in (0,1)$. By (A.2), $\{d\alpha\bar{c}, c\alpha\underline{c}\} \succ \{c\alpha\underline{c}\}$ for all $\alpha \in (0,1)$, and thus by Step 1, $U(d\alpha\bar{c}) + V(d\alpha\bar{c}) > U(c\alpha\underline{c})+V(c\alpha\underline{c})$ for all $\alpha \in (0,1)$. By continuity of $U+V$, $U(d)+V(d) \geq U(c)+V(c)$, a contradiction. ■

For any state s_{t+2}, $G \in \mathcal{C}_{t+1}$ and $L \subset \mathcal{M}(C_{t+2} \times \mathcal{C}_{t+2})$, define the set $Ls_{t+2}G$ of contingent menus by

$$Ls_{t+2}G = \left\{[G_{-s_{t+2}}, M] : M \in L\right\} \subset \mathcal{C}_{t+1}.$$

Define $\succeq_{t\,|s_{t+1},s_{t+2}}$ on closed subsets of $\mathcal{M}(C_{t+2} \times \mathcal{C}_{t+2})$ by: $L' \succeq_{t\,|s_{t+1},s_{t+2}} L$ iff

$$\left(c', \left[F_{-s_{t+1}},\ (c, L's_{t+2}G)\right]\right) \succeq_t \left(c', \left[F_{-s_{t+1}},\ (c, Ls_{t+2}G)\right]\right),$$

for some $c, c' \in C$, F in $\mathcal{C}_t$, and G in $\mathcal{C}_{t+1}$.

Lemma A.3. *Suppose that $(\succeq_t)$ satisfies Axioms 1-8 and that $\succeq_{t|s_{t+1},s_{t+2}}$ has a (U,V) representation with nonconstant U. Then $\succeq_{t|s_{t+1},s_{t+2}}$ is strategically rational.*

Proof. By Lemma A.1(b), we need only establish that for any $M, M' \in \mathcal{M}(C_{t+2} \times \mathcal{C}_{t+2})$,

$$\{M\} \succeq_{t|s_{t+1,t+2}} \{M'\} \Longrightarrow \{M\} \sim_{t|s_{t+1,t+2}} \{M, M'\}.$$

Observe that $\{M\} \succeq_{t|s_{t+1,t+2}} \{M'\} \Longleftrightarrow$
$\left(c',\ \left[F_{-s_{t+1}},\ \{(c, [G_{-s_{t+2}}, M])\}\right]\right) \succeq_t \left(c',\ \left[F_{-s_{t+1}}, \{\left(c, [G_{-s_{t+2}}, M']\right)\}\right]\right) \Longrightarrow^*$
$\left(c',\ \left[F_{-s_{t+1}}, \{\ c, [G_{-s_{t+2}}, M])\}\right]\right) \sim_t \left(c',\ \left[F_{-s_{t+1}}, \{(c, [G_{-s_{t+2}}, M]), (c, [G_{-s_{t+2}}, M'])\}\right]\right)$
$\Longrightarrow \{M\} \sim_{t|s_{t+1,t+2}} \{M, M'\}$, where the implication $\Longrightarrow^*$ is by RSR. ■

In the next Lemma, $\succeq_t$ and $\succeq_{t+1}$ are the preferences corresponding to histories $(s_1, ..., s_t)$ and $(s_1, ..., s_t, s_{t+1})$ respectively.

Lemma A.4. *Suppose that $(\succeq_t)$ satisfies Axioms 1-8. If $H, H' \in \mathcal{C}_{t+1}$ are such that $H(s'_{t+2}) = H'(s'_{t+2})$ for all $s'_{t+2} \neq s_{t+2}$, then for any s_{t+1}, c, c' and F,*

$$(c, H) \succeq_{t+1} (c, H') \Longleftrightarrow$$

$$(c', [F_{-s_{t+1}},\ \{(c, H)\}]) \succeq_t (c', [F_{-s_{t+1}},\ \{(c, H')\}]).$$

Proof. $\Longleftarrow$ follows from RSR. Conversely, suppose that $(c, H) \succeq_{t+1} (c, H')$ and $\left(c', [F_{-s_{t+1}},\ \{(c, H')\}]\right) \succ_t \left(c', [F_{-s_{t+1}},\ \{(c, H)\}]\right)$. Sophistication implies

$$(c', [F_{-s_{t+1}},\ \{(c, H'), (c, H)\}]) \preceq_t (c', [F_{-s_{t+1}},\ \{(c, H)\}]);$$

by Set-Betweenness, this weak preference is in fact indifference. Therefore,

$$(c', [F_{-s_{t+1}},\ \{(c, H')\}]) \succ_t (c',\ [F_{-s_{t+1}},\ \{(c, H'), (c, H)\}]),$$

which contradicts RSR. ■

A.2. Proof of Theorem 3.1

Necessity: We provide details only for some axioms. To verify Sophistication, suppose that $t < T$ and $(c, [G_{-s_{t+1}},\ \{f\}]) \succ_t (c, [G_{-s_{t+1}},\ \{g\}])$. Then by the representation and Step 1 of the proof of Lemma A.2,

$$(c, [G_{-s_{t+1}},\ \{f, g\}]) \succ_t (c, [G_{-s_{t+1}},\ \{g\}]) \Longleftrightarrow \mathcal{U}_{t+1}(f) > \mathcal{U}_{t+1}(g) \Longleftrightarrow f \succ_{t+1} g.$$

Risk Preference and State Independence follow from p_{t+1} having full support and the fact that for any $\ell \in \mathcal{L}_{t+1}$, $U_{t+1}(\ell, s_{t+1}) = u(\ell_{t+1}) + \delta\Sigma_{t+2}^{T+1}\ \delta^{\tau-(t+2)} u(\ell_\tau)$. To establish RSR, take any $H, H' \in \mathcal{C}_{t+1}$ such that $H(s'_{t+2}) = H'(s'_{t+2})$ for all $s'_{t+2} \neq s_{t+2}$. The hypothesis (3.6) implies $U_{t+2}(H'(s_{t+2}), s_{t+2}) \geq U_{t+2}(H(s_{t+2}), s_{t+2})$. Thus $U_{t+1}(\{(c, H'), (c, H)\}, s_{t+1}) =$

$u(c) + \max_{F \in \{H', H\}} \delta \int_{S_{t+2}} U_{t+2}(F(s_{t+2}), s_{t+2})\ d\left(p_{t+1} + \frac{1-\alpha_{t+1}}{\alpha_{t+1}} q_{t+1}\right)$

$- \max_{F' \in \{H', H\}} \frac{1-\alpha_{t+1}}{\alpha_{t+1}} \delta \int_{S_{t+2}} U_{t+2}(F'(s_{t+2}), s_{t+2})\ dq_{t+1}(s_{t+2})$

$= u(c) + \max_{F \in \{H', H\}} \delta \int_{S_{t+2}} U_{t+2}(F(s_{t+2}), s_{t+2})\ dp_{t+1}(s_{t+2})$

$= u(c) + \delta \int_{S_{t+2}} U_{t+2}(H'(s_{t+2}), s_{t+2})\ dp_{t+1}(s_{t+2})$

$= U_{t+1}(\{(c, H')\}, s_{t+1})$, which implies (3.7).

To see (3.8) note that since p_{t+1} has full support,

$U_{t+2}(H'(s_{t+2}), s_{t+2}) \geq U_{t+2}(H(s_{t+2}), s_{t+2}) \Longrightarrow$

$\int_{S_{t+2}} U_{t+2}(H'(s_{t+2}), s_{t+2})\ d\left(p_{t+1} + \frac{1-\alpha_{t+1}}{\alpha_{t+1}} q_{t+1}\right)$

$\geq \int_{S_{t+2}} U_{t+2}(H(s_{t+2}), s_{t+2})\ d\left(p_{t+1} + \frac{1-\alpha_{t+1}}{\alpha_{t+1}} q_{t+1}\right)$

$\Longrightarrow \mathcal{U}_{t+1}(c, H') > \mathcal{U}_{t+1}(c, H)$, as desired.

Sufficiency: The proof of sufficiency is by backward induction on t. Begin by showing that $\succeq_T$ is represented by the function

$$\mathcal{W}_T(c_T, F_T) = u(c_T) + \delta \int_{S_{T+1}} u(F_T(s_{T+1}))\ dm_T, \quad (c_T,\ F_T) \in C_T \times \mathcal{C}_T, \tag{A.3}$$

where $m_T \in \Delta(S_{T+1})$ and m_T has full support.

We claim that $\succeq_T$ may be represented by

$$\mathcal{W}_T(c, F_T) = u_1(c) + u_2(F_T), \tag{A.4}$$

where $u_1(\cdot)$ and $u_2(\cdot)$ are continuous and linear. Argue as follows: Since $C_T \times \mathcal{C}_T$ is a mixture space and $\succeq_T$ satisfies Order, Continuity and Independence, there exists a continuous linear representation $\mathcal{W}_T(\cdot)$ of $\succeq_T$. By definition of the mixture operation, for any $c, c' \in C$ and $F, F' \in \mathcal{C}_T$,

$$\frac{1}{2}(c, F) + \frac{1}{2}(c', F') = \frac{1}{2}(c', F) + \frac{1}{2}(c, F').$$

Thus

$$\mathcal{W}_T\left(\tfrac{1}{2}(c,F) + \tfrac{1}{2}(c',F')\right) = \mathcal{W}_T\left(\tfrac{1}{2}(c',F) + \tfrac{1}{2}(c,F')\right) \Longrightarrow$$
$$\tfrac{1}{2}\mathcal{W}_T(c,F) + \tfrac{1}{2}\mathcal{W}_T(c',F') = \tfrac{1}{2}\mathcal{W}_T(c',F) + \tfrac{1}{2}\mathcal{W}_T(c,F') \Longrightarrow$$
$$\mathcal{W}_T(c,F) = \mathcal{W}_T(c,F') + \mathcal{W}_T(c',F) - \mathcal{W}_T(c',F') \equiv u_1(c) + u_2(F).$$

Linearity and continuity of u_1 and u_2 are evident.

Return to the proof of (A.3). Take any c and define the order $\succeq$ on $\mathcal{C}_T$ by

$$F \succeq G \Longleftrightarrow (c, F) \succeq_T (c, G). \tag{A.5}$$

Verify that $\succeq$ satisfies the Anscombe-Aumann axioms: Order, Continuity and Independence are immediate. By Risk Preference and nonconstancy of $u(\cdot)$, there exists $c', c'' \in C_{T+1}$ such that $c' \not\sim c''$, and thus $\succeq$ satisfies the Anscombe-Aumann nondegeneracy condition. State Independence applied twice yields $(F_{-s_{T+1}}, c') \succeq (F_{-s_{T+1}}, c'') \Longrightarrow (F_{-s'_{T+1}}, c') \succeq (F_{-s'_{T+1}}, c'')$ for all $c', c'' \in C_{T+1}$ and $s_{T+1}, s'_{T+1} \in S_{T+1}$. Thus there exists $m_T \in \Delta(S_{T+1})$ and $u : C_{T+1} \longrightarrow \mathbb{R}$, nonconstant, continuous and linear, such that $\succeq$ is represented by $w(\cdot)$,

$$w(F_T) = \int_{S_{T+1}} u(F_T(s_{T+1}))\, dm_T, \qquad F_T \in \mathcal{C}_T.$$

Since $u_2(\cdot)$ is continuous, linear and (by (A.5)) ordinally equivalent to $w(\cdot)$, it follows that $u_2(\cdot) = aw(\cdot) + b$ for some $a > 0$. By Risk Preference, it is wlog to set $u(\cdot) = u_1(\cdot)$, $a = \delta$ and $b = 0$. State Independence, Risk Preference and the nonconstancy of $u(\cdot)$ imply that m_T has full support.

As the induction hypothesis, suppose that for some $t < T$ and every τ satisfying $t \leq \tau < T$, $\succeq_{\tau+1}$ is represented by

$$\mathcal{W}_{\tau+1}(c, F_{\tau+1}) = u(c) + \delta \int_{S_{\tau+2}} U_{\tau+2}(F_{\tau+1}(s_{\tau+2}), s_{\tau+2})\, dm_{\tau+1}, \qquad (c, F_{\tau+1}) \in C_{\tau+1} \times \mathcal{C}_{\tau+1},$$

where $m_{\tau+1}$ has full support, $U_{\tau+2}(\cdot, s_{\tau+2}) : \mathcal{M}(C_{\tau+2} \times \mathcal{C}_{\tau+2}) \longrightarrow \mathbb{R}^1$ is nonconstant, continuous, linear and is defined recursively via

$$U_{\tau+2}(M_{\tau+2}, s_{\tau+2}) =$$

$$\max_{(c,F_{\tau+2})\in M_{\tau+2}} \left\{ \begin{array}{c} u(c) + \delta \int_{S_{\tau+3}} U_{\tau+3}(F_{\tau+2}(s_{\tau+3}), s_{\tau+3})\, dp_{\tau+2} \\ + \frac{(1-\alpha_{\tau+2})}{\alpha_{\tau+2}} \left(u(c) + \delta \int_{S_{\tau+3}} U_{\tau+3}(F_{\tau+2}(s_{\tau+3}), s_{\tau+3})\, dq_{\tau+2} \right) \end{array} \right\}$$

$$- \max_{(c',F'_{\tau+2})\in M_{\tau+2}} \frac{(1-\alpha_{\tau+2})}{\alpha_{\tau+2}} \left\{ u(c') + \delta \int_{S_{\tau+3}} U_{\tau+3}\left(F'_{\tau+2}(s_{\tau+3}), s_{\tau+3}\right) dq_{\tau+2} \right\},$$

and the boundary condition

$$U_{T+1}(c_{T+1}, s_{T+1}) = u(c_{T+1}).$$

Above

$$\alpha_{\tau+2} \in (0,1],\ p_{\tau+2},\, q_{\tau+2} \in \Delta(S_{\tau+2}), \text{ each } p_{\tau+2} \text{ has full support,}$$
$$\text{and } m_{\tau+2} = \alpha_{\tau+2} p_{\tau+2} + (1-\alpha_{\tau+2})\, q_{\tau+2}.$$

We construct $\mathcal{W}_t$ having the appropriate form and representing $\succeq_t$.[17] The argument is divided into a series of steps.

Step 1: We define the "convex hull" of contingent menus.

For any mixture space, we have the usual notion of convex hull of a set M - the smallest set convex (mixture-closed) containing M. However, a mixture space framework is not adequate because, for example, $\mathcal{M}(C_T \times \mathcal{C}_T)$ is not a mixture space - $\lambda\left[\lambda' M + (1-\lambda')M'\right] + (1-\lambda)M' \neq \lambda\lambda' M + (1-\lambda\lambda')M'$ if M and M' are not convex . More generally, because $\alpha M + (1-\alpha)M \neq M$ in general, the "convex hull" of any M need not contain M. In fact, we are interested in the convex hull of contingent menus. Thus we define $co(F_t)$ for any F_t in $\mathcal{C}_t$ and we do so by backward induction.

For $t = T-1$, $\mathcal{C}_T = (C_{T+1})^{S_{T+1}}$, the set of (Anscombe-Aumann) acts over S_{T+1}, and both $\mathcal{C}_T$ and $C_T \times \mathcal{C}_T$ are mixture spaces. Thus so is $C_T \times \mathcal{C}_T$, and "convex hull of M_{T-1}" has the usual meaning - the smallest convex set containing M_{T-1}. For any contingent menu F_{T-1} in $\mathcal{C}_{T-1}$, define its convex hull, $co(F_{T-1})$, as the contingent menu that maps each s_T into $co(F_{T-1}(s_T))$. Let

$$\mathcal{D}_{T-1} = \{co\left(F'_{T-1}\right) : F'_{T-1} \in \mathcal{C}_{T-1}\} \subset \mathcal{C}_{T-1}.$$

Then $\mathcal{D}_{T-1}$ is a mixture space.

For the inductive step, supposing that $co(\cdot)$ has been defined on $\mathcal{C}_{t+1}$, and that

$$\mathcal{D}_{t+1} = \{co\left(F'_{t+1}\right) : F'_{t+1} \in \mathcal{C}_{t+1}\} \subset \mathcal{C}_{t+1}$$

is a mixture space. Let $F_t \in \mathcal{C}_t$, $s_{t+1} \in S_{t+1}$, and

$$N = \{(c_{t+1}, co(F_{t+1})) : (c_{t+1}, F_{t+1}) \in F_t(s_{t+1})\}.$$

Since $C_{t+1} \times \mathcal{D}_{t+1}$ is a mixture space, the smallest convex subset of $C_{t+1} \times \mathcal{D}_{t+1}$ containing N is well-defined. We define $co(F_t)(s_{t+1})$ to be that set. This defines $co(F_t)$. Note that it lies in $\mathcal{D}_t = \{co(F'_t) : F'_{t+1} \in \mathcal{C}_{t+1}\}$, and that the latter is a mixture space.

[17] For $t = 0$, the measure m_0 over S_1 that we construct can be denoted instead by p_0, as in the desired representation.

Step 2: Each $\succeq_t$ satisfies Indifference to Randomization, that is,

$$(c, F_t) \sim_t (c, co\,(F_t)). \tag{A.6}$$

Let $t = T-1$, corresponding to the 3-period setting in [5]. Then $\succeq_t$ is defined on $C_{T-1} \times (\mathcal{M}\,(C_T \times \mathcal{C}_T))^{S_T}$, which, as noted above, is a mixture space. Hence IR is implied by Order, Continuity and Independence (see Dekel, Lipman and Rustichini [3, Lemma 1]).[18]

However, the domain of $\succeq_t$ is not a mixture space if $t < T-1$. Fortunately, we can invoke Kopylov [9] to prove (A.6).[19] He extends the GP theorem to a domain, consisting of hierarchies of menus, that corresponds to our setting when the state space S is a singleton and when consumption occurs only at the terminal time. In achieving this extension, Kopylov proves a counterpart of (A.6) for his setting (see his Appendix B, especially Lemma B.4 and its discussion). The multiplicity of states and the presence of intermediate consumption are not germane to the validity of (A.6), and Kopylov's arguments are readily adapted to accommodate these features.

Step 3: The order $\succeq_t$ can be represented by $\widehat{\mathcal{W}}_t\,(\cdot)$ having the form

$$\widehat{\mathcal{W}}_t\,(c, F) \;=\; u_t^*(c) + \Sigma_{s_{t+1}} U_{t+1}^*\,(F\,(s_{t+1})\,, s_{t+1})\,, \tag{A.7}$$

where $u_t(\cdot)$ and $U_{t+1}^*\,(\cdot, s_{t+1})$ are nonconstant, continuous and linear on C_t and $\mathcal{M}\,(C_{t+1} \times \mathcal{C}_{t+1})$ respectively, and where

$$U_{t+1}^*\,(M, s_{t+1}) = U_{t+1}^*\,(co\,(M)\,, s_{t+1})\,, \text{ for } M \in \mathcal{M}\,(C_{t+1} \times \mathcal{C}_{t+1})\,. \tag{A.8}$$

To prove this, restrict attention first to $C_t \times \mathcal{D}_t$. Each F in $\mathcal{D}_t$ maps S_{t+1} into $\mathcal{M}^c\,(C_{t+1} \times \mathcal{D}_{t+1})$, the collection of *convex* (and closed) subsets of the mixture space $C_{t+1} \times \mathcal{D}_{t+1}$. But $\mathcal{M}^c\,(C_{t+1} \times \mathcal{D}_{t+1})$ is a mixture space. Since $\succeq_t$ satisfies Order, Continuity and Independence on $C_t \times \mathcal{D}_t$, $\succeq_{t|C_t \times \mathcal{D}_t}$ admits a utility representation by some $\widehat{\mathcal{W}}_t : C_t \times \mathcal{D}_t \to \mathbb{R}^1$ having the form (A.7) when restricted to $C_t \times \mathcal{D}_t$; additivity across c and F can be established as in (A.4), while the additive separability across states follows as in [10, Propn. 7.4], for example. Use (A.8) to extend (A.7) to all of $C_t \times \mathcal{C}_t$. Indifference to Randomization (Step 2) implies that $\widehat{\mathcal{W}}_t\,(\cdot)$ represents $\succeq_t$ on $C_t \times \mathcal{C}_t$.

Let $\succeq_{t\,|s_{t+1}}$ on $\mathcal{M}\,(C_{t+1} \times \mathcal{C}_{t+1})$ be the preference represented by $U_{t+1}^*\,(\cdot, s_{t+1})$.

Step 4: $\succeq_{t\,|s_{t+1}}$ satisfies GP axioms suitably translated to $\mathcal{M}\,(C_{t+1} \times \mathcal{C}_{t+1})$. Thus by their theorem and the extension provided by Kopylov [9],[20]

[18] Their result is formulated for preference defined on menus of lotteries, but the same argument can be used for menus of any compact metric mixture space. The contingent nature of menus in our case is of no significance because mixtures are defined statewise.

[19] We are grateful to Igor Kopylov for pointing out this line of argument.

[20] GP work with a domain of menus of lotteries. Their theorem would apply directly if we had adopted the larger domain obtained by replacing (2.2) with $F_t : S_{t+1} \longrightarrow \mathcal{M}\,(\Delta\,(C_{t+1} \times \mathcal{C}_{t+1}))$. However, adding an extra layer of lotteries can be avoided by invoking Kopylov, suitably extended to accommodate a finite (nonsingleton) state space and intermediate consumption. (His Temporal Set-Betweenness axiom is satisfied by our preference $\succeq_{t\,|s_{t+1}}$, by Lemma A.4 and Set-Betweenness.)

$$U^*_{t+1}(M, s_{t+1}) = \max_{(c,F)\in M} \left\{U^{GP}_{t+1}(c, F, s_{t+1}) + V^{GP}_{t+1}(c, F, s_{t+1})\right\} - \max_{(c',F')\in M} V^{GP}_{t+1}(c', F', s_{t+1}),$$

for some $U^{GP}_{t+1}(\cdot, s_{t+1})$ and $V^{GP}_{t+1}(\cdot, s_{t+1})$, continuous and linear functions on $C_{t+1} \times \mathcal{C}_{t+1}$. The subscript t indicates that these functions may depend also on the history s^t_1 underlying $\succeq_t$.

Step 5: Show that for some $A(s_{t+1}) > 0$,

$$U^{GP}_{t+1}(c, F, s_{t+1}) + V^{GP}_{t+1}(c, F, s_{t+1}) = A(s_{t+1}) \left[u(c) + \delta \int_{S_{t+2}} U_{t+2}(F_{t+1}(s_{t+2}), s_{t+2})\, dm_{t+1}\right] \tag{A.9}$$

By Risk Preference and State Independence, for any c, H, s_{t+1} there exists $\ell, \ell' \in \mathcal{L}_{t+1}$ such that

$$(c, \left[H_{-s_{t+1}},\ \{\ell\}\right]) \succ_t (c, \left[H_{-s_{t+1}},\ \{\ell'\}\right]) \text{ and } \ell \succ_{t+1} \ell'.$$

It follows from Sophistication that

$$(c, \left[H_{-s_{t+1}},\ \{\ell, \ell'\}\right]) \succ_t (c, \left[H_{-s_{t+1}},\ \{\ell'\}\right]).$$

In particular, the preference $\succeq_{t\,|s_{t+1}}$ on $\mathcal{M}(C_{t+1} \times \mathcal{C}_{t+1})$ satisfies $\{\ell, \ell'\} \succ_{t|s_{t+1}} \{\ell'\}$. By Step 4 this preference has a $(U^{GP}_{t+1}, V^{GP}_{t+1})$ representation, and thus by Sophistication and Order, Continuity and Independence for $\succeq_{t+1}$, Lemma A.2 implies that $U^{GP}_{t+1}(\cdot, s_{t+1}) + V^{GP}_{t+1}(\cdot, s_{t+1})$ represents $\succeq_{t+1}$. By the induction hypothesis, $\succeq_{t+1}$ is represented also by $\mathcal{W}_{t+1}(\cdot)$, and since both functions are continuous and linear, they must be cardinally equivalent. Thus (A.9) follows wlog.

Step 6: Let $V_{t+1}(c, F, s_{t+1}) = \frac{1}{A(s_{t+1})} V^{GP}_{t+1}(c, F, s_{t+1})$ and show that

$$V_{t+1}(c, F, s_{t+1}) = w_{t+1}(c, s_{t+1}) + \Sigma_{s_{t+2}} v_{t+1}(F(s_{t+2}), s_{t+1}, s_{t+2}), \tag{A.10}$$

where $w_{t+1}(\cdot, s_{t+1})$ and each $v_{t+1}(\cdot, s_{t+1}, s_{t+2})$ is continuous and linear on C_{t+1} and $\mathcal{M}(C_{t+2} \times \mathcal{C}_{t+2})$ respectively.

The function $M \longmapsto V_{t+1}\left(c, [F_{-s_{t+2}}, M], s_{t+1}\right)$ gives the (temptation) utility of the indicated consumption and contingent menu pair as a function of the menu M provided in state s_{t+2}. Similarly for the function $M \longmapsto \overline{U}_{t+1}\left(c, [F_{-s_{t+2}}, M], s_{t+1}\right)$, where

$$\overline{U}_{t+1}(c, F, s_{t+1}) = \frac{1}{A(s_{t+1})} U^{GP}_{t+1}(c, F, s_{t+1}).$$

Recall the order $\succeq_{t\,|s_{t+1}, s_{t+2}}$ defined prior to Lemma A.3. For any given c and F, it is represented by

$$L \longmapsto \max_{M\in L} \{\overline{U}_{t+1}\left(c, [F_{-s_{t+2}}, M], s_{t+1}\right) + V_{t+1}\left(c, [F_{-s_{t+2}}, M], s_{t+1}\right)\} - \max_{M'\in L} V_{t+1}\left(c, [F_{-s_{t+2}}, M'], s_{t+1}\right), \tag{A.11}$$

for any closed $L \subset \mathcal{M}(C_{t+2} \times \mathcal{C}_{t+2})$. By Risk Preference, State Independence and Lemma A.4, $\overline{U}_{t+1}\left(c, [F_{-s_{t+2}}, \cdot], s_{t+1}\right)$ is nonconstant, and so by Lemma A.3, $\succeq_{t\,|s_{t+1},s_{t+2}}$ is strategically rational. By Lemma A.1(a), if $V_{t+1}\left(c, [F_{-s_{t+2}}, \cdot], s_{t+1}\right)$ is nonconstant then it is ordinally equivalent to $\overline{U}_{t+1}\left(c, [F_{-s_{t+2}}, \cdot], s_{t+1}\right) + V_{t+1}\left(c, [F_{-s_{t+2}}, \cdot], s_{t+1}\right)$, which by Step 5 is ordinally equivalent to $U_{t+2}(\cdot, s_{t+2})$. Thus, if $V_{t+1}\left(c, [F_{-s_{t+2}}, \cdot], s_{t+1}\right)$ is nonconstant, then for all $M, M' \in \mathcal{M}(C_{t+2} \times \mathcal{C}_{t+2})$,

$$\begin{aligned}
V_{t+1}\left(c, [F_{-s_{t+2}}, M], s_{t+1}\right) &\geq V_{t+1}\left(c, [F_{-s_{t+2}}, M'], s_{t+1}\right) \Longleftrightarrow \\
U_{t+2}(M, s_{t+2}) &\geq U_{t+2}(M', s_{t+2}) \Longleftrightarrow \\
U_{t+2}(co(M), s_{t+2}) &\geq U_{t+2}(co(M'), s_{t+2}) \Longleftrightarrow \\
V_{t+1}\left(c, [F_{-s_{t+2}}, co(M)], s_{t+1}\right) &\geq V_{t+1}\left(c, [F_{-s_{t+2}}, co(M')], s_{t+1}\right),
\end{aligned} \tag{A.12}$$

where use has been made of (A.8). On the other hand, if $V_{t+1}\left(c, [F_{-s_{t+2}}, \cdot], s_{t+1}\right)$ is constant, then the equivalence of the first and last lines is clear. Conclude that for every F, c and s_{t+2},

$$V_{t+1}\left(c, [F_{-s_{t+2}}, M], s_{t+1}\right) = V_{t+1}\left(c, [F_{-s_{t+2}}, co(M)], s_{t+1}\right).$$

Repeated application of this equality for all states in S_{t+2} yields

$$V_{t+1}(c, F, s_{t+1}) = V_{t+1}(c, co(F), s_{t+1}),$$

a form of indifference to randomization for V_{t+1}. Thus one can argue as in Step 3 to derive (A.10).

Step 7: Show that for some $\gamma(s_{t+1}) \geq 0$, continuous linear function $w(\cdot, s_{t+1})$ on C_{t+1} and $q_{t+1} \in \Delta(S_{t+2})$,

$$V_{t+1}^{GP}(c, F, s_{t+1}) = A(s_{t+1})\left[w_{t+1}(c, s_{t+1}) + \gamma(s_{t+1}) \int_{S_{t+2}} U_{t+2}(M, s_{t+2})\, dq_{t+1}(s_{t+2})\right].$$

Begin by providing structure on each $v_{t+1}(\cdot, s_{t+1}, s_{t+2})$ in (A.10) - show that

$$v_{t+1}(\cdot, s_{t+1}, s_{t+2}) = a(s_{t+1}, s_{t+2})\, U_{t+2}(\cdot, s_{t+2}) + b(s_{t+1}, s_{t+2}), \tag{A.13}$$

for some $a(s_{t+1}, s_{t+2}) \geq 0$. Given (A.10), we can refine (A.12) into the statement that if $v_{t+1}(\cdot, s_{t+1}, s_{t+2})$ is nonconstant, then $v_{t+1}(\cdot, s_{t+1}, s_{t+2})$ is ordinally equivalent to $U_{t+2}(\cdot, s_{t+2})$. Given continuity and linearity of both functions, (A.13) holds for some $a(s_{t+1}, s_{t+2}) > 0$. If $v_{t+1}(\cdot, s_{t+1}, s_{t+2})$ is constant, then (A.13) holds with $a(s_{t+1}, s_{t+2}) = 0$.

Define $\gamma(s_{t+1})$ and the measure q_{t+1} over S_{t+2} by

$$\begin{aligned}
\gamma(s_{t+1}) &= \sum_{S_{t+2}} a(s_{t+1}, s_{t+2}) \geq 0, \\
q_{t+1}(s_{t+2}) &= \begin{cases} \frac{a(s_{t+1}, s_{t+2})}{\gamma(s_{t+1})} & \text{if } \gamma(s_{t+1}) > 0 \\ m_{t+1}(s_{t+2}) & \text{otherwise} \end{cases}.
\end{aligned}$$

Then

$$V_{t+1}\left(c,F,s_{t+1}\right)= \; w_{t+1}(c,s_{t+1})+\gamma(s_{t+1})\int_{S_{t+2}}U_{t+2}\left(M,s_{t+2}\right)dq_{t+1}(s_{t+2})+k,$$

where $k=\sum_{S_{t+2}} b\left(s_{t+1},s_{t+2}\right)$. Set $k=0$ wlog.

Step 8: Show that for some $0<\alpha_{t+1}(s_{t+1})\leq 1$,

$$V_{t+1}^{GP}\left(c,F,s_{t+1}\right)=A(s_{t+1})\;\left(1-\alpha_{t+1}(s_{t+1})\right)\;\left(u(c)+\delta\int_{S_{t+2}}U_{t+2}\left(F\left(s_{t+2}\right),s_{t+2}\right)\,dq_{t+1}\right). \tag{A.14}$$

By Risk Preference and State Independence, $U_{t+1}^{GP}(\ell,s_{t+1})$ is ordinally (and hence cardinally) equivalent to the continuous linear function $\ell\longmapsto\Sigma_{t+1}^{T+1}\;\delta^{\tau-(t+1)}u\left(\ell_{\tau}\right)$. Thus wlog

$$U_{t+1}^{GP}(\ell,s_{t+1})=A(s_{t+1})\alpha_{t+1}(s_{t+1})\left[\Sigma_{t+1}^{T+1}\;\delta^{\tau-(t+1)}u\left(\ell_{\tau}\right)\right] \tag{A.15}$$

for some $\alpha_{t+1}(s_{t+1})>0$. By Step 5,

$$U_{t+1}^{GP}\left(\ell,s_{t+1}\right)+V_{t+1}^{GP}\left(\ell,s_{t+1}\right)=A(s_{t+1})\left[\Sigma_{t+1}^{T+1}\;\delta^{\tau-(t+1)}u\left(\ell_{\tau}\right)\right].$$

Thus, $V_{t+1}^{GP}(\ell,s_{t+1})=A(s_{t+1})(1-\alpha_{t+1}(s_{t+1}))\left[\Sigma_{t+1}^{T+1}\;\delta^{\tau-(t+1)}u\left(\ell_{\tau}\right)\right]$, and by Step 7,
$w_{t+1}(\ell_{t+1},s_{t+1})+\gamma(s_{t+1})\left[\Sigma_{t+2}^{T+1}\;\delta^{\tau-(t+2)}u\left(\ell_{\tau}\right)\right]=$
$(1-\alpha_{t+1}(s_{t+1}))\left(u(\ell_{t+1})+\delta\left[\Sigma_{t+2}^{T+1}\;\delta^{\tau-(t+2)}u\left(\ell_{\tau}\right)\right]\right)\implies$

$$w_{t+1}(\ell_{t+1},s_{t+1})-(1-\alpha_{t+1}(s_{t+1}))u(\ell_{t+1})=\left[(1-\alpha_{t+1}(s_{t+1}))\delta-\gamma(s_{t+1})\right]\left[\Sigma_{t+2}^{T+1}\;\delta^{\tau-(t+2)}u\left(\ell_{\tau}\right)\right].$$

Since $u\left(\cdot\right)$ is nonconstant, deduce that $(1-\alpha_{t+1}(s_{t+1}))\delta=\gamma(s_{t+1})$ and $w_{t+1}(\ell_{t+1},s_{t+1})=(1-\alpha_{t+1}(s_{t+1}))u(\ell_{t+1})$. If $\gamma(s_{t+1})=0$, then $\delta>0$ implies $w_{t+1}(\ell_{t+1},s_{t+1})=0$, which yields (A.14) with $\alpha_{t+1}(s_{t+1})=1$. On the other hand, if $\gamma(s_{t+1})>0$, then $\delta>0$ implies (A.14) with $\alpha_{t+1}(s_{t+1})<1$.

Step 9: Show that the unique measure p_{t+1} over S_{t+2} satisfying $m_{t+1}=\alpha_{t+1}p_{t+1}+(1-\alpha_{t+1})q_{t+1}$ is a probability measure with full support and furthermore that

$$U_{t+1}^{GP}\left(c,F,s_{t+1}\right)=A(s_{t+1})\;\alpha_{t+1}(s_{t+1})\;\left(u(c)+\delta\int_{S_{t+2}}U_{t+2}\left(F\left(s_{t+2}\right),s_{t+2}\right)\,dp_{t+1}\right). \tag{A.16}$$

Steps 5 and 8 yield (A.16), given that p_{t+1} satisfies $m_{t+1}=\alpha_{t+1}p_{t+1}+(1-\alpha_{t+1})q_{t+1}$. Show next that p_{t+1} is a probability measure with full support. The definition of p_{t+1} implies that $\sum_{s_{t+2}}p_{t+1}(s_{t+2})=1$. To see that $p_{t+1}(s_{t+2})>0$ for all s_{t+2}, note that $U_{t+2}\left(\cdot,s_{t+2}\right)$ is nonconstant (by the induction hypothesis) and that for any $s_{t+1},s_{t+2},c',c,M',M,F$ and G,

$U_{t+2}(M', s_{t+2}) \geq U_{t+2}(M, s_{t+2}) \Longleftrightarrow$
$(c, [F_{-s_{t+2}}, M']) \succeq_{t+1} (c, [F_{-s_{t+2}}, M]) \Longleftrightarrow^*$
$(c', [G_{-s_{t+1}}, \{(c, [F_{-s_{t+2}}, M'])\}]) \succeq_t (c', [G_{-s_{t+1}}, \{(c, [F_{-s_{t+2}}, M])\}]) \Longleftrightarrow$
$U_{t+1}^{GP}(c, [F_{-s_{t+2}}, M'], s_{t+1}) \geq U_{t+1}^{GP}(c, [F_{-s_{t+2}}, M], s_{t+1}) \Longleftrightarrow$
$U_{t+2}(M', s_{t+2}) p_{t+1}(s_{t+2}) \geq U_{t+2}(M, s_{t+2}) p_{t+1}(s_{t+2}),$
where the equivalence $\Longleftrightarrow^*$ is implied by Lemma A.4.

Step 10: Complete the inductive step.

Since $A_t(s_{t+1})\alpha_{t+1}(s_{t+1}) > 0$ for all s_{t+1}, we have $\sum_{s_{t+1}} A_t(s_{t+1})\alpha_{t+1}(s_{t+1}) > 0$. Consider the positive affine transformation of $\widehat{\mathcal{W}_t}$ defined by $\mathcal{W}_t(c, F) =$

$$\begin{aligned}\frac{\delta}{\sum_{s_{t+1}} A_t(s_{t+1})\alpha_{t+1}(s_{t+1})}\widehat{\mathcal{W}_t}(c,F) &= \frac{\delta}{\sum_{s_{t+1}} A_t(s_{t+1})\alpha_{t+1}(s_{t+1})} u^*_{t+1}(c) + \\ &\quad \frac{\delta}{\sum_{s_{t+1}} A_t(s_{t+1})\alpha_{t+1}(s_{t+1})} \Sigma_{s_{t+1}} U^*_{t+1}(F(s_{t+1}), s_{t+1}),\end{aligned}$$

for all $(c, F) \in C_t \times \mathcal{C}_t$. Obviously, $\mathcal{W}_t(\cdot)$ represents $\succeq_t$. Define

$$\begin{aligned}u_{t+1}(c) &\equiv \frac{\delta}{\sum_{s_{t+1}} A_t(s_{t+1})\alpha_{t+1}(s_{t+1})}\; u^*_{t+1}(c), \\ U_{t+1}(M_{t+1}, s_{t+1}) &\equiv \frac{1}{A_t(s_{t+1})\alpha_{t+1}(s_{t+1})}\; U^*_{t+1}(M, s_{t+1}), \\ m_t(s_{t+1}) &= \frac{A_t(s_{t+1})\alpha_{t+1}(s_{t+1})}{\sum_{s_{t+1}} A_t(s_{t+1})\alpha_{t+1}(s_{t+1})} > 0 \text{ for each } s_{t+1}.\end{aligned}$$

Then m_t has full support and

$$\mathcal{W}_t(c, F) = u_{t+1}(c) + \delta \int_{S_{t+1}} U_{t+1}(F(s_{t+1}), s_{t+1})\, dm_t(s_{t+1}), \qquad F_t \in \mathcal{C}_t,$$

where $U_{t+1}(M_{t+1}, s_{t+1}) =$

$$\max_{(c, F_{t+1}) \in M_{t+1}} \left\{ \begin{array}{c} u(c) + \delta \int_{S_{t+2}} U_{t+2}(F_{t+1}(s_{t+2}), s_{t+2})\, dp_{t+1} \\ + \frac{(1-\alpha_{t+1})}{\alpha_{t+1}} \left(u(c) + \delta \int_{S_{t+2}} U_{t+2}(F_{t+1}(s_{t+2}), s_{t+2})\, dq_{t+1} \right) \end{array} \right\}$$
$$- \max_{(c', F'_{t+1}) \in M_{t+1}} \frac{(1-\alpha_{t+1})}{\alpha_{t+1}} \left\{ u(c') + \delta \int_{S_{t+2}} U_{t+2}(F'_{t+1}(s_{t+2}), s_{t+2})\, dq_{t+1}(s_{t+2}) \right\}.$$

It remains to show that $u_{t+1}(\cdot) = u(\cdot)$. By Risk Preference and the representation $\mathcal{W}_t(\cdot)$, the following functions are ordinally equivalent:

$$\begin{aligned}(c, \ell) &\longmapsto u(c) + \delta \left[\Sigma_{t+1}^{T+1} \delta^{\tau-(t+1)} u(\ell_\tau) \right] \\ (c, \ell) &\longmapsto u_{t+1}(c) + \delta \left[\Sigma_{t+1}^{T+1} \delta^{\tau-(t+1)} u(\ell_\tau) \right].\end{aligned}$$

Since both are continuous linear functions, they must be cardinally equivalent. An argument analogous to that used in Step 8 yields the desired result, and concludes the proof of sufficiency.

The proof for uniqueness is similar to that in [5], and thus is omitted. ■

B. APPENDIX: Proofs for Specific Biases

Proof of Corollary 4.1: Necessity of Prior-Bias: Given the representation, let

$$\mu_t(s_{t+2}) = \int p_{t+1}\left(s_{t+2} \mid s'_{t+1}\right) dm_t\left(s'_{t+1}\right).$$

Then the axiom can be translated into the statement:

$$\int \left[U_{t+2}\left(F\left(s_{t+2}\right), s_{t+2}\right) - U_{t+2}\left(G\left(s_{t+2}\right), s_{t+2}\right)\right] dp_{t+1} > 0 \text{ and}$$

$$\int \left[U_{t+2}\left(F\left(s_{t+2}\right), s_{t+2}\right) - U_{t+2}\left(G\left(s_{t+2}\right), s_{t+2}\right)\right] d\mu_t\left(s_{t+2}\right) = 0$$

imply

$$(1-\alpha_{t+1}) \int \left[U_{t+2}\left(F\left(s_{t+2}\right), s_{t+2}\right) - U_{t+2}\left(G\left(s_{t+2}\right), s_{t+2}\right)\right] dq_{t+1} \geq 0.$$

This is obviously satisfied given (4.6).

Sufficiency of Prior-Bias: If $\alpha_{t+1} = 1$, then any q_{t+1} is consistent with a representation for $\succeq_t$, including in particular q_{t+1} as in (4.6) with any λ_{t+1}. Suppose that $\alpha_{t+1} < 1$ and consider contingent menus F and G that lie in $\mathcal{C}_{t+1}^{c,+1}$. They provide perfect commitment, with all uncertainty resolved at $t+2$, and as in Section 2.3, $F(s_{t+2})$ and $G(s_{t+2})$ can be identified with deterministic consumption process $c^F(s_{t+2})$ and $c^G(s_{t+2})$ respectively. It follows that

$$U_{t+2}\left(F\left(s_{t+2}\right), s_{t+2}\right) = \Sigma_{\tau=t+2}\, \delta^{\tau-(t+2)}\, u\left(c_\tau^F\left(s_{t+2}\right)\right) \equiv \widehat{u}\left(c^F\left(s_{t+2}\right)\right),$$

and similarly for G. Write

$$x^{FG}\left(s_{t+2}\right) = \widehat{u}\left(c^F\left(s_{t+2}\right)\right) - \widehat{u}\left(c^G\left(s_{t+2}\right)\right).$$

Then by Prior-Bias,

$$\left[\int x^{FG}\left(s_{t+2}\right) dp_{t+1}\left(s_{t+2} \mid s_{t+1}\right) > 0 \text{ and} \int x^{FG}\left(s_{t+2}\right)\, d\mu_t(s_{t+2}) = 0\,\right] \implies$$

$$\int x^{FG}\left(s_{t+2}\right) dq_{t+1} \geq 0.$$

One can show that $x^{FG}(\cdot)$ can be made to vary sufficiently (over an open neighborhood of zero) as we range over F and G lying in $\mathcal{C}_{t+1}^{c,+1}$. Apply a Theorem of the Alternative [13, p. 34].

The arguments for the other axioms are similar. ■

Proof of Corollary 4.2: The proof is similar to that of the preceding corollary. We point out only that for G in $\mathcal{C}_t^{c,+1}$, $U_{t+1}\left(c_{t+1}, \left(\int G d\Psi_{t+1}\right)\left(s_{t+1}\right), s_{t+1}\right)$

$$= u\left(c_{t+1}\right) + \delta \int_{S_{t+2}} U_{t+2}\left(\int_{S_{t+2}} G\left(s'_{t+2}\right) d\Psi_{t+1}, s_{t+2}\right) dp_{t+1}\left(s_{t+2}\right)$$

$$= u(c_{t+1}) + \delta \int_{S_{t+2}} \int_{S_{t+2}} U_{t+2}\left(G\left(s'_{t+2}\right), s_{t+2}\right) d\Psi_{t+1}\left(s'_{t+2}\right) dp_{t+1}(s_{t+2})$$
$$= u(c_{t+1}) + \delta \int_{S_{t+2}} U_{t+2}\left(G\left(s'_{t+2}\right), s_{t+2}\right) d\Psi_{t+1}\left(s'_{t+2}\right),$$

because $U_{t+2}\left(G\left(s'_{t+2}\right), s_{t+2}\right)$ does not depend on s_{t+2}. ■

C. APPENDIX: Learning in the Long Run

Proof of Theorem 5.1: *(a)* First we show that $\log \mu_t(\theta^*)$ is a submartingale under P^*. Because

$$\log \mu_{t+1}(\theta^*) - \log \mu_t(\theta^*) = \log\left(\left(1-\gamma_{t+1}\right)\tfrac{\ell(s_{t+1}|\theta^*)}{m_t(s_{t+1})} + \gamma_{t+1}\right), \tag{C.1}$$

it suffices to show that

$$E^*\left[\log\left(\left(1-\gamma_{t+1}\right)\tfrac{\ell(s_{t+1}|\theta^*)}{m_t(s_{t+1})} + \gamma_{t+1}\right) \mid \mathcal{S}_t\right] \geq 0, \tag{C.2}$$

where E^* denotes expectation with respect to P^*. By assumption, γ_{t+1} is constant given $\mathcal{S}_t$. Thus the expectation equals

$$\sum_{s_{t+1}} \ell(s_{t+1} \mid \theta^*) \log\left(\left(1-\gamma_{t+1}\right)\tfrac{\ell(s_{t+1}|\theta^*)}{m_t(s_{t+1})} + \gamma_{t+1}\right) \geq$$
$$\sum_{s_{t+1}} \ell(s_{t+1} \mid \theta^*) \left(1-\gamma_{t+1}\right) \log\left(\tfrac{\ell(s_{t+1}|\theta^*)}{m_t(s_{t+1})}\right) =$$
$$\left(1-\gamma_{t+1}\right) \sum_{s_{t+1}} \ell(s_{t+1} \mid \theta^*) \log\left(\tfrac{\ell(s_{t+1}|\theta^*)}{m_t(s_{t+1})}\right) \geq 0$$

as claimed, where both inequalities are due to concavity of $\log(\cdot)$. (The second is the well-known entropy inequality.)

Clearly $\log \mu_t(\theta^*)$ is bounded above by zero. Therefore, by the martingale convergence theorem, it converges $P^* - a.s.$ From (C.1),

$$\log \mu_{t+1}(\theta^*) - \log \mu_t(\theta^*) = \log\left(\left(1-\gamma_{t+1}\right)\tfrac{\ell(s_{t+1}|\theta^*)}{m_t(s_{t+1})} + \gamma_{t+1}\right) \longrightarrow 0$$

and hence $\frac{\ell(s_{t+1}|\theta^*)}{m_t(s_{t+1})} \longrightarrow 1 \; P^* - a.s.$ ■

(b) $E^*\left[\left(\left(1-\gamma_{t+1}\right)\frac{\ell(s_{t+1}|\theta^*)}{m_t(s_{t+1})} + \gamma_{t+1}\right) \mid \mathcal{S}_t\right] = \left(1-\gamma_{t+1}\right) E^*\left[\frac{\ell(s_{t+1}|\theta^*)}{m_t(s_{t+1})} \mid \mathcal{S}_t\right] + \gamma_{t+1} \geq \left(1-\gamma_{t+1}\right) + \gamma_{t+1} = 1$. (The last inequality is implied by the fact that

$$min_X \left\{E^*\left[\tfrac{1}{X(s_{t+1})} \mid \mathcal{S}_t\right] : E^*\left[X(s_{t+1}) \mid \mathcal{S}_t\right] = 1\right\} = 1.$$

The minimization is over random variable X's, $X : S_{t+1} \longrightarrow \mathbb{R}^1_{++}$, and it is achieved at $X(\cdot) = 1$ because $\frac{1}{x}$ is a convex function on $(0, \infty)$.) Deduce that $E^* \left[\frac{\mu_{t+1}(\theta^*)}{\mu_t(\theta^*)} \mid \mathcal{S}_t \right] \geq 1$ and hence that $\mu_t(\theta^*)$ is a submartingale. By the martingale convergence theorem,

$$\mu_\infty(\theta^*) \equiv lim \mu_t(\theta^*) \quad \text{exists } P^* - a.s.$$

Claim: $\mu_\infty(\theta^*) > 0$ on a set with positive P^*-probability: By the bounded convergence theorem,

$$E^* \mu_t(\theta^*) \longrightarrow E^* \mu_\infty(\theta^*);$$

and $E^* \mu_t(\theta^*) \nearrow$ because $\mu_t(\theta^*)$ is a submartingale. Thus $\mu_0(\theta^*) > 0$ implies that $E^* \mu_\infty(\theta^*) > 0$, which proves the claim.

It suffices now to show that if $\mu_\infty(\theta^*) > 0$ along a sample path s_1^∞, then forecasts are eventually correct along s_1^∞. But along such a path, $\frac{\mu_{t+1}(\theta^*)}{\mu_t(\theta^*)} \longrightarrow 1$ and hence

$$\left(1 - \gamma_{t+1}\right) \left(\frac{\ell(s_{t+1}|\theta^*)}{m_t(s_{t+1})} - 1 \right) \longrightarrow 0.$$

By assumption, $\left(1 - \gamma_{t+1}\right)$ is bounded away from zero. Therefore,

$$\left(\frac{\ell(s_{t+1}|\theta^*)}{m_t(s_{t+1})} - 1 \right) \longrightarrow 0. \qquad \blacksquare$$

Part (c) calls for two examples.

Example 1: Convergence to wrong forecasts may occur with P^*-positive probability when $\gamma_{t+1} < 0$, even where γ_{t+1} is $\mathcal{S}_t$-measurable (overreaction); in fact, we take $(\alpha_{t+1}, \lambda_{t+1}) = (\alpha, \lambda)$ and hence also $\gamma_{t+1} = \gamma$ to be constant over time and states.

Think of repeatedly tossing an unbiased coin that is viewed at time 0 as being either unbiased or having probability of Heads equal to b, $0 < b < \frac{1}{2}$. Thus take $S = \{H, T\}$ and $\ell(H \mid \theta) = \theta$ for $\theta \in \Theta = \{b, \frac{1}{2}\}$. Assume also that

$$1 < -\gamma < \frac{b}{\frac{1}{2} - b}. \tag{C.3}$$

The inequality $\gamma < -1$ indicates a sufficient degree of overreaction.

To explain the reason for the other inequality, note that the model requires that (ν_t) solving (5.6) be a probability measure (hence non-negative valued). This is trivially true if $\lambda_{t+1} \geq 0$ but otherwise requires added restrictions: $\nu_{t+1} \geq 0$ if

$$\frac{\ell(s_{t+1} \mid \theta)}{m_t(s_{t+1})} = \frac{dBU(\mu_t; s_{t+1})(\theta)}{d\mu_t} \geq -\frac{\lambda_{t+1}}{1 + \lambda_{t+1}}.$$

In the present example $min_{s,\theta} \frac{\ell(s|\theta)}{m_t(s)} \geq 2b$, and thus it suffices to have

$$-\frac{\lambda}{1+\lambda} \leq 2b. \tag{C.4}$$

Because only values for α in $(0,1]$ are admissible, $\gamma = \lambda(1-\alpha)$ is consistent with (C.4) if and only if $-\gamma < b/\left(\frac{1}{2}-b\right)$.

We show that if (C.3), then

$$m_t(\cdot) \longrightarrow \ell(\cdot \mid b) \quad \text{as } t \longrightarrow \infty,$$

with probability under P^* at least $\frac{1}{2}$.

Abbreviate $\mu_t\left(\frac{1}{2}\right)$ by μ_t^*.

Claim 1: $\mu_\infty^* \equiv lim \mu_t^*$ exists $P^* - a.s.$ and if $\mu_\infty^* > 0$ for some sample realization s_1^∞, then $m_t(H) \longrightarrow \frac{1}{2}$ and $\mu_t^* \longrightarrow 1$ along s_1^∞. (The proof is analogous to that of part (b).) Deduce that

$$\mu_\infty^* \in \{0,1\} \quad P^* - a.s.$$

Claim 2: $f(z) \equiv \left[(1-\gamma)\frac{\frac{1}{2}}{z}+\gamma\right]\left[(1-\gamma)\frac{1-\frac{1}{2}}{(1-z)}+\gamma\right] \leq 1$, for all $z \in [b, \frac{1}{2}]$. Argue that $f(z) \leq 1$ $\Longleftrightarrow$ $g(z) \equiv [(1-\gamma)+2\gamma z]\,[(1-\gamma)+2\gamma(1-z)] - 4z(1-z) \leq 0$. Compute that $g\left(\frac{1}{2}\right) = 0$, $g'\left(\frac{1}{2}\right) = 0$ and g is concave because $\gamma < -1$. Thus $g(z) \leq g(0) = 0$.

Claim 3: $E^*\left[\log\left((1-\gamma)\frac{\ell\left(s_{t+1} \mid \frac{1}{2}\right)}{m_t(s_{t+1})}+\gamma\right) \mid \mathcal{S}_t\right]$

$= \frac{1}{2}\log\left((1-\gamma)\frac{\frac{1}{2}}{b+\left(\frac{1}{2}-b\right)\mu_t^*}+\gamma\right) + \frac{1}{2}\log\left((1-\gamma)\frac{1-\frac{1}{2}}{\left(1-b-\left(\frac{1}{2}-b\right)\mu_t^*\right)}+\gamma\right)$

$= \frac{1}{2}\log\left(f\left(b+\left(\frac{1}{2}-b\right)\mu_t\left(\frac{1}{2}\right)\right)\right) \leq 0$, by Claim 2.

By Claim 1, it suffices to prove that $\mu_\infty^* = 1$ $P^* - a.s.$ is impossible. Compute that

$$\mu_t^* = \mu_0^*\left[\Pi_{k=0}^{t-1}\left((1-\gamma)\frac{\ell\left(s_{k+1} \mid \frac{1}{2}\right)}{m_k(s_{k+1})}+\gamma\right)\right],$$

$$log\mu_t^* = log\mu_0^* + \Sigma_{k=0}^{t-1}\, log\left((1-\gamma)\frac{\ell\left(s_{k+1} \mid \frac{1}{2}\right)}{m_k(s_{k+1})}+\gamma\right)$$

$$= log\mu_0^* + \Sigma_{k=0}^{t-1}\left(logz_{k+1} - E\left[logz_{k+1} \mid \mathcal{S}_k\right]\right) + \Sigma_{k=0}^{t-1} E\left[logz_{k+1} \mid \mathcal{S}_k\right],$$

where $z_{k+1} = (1-\gamma)\frac{\ell\left(s_{k+1} \mid \frac{1}{2}\right)}{m_k(s_{k+1})}+\gamma$. Therefore, $log\mu_t^* \geq \frac{1}{2}log\mu_0^*$ iff

$$\Sigma_{k=0}^{t-1}\left(logz_{k+1} - E\left[logz_{k+1} \mid \mathcal{S}_k\right]\right) \geq -\tfrac{1}{2}log\mu_0^* - \Sigma_{k=0}^{t-1} E\left[logz_{k+1} \mid \mathcal{S}_k\right] \equiv a_k.$$

By Claim 3, $a_k > 0$. The random variable $logz_{k+1} - E\left[logz_{k+1} \mid \mathcal{S}_k\right]$ takes on two possible values, corresponding to $s_{k+1} = H$ or T, and under the truth they are equally likely and average to zero. Thus

$$P^*\left(logz_{k+1} - E\left[logz_{k+1} \mid \mathcal{S}_k\right] \geq a_k\right) \leq \tfrac{1}{2}.$$

Deduce that

$$P^*\left(log\mu_t^* \geq \frac{1}{2}log\mu_0^*\right) \leq \frac{1}{2}$$

and hence that

$$P^*\left(log\mu_t^* \longrightarrow 0\right) \leq \frac{1}{2}. \quad \blacksquare$$

Example 2: Convergence to wrong forecasts may occur with P^*-positive probability when $\gamma_{t+1} > 0$ (Positive Prior-Bias), if γ_{t+1} is only $\mathcal{S}_{t+1}$-measurable.

The coin is as before - it is unbiased, but the agent does not know that and is modeled via $S = \{H, T\}$ and $\ell(H \mid \theta) = \theta$ for $\theta \in \Theta = \{b, \frac{1}{2}\}$. Assume further that α_{t+1} and λ_{t+1} are such that

$$\gamma_{t+1} \equiv \lambda_{t+1}(1-\alpha_{t+1}) = \begin{cases} w & \text{if } s_{t+1} = H \\ 0 & \text{if } s_{t+1} = T, \end{cases}$$

where $0 < w < 1$. Thus, from (5.9), the agent updates by Bayes' Rule when observing T but attaches only the weight $(1-w)$ to last period's prior when observing H. Assume that

$$w > 1 - 2b.$$

Then

$$m_t(\cdot) \longrightarrow \ell(\cdot \mid b) \quad \text{as } t \longrightarrow \infty,$$

with probability under P^* at least $\frac{1}{2}$.

The proof is similar to that of Example 1. The key is to observe that $E^*\left[\log\left((1-\gamma)\frac{\ell\left(s_{t+1} \mid \frac{1}{2}\right)}{m_t(s_{t+1})} + \gamma\right) \mid \mathcal{S}_t\right] \leq 0$ under the stated assumptions.

The proof of Theorem 5.2 requires the following lemmas:

Lemma C.1. *(Freedman (1975)) Let $\{z_t\}$ be a sequence of uniformly bounded $\mathcal{S}_t$-measurable random variables such that for every $t \geqslant 1$, $E^*(z_{t+1}|\mathcal{S}_t) = 0$. Let $V_t^* \equiv VAR(z_{t+1}|\mathcal{S}_t)$ where VAR is the variance operator associated with P^*. Then,*

$$\sum_{t=1}^{n} z_t \text{ converges to a finite limit as } n \to \infty,\ P^*\text{-a.s. on } \left\{\sum_{t=1}^{\infty} V_t^* < \infty\right\}$$

and

$$\sup_n \sum_{t=1}^{n} z_t = \infty \text{ and } \inf_n \sum_{t=1}^{n} z_t = -\infty,\ P^*\text{-a.s. on } \left\{\sum_{t=1}^{\infty} V_t^* = \infty\right\}.$$

Definition C.2. *A sequence of $\{x_t\}$ of $\mathcal{S}_t$-measurable random variables is eventually a submartingale if, $P^* - a.s.$, $E^*(x_{t+1}|S_t) - x_t$ is strictly negative at most finitely many times.*

Lemma C.3. *Let $\{x_t\}$ be uniformly bounded and eventually a submartingale. Then, $P^* - a.s.$, x_t converges to a finite limit as t goes to infinity.*

Proof. Write

$$x_t = \sum_{j=1}^{t} \left(r_j - E^*\left(r_j|\mathcal{S}_{j-1}\right)\right) + \sum_{j=1}^{t} E^*\left(r_j|\mathcal{S}_{j-1}\right) + x_0, \text{ where } r_j \equiv x_j - x_{j-1}.$$

By assumption, $P^* - a.s.$, $E^*\left(r_j|\mathcal{S}_{j-1}\right)$ is strictly negative at most finitely many times. Hence, $P^* - a.s.$,

$$\inf_t \sum_{j=1}^{t} E^*\left(r_j|\mathcal{S}_{j-1}\right) > -\infty.$$

Given that x_t is uniformly bounded, $P^* - a.s.$,

$$\sup_t \sum_{j=1}^{t} z_j < \infty, \text{ where } z_j \equiv r_j - E^*\left(r_j|\mathcal{S}_{j-1}\right).$$

It follows from Freedman's result that $P^* - a.s.$,

$$\sum_{j=1}^{t} z_j \text{ converges to a finite limit as } t \to \infty.$$

It now follows from x_t uniformly bounded that $\sup_t \sum_{j=1}^{t} E^*\left(r_j|\mathcal{S}_{j-1}\right) < \infty$. Because $E^*\left(r_j|\mathcal{S}_{j-1}\right)$ is strictly negative at most finitely many times,

$$\sum_{j=1}^{t} E^*\left(r_j|\mathcal{S}_{j-1}\right) \text{ converges to a finite limit as } t \to \infty.$$

Therefore, $P^* - a.s.$, x_t converges to a finite limit as t goes to infinity. ■

Proof of Theorem 5.2:

Claim 1: Define $f(\theta, m) = \sum_k \theta_k^* \frac{\theta_k}{m_k}$ on the interior of the $2K$-simplex. There exists $\delta' \in \mathbb{R}_{++}^K$ such that

$$\mid \theta_k - \theta_k^* \mid < \delta_k' \text{ for all } k \implies \quad f(\theta, m) - 1 \geq -\underline{\gamma} K^{-1} \sum_k \mid m_k - \theta_k \mid.$$

Proof: $f(\theta, \theta) = 1$, $f(\theta, \cdot)$ is convex and hence

$$\begin{aligned} f(\theta, m) - 1 &\geq \sum_{k \neq K} \left(\frac{\partial f(\theta, m)}{\partial m_k} - \frac{\partial f(\theta, m)}{\partial m_K} \right) |_{m=\theta} \ (m_k - \theta_k) \\ &= \sum_{k \neq K} \left(-\tfrac{\theta_k^*}{\theta_k} + \tfrac{\theta_K^*}{\theta_K} \right) (m_k - \theta_k). \end{aligned}$$

But the latter sum vanishes at $\theta = \theta^*$. Thus argue by continuity.

Given any $\delta \in \mathbb{R}^K_{++}$, $\delta << \delta'$, define $\Theta^* = (\theta^* - \delta, \theta^* + \delta) \equiv \Pi_{k=1}^K (\theta_k^* - \delta_k, \theta_k^* + \delta_k)$ and $\mu_t^* = \Sigma_{\theta \in \Theta^*} \mu_t(\theta)$.

Claim 2: Define $m_t^*(s^k) = \Sigma_{\theta \in \Theta^*} \theta_k \mu_t(\theta) / \mu_t^*(\theta)$. Then

$$\mid m_t(s^k) - m_t^*(s^k) \mid \leq 1 - \mu_t^*.$$

Proof: $m_t(s^k) - m_t^*(s^k) = \frac{\Sigma_{\theta \in \Theta^*} \theta_k \mu_t(\theta)}{\mu_t^*}(\mu_t^* - 1) + \Sigma_{\theta \notin \Theta^*} \theta_k \mu_t(\theta)$. Therefore, $(\mu_t^* - 1) \leq m_t^*(s^k)(\mu_t^* - 1) = \frac{\Sigma_{\theta \in \Theta^*} \theta_k \mu_t(\theta)}{\mu_t^*}(\mu_t^* - 1) \leq m_t(s^k) - m_t^*(s^k) \leq \Sigma_{\theta \notin \Theta^*} \theta_k \mu_t(\theta) \leq 1 - \mu_t^*$.

Claim 3: For any $\delta << \delta'$ as above,

$$\sum_k \theta_k^* \frac{m_t^*(s^k)}{m_t(s^k)} - 1 \geq -\underline{\gamma}(1 - \mu_t^*).$$

Proof: Because $\mid m_t^*(s^k) - \theta_k^* \mid < \delta_k < \delta_k^{'}$, we have that

$$\sum_k \theta_k^* \frac{m_t^*(s^k)}{m_t(s^k)} - 1 \geq -\underline{\gamma} K^{-1} \sum_k \mid m_t(s^k) - m_t^*(s^k) \mid.$$

Now Claim 3 follows from Claim 2.

Compute that

$$E^*\left[\mu_{t+1}(\theta) \mid \mathcal{S}_t\right] = (1 - \gamma_{t+1}) \left[\sum_k \theta_k^* \frac{\theta_k}{m_t(s^k)}\right] \mu_t(\theta) + \gamma_{t+1} E^*\left[\psi_{t+1}(\theta) \mid \mathcal{S}_t\right], \tag{C.5}$$

where use has been made of the assumption that γ_{t+1} is $\mathcal{S}_t$-measurable. Therefore,

$$E^*\left[\mu_{t+1}^*(\theta) \mid \mathcal{S}_t\right] - \mu_t^* = (1 - \gamma_{t+1}) \sum_k \left(\theta_k^* \frac{m_t^*(s^k)}{m_t(s^k)}\right) \mu_t^* + \gamma_{t+1} \Sigma_{\theta \in \Theta^*} E^*\left[\psi_{t+1}(\theta) \mid \mathcal{S}_t\right] - \mu_t^*$$

$$= (1 - \gamma_{t+1}) \left[\sum_k \left(\theta_k^* \frac{m_t^*(s^k)}{m_t(s^k)}\right) - 1\right] \mu_t^* + \gamma_{t+1} \Sigma_{\theta \in \Theta^*} E^*\left[\psi_{t+1}(\theta) \mid \mathcal{S}_t\right] - \gamma_{t+1} \mu_t^*.$$

By the LLN, $P^* - a.s.$ for large enough t the frequency of s^k will eventually be θ_k^* and

$$\Sigma_{\theta \in \Theta^*} E^*\left[\psi_{t+1}(\theta) \mid \mathcal{S}_t\right] = 1.$$

Eventually along any such path,

$$E^*\left[\mu_{t+1}^*(\theta) \mid \mathcal{S}_t\right] - \mu_t^* = (1 - \gamma_{t+1}) \left[\sum_k \left(\theta_k^* \frac{m_t^*(s^k)}{m_t(s^k)}\right) - 1\right] \mu_t^* + \gamma_{t+1}(1 - \mu_t^*)$$

$$\geq \left[-\underline{\gamma}\left(1-\gamma_{t+1}\right)\mu_t^* + \gamma_{t+1}\right]\left(1-\mu_t^*\right) \geq 0,$$

where the last two inequalities follow from Claim 3 and the hypothesis $\underline{\gamma} \leq \gamma_{t+1}$.

Hence (μ_t^*) is eventually a P^*-submartingale. By Lemma C.3, $\mu_\infty^* \equiv \overline{lim}\mu_t^*$ exists $P^* - a.s.$ Consequently, $E^*\left[\mu_{t+1}^*(\theta) \mid \mathcal{S}_t\right] - \mu_t^* \longrightarrow 0$ $P^* - a.s.$ and from the last displayed equation, $\left[-\underline{\gamma}\left(1-\gamma_{t+1}\right)\mu_t^* + \gamma_{t+1}\right]\left(1-\mu_t^*\right) \longrightarrow 0$ $P^* - a.s.$ It follows that $\mu_\infty^* = 1$. Finally, $m_t(\cdot) = \int \ell(\cdot \mid \theta)\, d\mu_t$ eventually remains in $\Theta^* = (\theta^* - \delta, \theta^* + \delta)$.

Above δ is arbitrary. Apply the preceding to $\delta = \frac{1}{n}$ to derive a set Ω_n such that $P^*(\Omega_n) = 1$ and such that for all paths in Ω_n, m_t eventually remains in $\left(\theta^* - \frac{1}{n}, \theta^* + \frac{1}{n}\right)$. Let $\Omega \equiv \bigcap_{n=1}^{\infty} \Omega_n$. Then, $P^*(\Omega) = 1$ and for all paths in Ω, m_t converges to θ^*. ■