Dynamic Concern for Misspecification*

Giacomo Lanzani (Harvard)

August 2023 - Click here for the latest version.

Abstract

We consider an agent who posits a set of probabilistic models for the payoff-relevant outcomes. The agent has a prior over this set but fears the actual model is omitted and hedges against this possibility. The concern for misspecification is endogenous: If a model explains the previous observations well, the concern attenuates. We show that different static preferences under uncertainty (subjective expected utility, maxmin, robust control) arise in the long run, depending on how quickly the agent becomes unsatisfied with unexplained evidence and whether they are misspecified. The misspecification concern's endogeneity naturally induces behavior cycles, and we characterize the limit action frequency. This model is consistent with the empirical evidence on monetary policy cycles and choices in the face of complex tax schedules. Finally, we axiomatize in terms of observable choices this decision criterion and how quickly the agent adjusts their misspecification concern.

1 Introduction

Bayesian rationality requires that an agent uncertain about the data-generating process postulates multiple probabilistic descriptions of the environment and uses Bayes rule to adjust their relative weights. However, even rational agents may fear that they are misspecified and that none of these descriptions is correct. This concern is remarkably natural in complex and high-dimensional settings, where uncertainty needs to be simplified to obtain well-behaved optimization and learning procedures.

For example, none of the model economies considered by a central bank perfectly describes the underlying data-generating process for output and inflation. Similarly, the consumer response models that a firm uses to set prices and qualities are unlikely to include one that considers all relevant decision factors. Moreover, the diffusion of complex and not explicitly described machine learning algorithms naturally creates new reasons for misspecification. Indeed, consumers increasingly rely on automated recommendations. Although they may have some conjecture on how the alternative's features translate into a score or a "match quality" with their profile, they certainly do not consider the specific algorithm used by these recommendation systems. Misspecification is even more relevant when dealing with entirely novel issues, such as those faced by a regulatory body that tries to mitigate the effect of climate change using theoretical models that take into account human impacts never experienced in history.

Misspecification has been analyzed from two distinct perspectives. On the one hand, several papers have studied the long-run implications of subjective expected utility (SEU) maximizers learning with misspecified beliefs (see, e.g., Esponda and Pouzo, 2016, Fudenberg et al., 2021, Frick et al., 2023, and the references therein). These works assume that the agents have no concern about being misspecified. Here we show that the absence of such concern is normatively unappealing, as it can induce long-run average payoffs lower than a safe guarantee. It also seems descriptively unrealistic, as the widely documented ambiguity-averse behavior may be seen as a way to hedge against the incorrect specification of the model. On the other hand, the robust control literature in macroeconomics pioneered by Hansen and Sargent (2001) considers agents who fear model misspecification. In particular, the first axioms-based decision criterion that accounts for model misspecification was proposed in Cerreia-Vioglio et al. (2022).[1]

---

[1] It has as a particular case the robust control model of Hansen and Sargent (2001) axiomatized by Strzalecki (2011). Since in Strzalecki (2011) the reference probability is subjective, it can also be interpreted as an axiomatization of robust prior analysis, see Hansen and Sargent (2022).

This work reconciles these approaches and shows how popular decision criteria such as maxmin expected utility, robust control preferences, and subjective expected utility arise as the limit behavior of an agent concerned about misspecification and learning about the actual data-generating process (DGP). We consider an agent that repeatedly chooses among actions whose payoffs have an unknown distribution. This choice is taken using an average of robust control assessments, where each assessment takes a different structured model as the benchmark. We introduce endogeneity in the misspecification concern: the better the structured models explain the past, the less concerned the agent is.

There are two critical determinants for the long-run dynamics: whether the agent is correctly specified and how demanding they are in evaluating their models' performance. First, we consider the case of a correctly specified agent. In that case, the behavior converges to a self-confirming equilibrium, regardless of how demanding the agent is in evaluating their model. A self-confirming equilibrium means that they play an SEU best reply to a belief supported over the data-generating processes that are observationally equivalent to the true one given the chosen action.

Instead, to characterize the limit behavior under misspecification, a taxonomy of how demanding the agent is turns out to be crucial. We first establish a normative benchmark: a "statistically sophisticated" behavior that performs a likelihood ratio evaluation of the agent's subjective model. More precisely, we show that the achievement of two desirable properties uniquely characterizes such behavior: global safety (i.e., guaranteeing the minmax payoff even when misspecified) and consistency under correct specification (i.e., no regret when correctly specified).[2]

We allow departures from this normative benchmark to obtain descriptive predictions on the effect of an endogenous misspecification concern. We consider agents that are too demanding in evaluating the models' performance (this case includes believers in the Law of Small Numbers, Tversky and Kahneman, 1971, that treat failures in explaining early realizations as a statistician treats long-run failures). Similarly, we allow the opposite case in which the agent is too lenient in evaluating their model and attributes too much unexplained evidence to sampling variability.

We then characterize the long-run behavior of these different types of misspecified agents. The actions of the lenient type converge to a Berk-Nash equilibrium, i.e., to an SEU best reply to beliefs supported on the models closest to the actual data-generating process. Instead, overemphasis on the model's failures in explaining the data by the

---

[2]Moreover, we observe that SEU maximization and the original robust control of Hansen and Sargent (2001) fail to jointly satisfy these requirements.

demanding type induces convergence to a maxmin best reply to the models that are absolutely continuous with respect to the true one.

In contrast, a statistically sophisticated type maintains a nonextreme concern for misspecification. If their behavior converges, it converges to a robust control best reply to the models closest to the actual data-generating process. Moreover, the misspecification concern is endogenously determined by how well the best models fit the evidence generated by the limit action.

Therefore, our learning results provide several novel predictions about the relation between uncertainty attitudes and other individual traits.[3] First, the extent of long-run uncertainty aversion positively correlates with the agent being initially misspecified and their belief in the Law of Small Numbers. Second, these correlations are causal: repeated failures to explain the data (misspecification) and demanding evaluation of these failures induce the agent to shift to cautious behavior. Third, even keeping constant the misspecification and understanding of probability rules, the limit uncertainty attitudes are stochastic. Initial realizations leading to a limit action with consequences poorly explained by the agent's models induce a long-run uncertainty aversion higher than realizations leading to a limit action whose consequences are well explained.

We thus use the equilibrium behavior predicted by an endogenous concern for misspecification to rationalize the labor supply in the face of complex tax schedules documented in Rees-Jones and Taubinsky (2020). In particular, they show that around 40% of the agents have beliefs corresponding to a heuristic that simplifies the tax schedule to a linear one but that 20% fewer agents act accordingly to this heuristic. This is predicted by an endogenous concern for misspecification, as agents with an incorrect model are less prone to base their decisions on the conclusions they reach within the model.

In general, the behavior of a statistically sophisticated type is not guaranteed to converge. Indeed, it is possible that their behavior cycles between phases of different misspecification concerns. Still, we characterize the limit action frequency and concern for misspecification. We apply this result to revisit the cyclical behavior of monetary policies documented in Sargent (1999) and Sargent (2008). Intuitively, the cycles have the following structure. The agent spends some time playing an action whose consequences are well explained by one of their structured models (a conservative monetary policy in the application). Playing this action lowers the concern for misspecification and eventually leads to a more misspecification-vulnerable action (a more aggressive monetary policy).

---

[3]The empirical study of the correlation between behavioral biases is an active area of recent development. See, e.g., Dean and Ortoleva (2019), Snowberg and Yariv (2021), and the references therein.

Failures to explain the distribution of outcomes observed under this action lead to a return to the more misspecification robust action.

We also obtain two results that provide a testable foundation for the model employed in the learning part of the paper: An axiomatization of the static average robust control criterion and testable axioms for when the agent is of the lenient, statistically sophisticated, or demanding type. Two primary axioms pin down the static decision criterion. The first is a weaker form of the Sure-Thing Principle imposed only on bets on the data-generating process (e.g., bets on the urn composition) and bets conditional on the data-generating process (e.g., bets on the ball color conditional on having been told the urn composition). The second requires that conditional on being told the best-fitting model, the agent is equally concerned about misspecification, regardless of which one it is.

For the dynamic representation, a dynamic consistency axiom on the acts that bet on the data-generating process is shown to guarantee Bayesian updating over models. More interestingly, the preference adjustment of a statistically sophisticated type is pinned down by a novel Asymptotic Frequentism axiom, requiring arbitrarily similar preferences conditional to sufficiently long histories with the same outcome frequency.

The rest of the paper is structured as follows. Section 2 introduces the average robust control decision criterion and how preferences are adjusted. Section 3 studies what attitudes toward model failures induce good payoff performance and provides a learning foundation for the different uncertainty attitudes. Section 4 characterizes the limit frequency of time spent using the different actions when the behavior does not converge and applies the result to a central banking problem. Section 5 provides the axiomatization to the static decision criterion and how preferences are updated. Section 6 discusses the related literature and possible extensions. Section 7 concludes. All proofs are collected in the Appendix.

2 Decision Criterion

2.1 Static Decision Criterion

We describe the criterion used in the repeated decision problem and defer its axiomatization to Section 5. We consider an agent who chooses from a finite number of actions $a \in A$ and let Y be a compact metric space representing the set of possible outcomes. The agent has a continuous utility index $u : A \times Y \to \mathbb{R}$ over the action-outcome pairs that captures their preference when the uncertainty is resolved. However, the realized outcome is

stochastic and endogenous as each action $a \in A$ induces an objective probability measure $p_a^* \in \Delta(Y)$ over outcomes.[4]

Subjective Beliefs The agent correctly believes that the map from actions to probability distributions over outcomes is fixed and depends only on their current action. Still, they do not know $p^* = (p_a^*)_{a \in A}$ and deal with this uncertainty in a quasi-Bayesian way. The agent postulates a set $Q \subseteq \Delta(Y)^A$ of *structured models,* i.e., action-dependent probability measures over outcomes $q = (q_a)_{a \in A}$. They have a prior belief $\mu \in \Delta(Q)$ with support Q that describes the relative likelihood assigned to these models. For example, the agent may be a central bank that considers a Keynesian Samuelson-Solow model where the monetary policy affects the unemployment rate or a new classical Lucas-Sargent model with no systematic effect of inflation on unemployment.[5]

We must impose a few regularity conditions.

Assumption 1. Q is compact and for every $a \in A$: (i) For all $q \in Q$, $\tilde{q}_a := \frac{\mathrm{d}q_a}{\mathrm{d}p_a^*}$ exists, is continuous, and is p_a^*-a.s. bounded away from 0, uniformly in Q, (ii) For p_a^*-almost every $y \in Y$ the map $q \mapsto \tilde{q}_a(y)$ is continuous.

Condition (i) allows us to compute the relevant expectations while allowing for both discrete and continuous outcome spaces and guarantees that no subjective model of the agent is ruled out in finite time.[6] Continuity of the map from models to outcome distributions is a standard requirement for parametric models.

A myopic Bayesian agent with complete trust in their models evaluates action a according to its subjective expected utility (see, e.g., Cerreia-Vioglio et al., 2013b):

$$\int_Q \mathbb{E}_{q_a}[u(a, y)] \, \mathrm{d}\mu(q).$$

That is, they compute a two-stage expectation of the utility function: they evaluate the

---

[4]For every subset C of a metric space, we denote as $\Delta(C)$ the Borel probability measures on C, endowed with the topology of weak convergence of measures.

[5]This formulation follows the recent literature on misspecified learning in assuming that both the true data-generating process and the subjective models the agent considers are i.i.d. conditionally on the agent's behavior. This makes the true extent of misspecification time-invariant and "measurable". It is important to notice that this time invariance is often relaxed in the literature on dynamic decisions with robust control preferences that follows Hansen and Sargent (2001).

[6]Condition (i) also plays a technical role in guaranteeing the existence of the equilibrium concepts we consider. It is known it can be relaxed; see Anderson et al. (2022), but this relaxation comes at the cost of requiring nonstandard analysis techniques (where nonstandard means using infinitesimal numbers), something beyond this paper's scope.

utility of the action given the candidate model q, $\mathbb{E}_{q_a}[u(a,y)]$, and then they average over the models with weights given by their subjective belief μ.

However, we are interested in agents concerned with the possibility that none of these models is the exact description of the data-generating process but only a valid approximation, i.e., that are concerned that there is no $q \in Q$ with $q = p^*$. Therefore, in the spirit of the robustness criterion advocated by Hansen and Sargent (2001), they penalize actions that perform poorly under alternative distributions that are close in relative entropy $R(\cdot||\cdot)$ to some of the structured models.[7]

With this, an agent evaluates each action $a \in A$ accordingly to the *average robust control* criterion:

$$\int_Q \min_{p_a \in \Delta(Y)} \left(\mathbb{E}_{p_a}[u(a,y)] + \frac{1}{\lambda} R(p_a||q_a) \right) d\mu(q) \tag{1}$$

where $\lambda > 0$ is a parameter that trade-offs between decision robustness and performance under the structured models.[8]

The original robust control model introduced by Hansen and Sargent (2001) is the case in which μ is a Dirac measure (that in macroeconomics applications is often assumed to satisfy rational expectations, i.e., to be degenerate on the actual data-generating process). As described in Hansen et al. (2006), this case corresponds to when "[...] a maximizing player ('the decision maker') chooses a best response to a malevolent player ('nature') who can alter the stochastic process within prescribed limits. The minimizing player's malevolence is the maximizing player's tool for analyzing the fragility of alternative decision rules." Equation (1) follows Hansen and Sargent (2007) and Cerreia-Vioglio et al. (2022) in extending this interpretation to a situation in which the agent is still uncertain about the best-approximating model (i.e., μ is nondegenerate), allowing the malevolent nature to alter each of the candidate structured models.

The representation adopts the distinction between two levels of uncertainty. At the first level, given a probabilistic model q, the uncertainty about the exact specification of the model is captured by minimizing the expected utility for probabilities that are not too far away from q. At a higher level, the agent is also uncertain about the identity of the best structured model and posits a prior probability μ over them. While the higher level of uncertainty is already present under subjective expected utility, the lower level captures

---

[7]For every $p, q \in \Delta(Y)$, $p \gg q$ means that q is absolutely continuous with respect to p, and $p \sim q$ means that they are mutually absolutely continuous. Recall that for every $p, p' \in \Delta(Y)$, $R(p||p') = \int_Y \log\left(\frac{dp}{dp'}\right) dp$ if $p' \gg p$ and $R(p||p') = \infty$ otherwise.

[8]Lemma 1 justifies the use of a min rather than an inf in equation (1) and throughout the paper.

the agent's concern for misspecification.

2.2 Preference Evolution

The average robust control criterion of equation (1) describes how the agent chooses for a *given* belief and level of misspecification concern. However, the behavior responds to the received information. Formally, time is discrete, and a history is a finite vector of past actions and outcomes. In particular, the set of histories of length $t \in \mathbb{N}$ is $\mathcal{H}_t = (A \times Y)^t$, and the set of all histories is $\mathcal{H} = \bigcup_{t=0}^{\infty} \mathcal{H}_t$. We will denote with $\mathbf{a}_t, \mathbf{y}_t$, and $\mathbf{h}_t$ the random variables corresponding to the action, outcome, and history at time t, and we use the non-bold version for their realizations.

On the one hand, we stick to the classical treatment of beliefs about the possible data-generating processes. We let the belief be updated through standard Bayesian updating. That is, for every measurable subset C of Q, we denote by

$$\mu(C \mid (a^t, y^t)) = \frac{\int_C \prod_{\tau=1}^t \tilde{q}_{a_\tau}(y_\tau) \mathrm{d}\mu(q)}{\int_Q \prod_{\tau=1}^t \tilde{q}_{a_\tau}(y_\tau) \mathrm{d}\mu(q)} \tag{Bayes Rule}$$

the subjective belief the agent obtains using Bayes rule after history $(a^t, y^t) \in \mathcal{H}_t$.[9]

On the other hand, we introduce an endogenous and time-evolving concern for misspecification, i.e., we let λ depend on the realized history through a function $\Lambda : \mathcal{H} \to \mathbb{R}_+$.

Log-Likelihood Ratio In particular, we want to capture the idea that the concern for misspecification is a function of how well the structured models explain the current history. In statistics, a common measure of fit of a set of distributions Q against a set of unstructured alternatives $N(Q) \subseteq \Delta(Y)^A$ is the log-likelihood ratio:[10]

$$LLR((a^t, y^t), Q) = -\log \left(\frac{\max_{q \in Q} \prod_{\tau=1}^t \tilde{q}_{a_\tau}(y_\tau)}{\max_{p \in N(Q)} \prod_{\tau=1}^t \tilde{p}_{a_\tau}(y_\tau)} \right) \qquad \forall t \in \mathbb{N}, \forall (a^t, y^t) \in \mathcal{H}_t.$$

Here we want to take a conservative approach and not impose structure over the set

---

[9]By Assumption 1 (i), the posterior is well-defined after every positive probability history. We allow for arbitrary belief revisions after events with zero ex-ante subjective probability.

[10]The Neyman-Pearson Lemma establishes the performance of the log-likelihood ratio test under correct specification. At the same time, Foutz and Srivastava (1977) and Vuong (1989) contain the classical results about the informativeness of the LLR under misspecification. Schwartzstein and Sunderam (2021) is a recent paper that models agents in a persuasion problem who perform model selection using this statistic. Lemma 1 justifies the use of max in the definition of the LLR under Assumption 2.

of alternative unstructured distributions $N(Q)$ used to evaluate the model's fit. If Y is finite and all outcomes have positive probability, there is a natural way to do so, i.e., to consider as the set of alternatives unstructured distributions the entire (action-indexed) simplex $\Delta(Y)^A$. However, considering a completely unrestricted set of distributions with a continuum of outcomes leads to an utterly uninformative test of the model, as the (discrete) empirical distribution is an infinitely better fit to itself than any continuous distribution, i.e., the log-likelihood ratio always returns $+\infty$. To maintain informativeness, $N(Q)$ must then include only distributions that are mutually absolutely continuous with respect to the ones in Q. In particular, all our results are invariant to the $N(Q)$ choice as long as the following assumption is satisfied.

Assumption 2. (i) $p^* \in N(Q) \supseteq Q$. (ii) For every $a \in A$, the family of densities $\{\tilde{p}_a : p \in N(Q)\}$ is a compact set of continuous functions.

We require that the unstructured set is a relaxation sufficiently large to include the actual distribution and a continuity condition that rules out a Q that only contains continuous distributions and an $N(Q)$ that includes discrete distributions.

With this, an important role will be played by the rule

$$\Lambda(h_t) = \frac{\text{LLR}(h_t, Q)}{ct} \qquad \forall t \in \mathbb{N}, \forall h_t \in \mathcal{H}_t \tag{2}$$

where $c \in \mathbb{R}$. Observe that under this rule, the concern for misspecification is proportional to the average log-likelihood ratio. This *average* log-likelihood ratio is often used in statistics to measure the extent of model misspecification.[11] Therefore, we informally refer to an agent who uses such a rule as a "statistically sophisticated type". Of course, the estimation goal of a statistician can be very different from that of an agent involved in a repeated decision problem under uncertainty. Proposition 1 confirms that this rule is also a rationality benchmark in repeated decision problems, as it uniquely identifies the behavior that induces no regret when the agent is correctly specified and is always maxmin safe.

---

[11] This use of the *average* LLR complements the classical role of the *total* LLR in deciding whether to reject or accept a model. In particular, Wilks Theorem (see, e.g., Theorem 10.3.3 in Casella and Berger, 2021) shows that *under correct specification,* the likelihood ratio test statistic converges to a χ^2 distribution. However, it says nothing about the distribution of the LLR if the model is misspecified. See Hausman (1978) and the subsequent literature for a complementary approach to the measurement of model misspecification when the statistician can compute a consistent quasi-maximum-likelihood estimator.

3 Long-run Payoffs and Actions

In this section, we study the long-run consequences of using the decision criterion above. Our primary interest is in what attitudes towards unexplained evidence, i.e., what Λ, induce good payoff performance across environments and what are the limit actions and preferences under uncertainty attitudes that arise given a specific attitude.

Let $BR^{\lambda}(\nu)$ denote the set of average robust control best replies to belief ν when the concern for misspecification is λ, i.e.,[12]

$$BR^{\lambda}(\nu) = \operatorname*{argmax}_{a \in A} \int_Q \min_{p_a \in \Delta(Y)} \left(\mathbb{E}_{p_a}[u(a,y)] + \frac{R(p_a||q_a)}{\lambda} \right) d\nu(q) \qquad \forall \lambda \in \mathbb{R}_+, \forall \nu \in \Delta(Q).$$

Also, let

$$BR^{Seu}(\nu) = \operatorname*{argmax}_{a \in A} \int_Q \mathbb{E}_{q_a}[u(a,y)] \, d\nu(q) \qquad \forall \nu \in \Delta(Q)$$

denote the actions that maximize the (classical) subjective expected utility of an agent with belief ν and

$$BR^{Meu}(C) = \operatorname*{argmax}_{a \in A} \inf_{p \in C} \mathbb{E}_{p_a}[u(a,y)]$$

denote the actions preferred by a maxmin agent a la Gilboa and Schmeidler (1989) with models $C \subseteq \Delta(Y)^A$.

A (pure) *policy* is a measurable $\Pi : \mathcal{H} \to A$ that specifies an action for every history. The objective action-contingent probability distribution and a policy Π induce a probability measure $\mathbb{P}_\Pi$ on $(A \times Y)^{\mathbb{N}}$.[13] Our interest is in policies derived from maximizing the value in equation (1) for some rule Λ determining how the concern for misspecification is adjusted.

Definition 1. Policy Π is Λ-*optimal* if for all $h_t \in \mathcal{H}$, $\Pi(h_t) \in BR^{\Lambda(h_t)}(\mu(\cdot|h_t))$.

3.1 Normative Benchmark

We have mentioned that using the rule in equation (2) has the good statistical property of keeping Λ asymptotically informative about the fit of the model. Now, we provide a normative justification for considering it the relevant benchmark of rationality, showing that it satisfies the desirable properties of safety and consistency (cf. Fudenberg and

---

[12]Throughout the paper, we use the convention $0 \cdot \infty = 0$.

[13]We spell out the $\mathbb{P}_\Pi$ derivation in Appendix A.1.1.

Levine, 1995; Hart and Mas-Colell, 2000) across all possible decision problems the agent can face.

Definition 2. Let $\varepsilon > 0$. Λ is ε-*safe* for the decision problem (u, A, Y) if for every Λ-optimal policy Π and DGP $p^* \in \Delta (Y)^A$

$$\liminf_{t \to \infty} \frac{\sum_{i=1}^{t} u\left(\mathbf{a}_i, \mathbf{y}_i\right)}{t} \geq \max_{a \in A} \min_{y \in Y} u\left(a, y\right) - \varepsilon \qquad \mathbb{P}_{\Pi}\text{-a.s.} \tag{3}$$

This is a very mild condition that only requires the agent to obtain an average payoff at least ε close to what they can guarantee against *every* possible outcome. However, when paired with misspecification, ε-safety has a significant bite: a Bayesian SEU agent fails it in many decision problems. Indeed, such failures have been the basis of many critiques of learning under misspecification with Bayesian SEU agents.

Example 1 (Unsafe SEU). *Suppose $A = \{Bet\ Heads,\ Bet\ Tails,\ Out\}$ and $Y = \{Heads,\ Tails\}$. The utility is 0 if Out, 1 if the action matches the outcome, and -1 if there is a mismatch. Each agent's model is an action-independent probability of Heads. So identify $Q = \{0.9, 0.4\}$, and let $p_a^*(Heads) = 0.6$, and $\mu(0.9) = \frac{1}{2} = \mu(0.4)$. The actions of a Bayesian SEU maximizer converge to Bet Tails with an average performance of -0.2 versus a safe payoff of 0 under action Out.*

This is the simplest possible example, but safety also fails in the more economically motivated case of overconfidence, the key aspect being that good statistical fit does not necessarily induces good decisions.

Definition 3. Let $\varepsilon > 0$. Λ is ε-*consistent under correct specification* for the decision problem (u, A, Y) if there exists $\delta > 0$ such that for every Λ-optimal policy Π and DGP $p^* \in \Delta (Y)$

$$\min_{q \in Q} \max_{a \in A} R\left(p_a^* || q_a\right) < \delta \implies \liminf_{t \to \infty} \frac{\sum_{i=1}^{t} u\left(\mathbf{a}_i, \mathbf{y}_i\right)}{t} \geq \max_{a \in A} \mathbb{E}_{p_a^*}\left[u\left(a, y\right)\right] - \varepsilon \qquad \mathbb{P}_{\Pi}\text{-a.s.}$$

ε-consistency under correct specification requires that sufficiently low levels of misspecification (i.e., the existence of a model q with distance less than δ from the true data generating process) cannot induce considerable ex-post regret (i.e., a limit average payoff more than ε lower than the expected payoff of the objectively optimal action). Intuitively, we want that if the misspecification is minor, in the long run, the agent approximately identifies the actual model and starts best replying to it. Therefore, in classical statistics

terms, it can be interpreted as requiring that the long run, the probability of Type I error goes to 0.[14]

Proposition 1. *1. For every $\varepsilon > 0$ and decision problem where (i) $\{q^*\} = \operatorname{argmin}_{q \in Q} Q(a)$ and (ii)*

$$\lim_{n \to \infty} q_a^n = q_a^* \implies \lim_{n \to \infty} q_{a'}^n = q_{a'}^* \qquad \forall a, a' \in A, \forall (q^n)_{n \in \mathbb{N}} \in Q^{\mathbb{N}} \tag{4}$$

there exists $c > 0$ such that if

$$\Lambda(h_t) = \frac{LLR(h_t, Q)}{ct} \qquad \forall t \in \mathbb{N}, \forall h_t \in \mathcal{H}_t,$$

Λ is both ε-safe and ε-consistent under correct specification.

2. There exist $\varepsilon > 0$ and a decision problem that satisfies (i) and (ii) for which there is no ε-safe and ε-consistent under correct specification Λ with either[15]

$$\Lambda(h_t) = o\left(\frac{LLR(h_t, Q)}{t}\right) \qquad \forall (h_t)_{t \in \mathbb{N}} \in \times_{t \in \mathbb{N}} \mathcal{H}_t.$$

or

$$o(\Lambda(h_t)) = \frac{LLR(h_t, Q)}{t} \qquad \forall (h_t)_{t \in \mathbb{N}} \in \times_{t \in \mathbb{N}} \mathcal{H}_t.$$

The safety and consistency conditions we require are weak but are enough to single out the statistically sophisticated type. When the concern for misspecification is adjusted accordingly to equation (2), the combination of Bayesian updating over parameters and dynamically adjusted concern for misspecification is consistent with Savage's distinction between small and large worlds (see pages 82-91 in Savage, 1954). Indeed, Savage advocates reducing the large-world uncertainty to small worlds (for us, the structured Q) where Bayesian updating has appealing properties, but being aware that this description is incomplete and that the agent should evaluate the fit of that simplification (for us, using

---

[14]Observe that, in the language of Fudenberg and Levine (1995), this consistency is not "universal" as it is only required to hold against the DGPs in the support. That is, it captures the idea of a DM willing to incur some long-run payoff loss to use a simpler model. But they want that if the simple model is correct, this loss does not arise.

[15]Recall that given two sequences $(x_n)_{n \in \mathbb{N}}, (x'_n)_{n \in \mathbb{N}}$ of real numbers, $x_n = o(x'_n)$ means that $\lim_{n \to \infty} \frac{x_n}{x_n} = 0$. Here we use the convention that $\frac{0}{0} = 0$, making the definition of o more permissive. Since o will always appear as a requirement for a sequence in the hypothesis of our statements, such convention makes our results stronger and able to cover a larger range of cases.

a test that can measure the failures of this description). In terms of performance, this procedure results in a behavior that is safe and consistent under correct specification.

At the same time, some less normatively appealing but descriptively relevant phenomena are captured by other rules. On the one hand, a rule such that $\lim_{t\to\infty} \frac{\Lambda(h_t)}{\text{LLR}(h_t,Q)/t} = \infty$, e.g., $\Lambda(h_t) = \frac{\text{LLR}(h_t,Q)}{\sqrt{t}}$, overly penalizes minor imperfections of the model, expecting that the frequency quickly converges to its theoretical value, as in the fallacy called the Law of Small Numbers. On the other hand, an agent for which $\lim_{t\to\infty} \frac{\Lambda(h_t)}{\text{LLR}(h_t,Q)/t} = 0$ applies an excessively lenient adjustment to the likelihood ratio statistic and attributes too much of the unexplained evidence to sampling variability. In this regard, Proposition 1 tells us that there are decision problems where any way to adjust the concern for misspecification that is globally more demanding or lenient than the average LLR violates either ε-safety or ε-consistency under correct specification. Moreover, standard SEU maximization is not safe, while always using a maxmin best reply to Q induces a behavior that is not consistent under correct specification.[16]

Safety and consistency for the LLR-based rule are established under the conditions of identification, i.e., that there is a unique minimizer q^* regardless of the action, and a minimal continuity requirement between the consequences induced by different actions under the structured models. It allows for situations where the data-generating process is (subjectively) perceived as exogenous, where the structured models represent different Gaussian shifts or linear intercepts and, more generally, every example in the paper. However, it rules out (subjective) bandit problems with infinite Q, as in that case, structured models that are very similar under an action may be completely different under a different action. This is natural, as we should not expect consistency of an impatient agent in a bandit problem. In the extension of this paper to patient agents, we confirm the consistency of the LLR-based rule in bandit settings for sufficiently patient agents.

3.2 Positive Long-run Predictions

We are interested in the actions that can arise as the long-run behavior of agents with an evolving concern for misspecification. The main results of this section show that we can describe this limit behavior through fixed point conditions involving the agent's action, belief, and concern for misspecification. To this end, let $Q(a) = \text{argmin}_{q\in Q} R(p_a^* || q_a)$ be

---

[16]This observation about the inconsistency of a misspecified single agent complements the results of Fudenberg and Kreps (1993) and Fudenberg and Levine (1995) about the inconsistency of a correctly specified SEU player in games.

the structured models that best fit the actual data-generating process when action a is played.

Definition 4. Action a^* is a:

1. *Self-confirming equilibrium* (SCE) if there exists $\nu \in \Delta(Q)$ with

$$\operatorname{supp}\nu \subseteq \{q \in Q : q_{a^*} = p_{a^*}^*\} \text{ and } a^* \in BR^{Seu}(\nu).$$

2. *Berk-Nash equilibrium* if there exists $\nu \in \Delta(Q)$ with

$$\operatorname{supp}\nu \subseteq Q(a^*) \text{ and } a^* \in BR^{Seu}(\nu).$$

3. *Maxmin equilibrium* if

$$a^* \in BR^{Meu}\left(\left\{p \in \Delta(Y)^A : \exists q \in Q, \forall a \in A, q_a \gg p_a\right\}\right).$$

4. *c-robust equilibrium* if there exists $\nu \in \Delta(Q)$ with

$$\operatorname{supp}\nu \subseteq Q(a^*), \ a^* \in BR^\lambda(\nu), \text{ and } \lambda = \min_{q \in Q} R(p_{a^*}^* \| q_{a^*})/c.$$

Self-confirming equilibrium (Battigalli, 1987 and Fudenberg and Levine, 1993) describes a stable situation where the agent's action is a best reply to a belief that is on-path confirmed, in the sense of being concentrated over models that perfectly match the distribution over outcomes induced by the equilibrium action.

Berk-Nash equilibrium (Esponda and Pouzo, 2016) relaxes the confirmed beliefs condition of SCE by only requiring that the supporting beliefs are concentrated on the models that provide the best fit to the outcome distribution induced by the equilibrium action. Importantly, this fit is not required to be perfect.

In a maxmin equilibrium, the agent evaluates each action under the worst-case scenario that is minimally consistent with their structured descriptions of the environment (i.e., those scenarios that do not assign positive probability to events impossible for every structured models).

c-robust equilibrium is similar to Berk-Nash in requiring best reply to the best-fitting models. However, the best reply is the average robust control, with misspecification concern that decreases in how well the models fit the true DGP at the equilibrium.

We are interested in what actions have a positive probability of becoming the long-run behavior of the agent. The following definition captures this requirement.

Definition 5. Action a is a Λ-*limit action* if there is a Λ-optimal policy Π such that $\mathbb{P}_{\Pi}\left[\sup\{\mathbf{t}\colon \mathbf{a_t} \neq a\} < \infty\right] > 0$.

Our first limit result is a consistency check: Concern for misspecification is irrelevant in environments with a finite number of outcomes if the agent is correctly specified about the consequences induced by the limit action.

Proposition 2. *If Y is finite, a^* is a Λ-limit action with $p_{a^*}^* \in \{q_{a^*}\}_{q \in Q}$, and for every history sequence $(h_t)_{t \in \mathbb{N}} \in \times_{t \in \mathbb{N}} \mathcal{H}_t$*

$$\lim_{t \to \infty} \frac{LLR(h_t, Q)}{t} = 0 \implies \lim_{t \to \infty} \Lambda(h_t) = 0$$

then a^* is a self-confirming equilibrium.

Instead, how quickly the agent becomes unsatisfied with their model plays a key role when misspecified.

Theorem 1. *Let a^* be a Λ-limit action with $p_{a^*}^* \notin \{q_{a^*}\}_{q \in Q}$.*

1. *If*

$$\Lambda(h_t) = o\left(\frac{LLR(h_t, Q)}{t}\right) \qquad \forall (h_t)_{t \in \mathbb{N}} \in \times_{t \in \mathbb{N}} \mathcal{H}_t, \tag{5}$$

   then a^* is a Berk-Nash equilibrium.

2. *If*

$$o(\Lambda(h_t)) = \frac{LLR(h_t, Q)}{t} \qquad \forall (h_t)_{t \in \mathbb{N}} \in \times_{t \in \mathbb{N}} \mathcal{H}_t, \tag{6}$$

   then a^* is a maxmin equilibrium.

3. *If*

$$\Lambda(h_t) = \frac{LLR(h_t, Q)}{ct} \qquad \forall t \in \mathbb{N}, \forall h_t \in \mathcal{H}_t,$$

   then a^* is a c-robust equilibrium.

The theorem characterizes the possible limit actions of all types of agents. At one extreme, the concept of Berk-Nash equilibrium, introduced for subjective expected utility maximizers, is still sufficient to describe the long-run behavior of lenient types. At the

14

other extreme, the repeated failures in explaining the observed data lead demanding agents to a highly pessimistic behavior and consider the worst-case scenario among all the DGPs that are minimally consistent with the structured models.

Finally, if the behavior of the statistically sophisticated type converges, the limit action is a best reply to beliefs that are supported on the relative entropy minimizers. Here the misspecification concern is determined by the relative entropy between the actual DGP and the best-fitting model.

3.3 Equilibrium Illustrations

In this section, we revisit two of the main biases that have been justified as a consequence of misspecified learning (see Esponda and Pouzo, 2016). Within each example, adding an endogenous concern for misspecification predicts a change in a clear direction. However, one bias is reduced while the other is enhanced. Both changes are broadly consistent with the documented evidence. The first example shows how the endogenous misspecification concern moderates the Berk-Nash equilibrium's prediction that a more complicated tax schedule induces a higher labor supply.

Example 2 (Bias Reduction under Misperceived Taxation, Sobel, 1984 and Esponda and Pouzo, 2016). *An agent chooses effort $a \in A$ at cost $c(a)$ and obtains income $z = a + \omega_a$, where ω_a is a stochastic term with $\mathbb{E}_{p_a^*}[\omega_a] = 0$ for all $a \in A$. The agent pays taxes $t = \tau(z) + l\varepsilon_1$, where $\tau : \mathbb{R} \to \mathbb{R}$ is a convex tax schedule. Here $y = (z, t)$, and the payoff is $u(a, y) = z - t - c(a)$. The agent believes in a random coefficient model, $t = (\theta + \varepsilon_{2,a}) z$, in which the marginal and average tax rates are equal. The stochastic terms $\varepsilon_1, \varepsilon_{2,a} \sim N(0, 1)$ measure respectively actual and conjectured uncertain aspects of the tax schedule, and the $(\omega_a, \varepsilon_{2,a})_{a \in A}$ and ε_1 are independent.[17] See Liebman and Zeckhauser (2004) and Rees-Jones and Taubinsky (2020) for the empirical evidence supporting this "schmeduling" bias.*

Simple computations show that $Q(a) \sim \left\{ \mathbb{E}_{p_a^*} \left[\frac{\tau(a + \omega_a)}{a + \omega_a} \right] \right\}$ for l small, i.e., the best fitting marginal taxation is equal to the (lower) average taxation.[18] Therefore, as pointed out by Esponda and Pouzo (2016), in any Berk-Nash equilibrium, the agent ends up exerting higher effort than the optimal. Moreover, the more complex (i.e., convex) the tax code is, the more significant the gap between the average and marginal rate and the higher the excess effort of the agent.

---

[17]Formally, ε normally distributed implies that Y is not compact, in contrast with the primary analysis of the paper. Still, the conclusions below are unaffected by considering ε with a symmetrically truncated normal distribution that allows remaining in our main framework.

[18]See Appendix A.3 for the computations supporting the claims of the examples.

In every c-robust equilibrium, this bias is reduced. To see this observe that since the agent is not perfectly able to explain the equilibrium data, i.e., $\min_{\theta \in \Theta} R\left(p_a^* || q_a^\theta\right) > 0$, they maintain a positive level of concern for misspecification. However, higher efforts are perceived as more exposed to the uncertainty in the marginal rate (as the stochastic tax rate gets multiplied by an, on average, higher z).

Therefore, c-robust equilibrium provides a natural force that reduces the counterintuitive prediction that complicated nonlinear taxation codes induce more effort: failures to rationalize the received tax bill reduce effort. Moreover, the more complicated the tax code is, i.e., the more nonlinear τ is, the larger the correction size. This set of predictions is consistent with Rees-Jones and Taubinsky (2020), where it is shown that around 40% of the agents have beliefs *(elicited in an incentive-compatible way) corresponding with the schmeduling heuristic but that there are 20% fewer agents who* act *accordingly to the heuristic.*[19]

The second example shows that an endogenous concern for misspecification can enhance some biases. In particular, this is the case for Correlation Neglect, a bias that is indeed widely documented (see Enke and Zimmermann, 2019 and the references therein).

Example 3 (Bias Increase under Correlation Neglect, Esponda, 2008). *A buyer with valuation $v \in V$ and a seller submit a (bid) price $a \in A$, and an ask price $s \in S \subseteq \mathbb{R}_+$, respectively. They play a double auction with price at the buyer's bid, so the seller sets their ask s equal to their value, and a sale occurs if the buyer's bid a is at least s. The payoff for the buyer is*

$$u\left(a, v, s\right) = \begin{cases} v - a & a \geq s \\ 0 & otherwise. \end{cases}$$

The buyer mistakenly believes that the ask price and valuation are independent: $Q = \Delta\left(V\right) \times \Delta\left(S\right)$. Easy computations show that for every $a^* \in A$,

$$Q\left(a^*\right) = \left\{q \in Q : \forall a \in A, \forall s \in S, q_a\left(s\right) = p_a^*\left(s\right), q_a\left(v\right) = p_a^*\left(v\right)\right\}.$$

Therefore, in the Berk-Nash equilibrium, the agent makes a bid a^* lower than the optimal one, not realizing that higher successful bids are, on average, associated with higher-quality

---

[19] In this discussion we followed Rees-Jones and Taubinsky (2020) preferred interpretation in terms of an heterogeneous population. They observe that their data are also compatible with all the agents having beliefs induced by the schmeduling heuristic but under-responding to this biased estimation of the marginal tax rate. This explanation is consistent with a c-robust equilibrium and inconsistent with a Berk-Nash equilibrium, too.

goods. In this case, the bias is reinforced in a c-robust equilibrium: a complete unraveling of the market where the buyer bids 0 *is easier to achieve with an endogenous concern for misspecification. The correlation between valuations and prices results in a positive* $\min_{q \in Q} R(p_{a^*}^* || q_{a^*}) > 0$ *and makes the agent less confident in their model. Since offering* 0 *gives a certain payoff, it is less sensitive to the misspecification concern, and, therefore, this positive concern makes market participation less desirable.*

4 Cycles

Part 3 of Theorem 1 provides a necessary condition for the limit actions of the statistically sophisticated type. However, as momentarily illustrated by the monetary policy application of Section 4.1, there is no guarantee that such an action exists. In these cases, we know by Theorem 1 that the agent behavior cannot stabilize. We now propose a generalization of c-robust equilibrium, show that it always exists, and prove that it characterizes a weaker form of behavior convergence. Formally, for every $\alpha \in \Delta(A)$, let

$$Q(\alpha) = \text{argmin}_{q \in Q} \sum_{a \in A} \alpha(a) R(p_a^* || q_a)$$

be the set of parameters with the lowest average relative entropy from the actual data-generating process, where the average is computed using α.

Definition 6. A mixed action $\alpha^* \in \Delta(A)$ is a *mixed c-robust equilibrium* if there exists $\nu \in \Delta(Q)$ with

$$\text{supp}\nu \subseteq Q(\alpha^*), \ \alpha^* \in \Delta\left(BR^\lambda(\nu)\right), \ \text{and} \ \lambda = \min_{q \in Q} \sum_{a \in A} \alpha^*(a) R(p_a^* || q_a) / c.$$

A mixed robust equilibrium allows multiple actions to be played but requires that the beliefs and the concern for misspecification are determined by the probability assigned to each action. Intuitively, suppose actions for which the models in Q do not satisfactorily explain the consequences are played more often. In that case, the mixed action α^* must best reply to a more significant misspecification concern.

Proposition 3. *For every* $c > 0$ *there exists a mixed c-robust equilibrium.*

Existence is established by proving that the conditions characterizing a mixed c-robust (single-agent) equilibrium are equivalent to the ones of a Nash equilibrium in a game

among the agent and two adversarial Nature players. The result is then obtained by showing that this game satisfies the conditions that guarantee existence in Reny (1999).

Theorem 1 assumes convergence and characterizes the possible limit actions. However, there are natural environments where the action process almost surely does not converge. In that case, it is important to study a weaker form of behavior stabilization, i.e., the convergence of the empirical action frequency, which allows for persistent changes in actions and misspecification concerns. Let $\alpha_t(h_t) \in \Delta(A)$ be the empirical action frequency in history h_t, defined as

$$\alpha_t(h_t)(a) = \frac{\sum_{\tau=1}^{t} \mathbb{I}_{\{a\}}(a_\tau)}{t} \qquad \forall a \in A, \forall t \in \mathbb{N}, \forall h_t \in \mathcal{H}_t.$$

Definition 7. A mixed action $\alpha \in \Delta(A)$ is a Λ-*limit frequency* if there is a Λ-optimal policy Π such that $\mathbb{P}_\Pi[\lim_{t\to\infty} \alpha_t(\mathbf{h}_t) = \alpha] > 0$.

The following result shows that mixed robust equilibrium is the relevant equilibrium concept to capture the long-run stabilization of the average time spent playing each action.

Theorem 2. *If*

$$\Lambda(h_t) = \frac{LLR(h_t, Q)}{ct} \qquad \forall t \in \mathbb{N}, \forall h_t \in \mathcal{H}_t$$

and α^* is a Λ-limit frequency, then α^* is a mixed c-robust equilibrium.

To interpret Theorem 2, consider the case where α^* is supported over two actions a, a' such that Q explains very well the consequences of a, —i.e., $\min_{q\in Q} R(p_a^*||q_a)$ is low— but it explains poorly the consequences of a' —i.e., $\min_{q\in Q} R(p_{a'}^*||q_{a'})$ is high. Suppose also that a is a best reply to a high misspecification concern, while a' is a best reply to a low misspecification concern. Then, the agent oscillates between periods with great concern for misspecification, when they play a, and phases in which the excellent data fit leads them to experiment with action a'.

Whenever cycles are involved, a natural concern is whether the agent can predict them and whether they have the incentive to break them.[20] This is not the case in this model for two orders of reasons. First, the oscillations in behavior are stochastic, and the agent cannot predict and anticipate the changes perfectly. Second and more important, although the agent behavior does not converge, whenever $Q(\alpha)$ is a singleton, the agent's preferences converge.[21] They are approaching indifference between all the actions with

---

[20]For example, this is what happens in fictitious play.

[21]A singleton $Q(\alpha)$ is a mild requirement satisfied in many cases. See Fudenberg et al. (2021) for a discussion.

positive frequency. This asymptotic indifference dramatically reduces the incentives to try to detect and break the probabilistic cycles.

4.1   Application: Monetary Policy Cycles

Here we consider a monetary policy model taken from Sargent (1999), Cogley and Sargent (2005), and Sargent (2008) and in particular its adaptation in Battigalli et al. (2022).[22] A central bank is trying to control a two-dimensional consequence, $Y \subseteq \mathbb{R}^2$, where the y_U component is unemployment and the y_π component is inflation. The policy is aggressive $a = 1$ or conservative $a = 0$.[23] The central bank models are parametrized by the vector θ, with the following specification:

$$y_U = \theta_0 + \theta_{1\pi}y_\pi + \theta_{1a}a + \theta_2\varepsilon_U$$
$$y_\pi = a + \theta_3\varepsilon_\pi$$

where ε_U and ε_π are independent, zero-mean random shocks normalized to have the same support $[-1, 1]$. Here $\theta_0 > 0$ is the natural unemployment level, $\theta_{1\pi} < 0$ is the impact of the actual inflation on unemployment, and $\theta_{1a} > 0$ is the impact of the planned inflation on unemployment, a reduced form of the fact that the market participants (partially) incorporate the central bank actions in their inflation expectations. In particular, if $\theta_{1\pi} + \theta_{1a} = 0$, this is a Lucas-Sargent model with no (structural) exploitable employment-inflation trade-off. If $\theta_{1\pi} + \theta_{1a}$ is negative, this is a Samuelson-Solow model with a structural exploitable employment-inflation trade-off.

The agent's model is misspecified in that it misses the fact that an aggressive monetary policy, beyond raising its baseline level, also increases the inflation variability:

$$y_U = \theta_0^* + \theta_{1\pi}^*y_\pi + \theta_{1a}^*a + \theta_2^*\varepsilon_U$$
$$y_\pi = a + \theta_3^*f_a(\varepsilon_\pi)$$

where f_0 is the identity function, while f_1 is a continuous, strictly increasing, and odd function with $f_1(1) = 1$ that is strictly concave on $\mathbb{R}_{++}$, i.e., that amplifies the inflation-specific

---

[22]Spiegler (2020) also considers the effect of a misspecified simpler model in the context of Philipps curve estimation, with the difference being that the misspecification is on the side of the market rather than the central bank.

[23]Two actions are assumed for simplicity, but a finite A is needed to apply our results.

shocks.[24] This form of misspecification is motivated by the findings in Primiceri (2005) and Sims and Zha (2006) and recent inflation consequences of an aggressive monetary policy.

The central bank is endowed with standard quadratic preferences:

$$u(a, (y_U, y_\pi)) = -y_U^2 - y_\pi^2.$$

Assumption 3. i) Some trade-off is present: $(\theta_{1\pi}^* + \theta_{1a}^* + \theta_0^*)^2 + 1 < (\theta_0^*)^2$. ii) Inflation is more volatile than unemployment under the aggressive monetary policy: $\text{essinf}_{p_1^*} u(1, y) < \text{essinf}_{p_0^*} u(0, y)$. iii) Θ is a product set that includes θ^* and for all $\theta \in \Theta$, $(\theta_{1a}, \theta_2, \theta_3) = (\theta_{1a}^*, \theta_2^*, \theta_3^*)$.

Observe that the exploitable trade-off required by (i) may be so small that the reduced inflation variability under a conservative policy makes the latter optimal. Condition (ii) requires that the additional inflation volatility induced by the aggressive policy is enough to have the worst tail payoffs. Condition (iii) allows us to focus on the cycles induced by the oscillation in the concern for misspecification. Without that, one would get the same insights with other oscillations of beliefs that push even more towards cycles, a channel pointed out by Nyarko (1991) in a monopoly pricing setting.

Corollary 1. *There is $\bar{c} > 0$ such that for all $c \leq \bar{c}$*

1. *There is no c-robust equilibrium.*

2. *There exists a mixed c-robust equilibrium.*

3. *The maximal and minimal equilibria are such that $\alpha^*(0)$ is increasing in $\theta_{1\pi}^* + \theta_{1a}^*$.*

Playing the conservative policy is the best reply to a high misspecification concern and θ^* but induces a low concern as its consequences are well explained. In contrast, the aggressive policy is a best reply to a low misspecification concern and θ^* but induces

---

[24]An alternative, more parametric specification would have

$$
\begin{aligned}
y_U &= \theta_0^* + \theta_{1\pi}^* y_\pi + \theta_{1a}^* a + \theta_2^* \varepsilon_U \\
y_\pi &= \left(1 + \sigma_{\pi a}^2 \varepsilon_{\pi a}\right) a + \theta_3^* \varepsilon_\pi
\end{aligned}
$$

where $\varepsilon_{\pi a}$ is an independent error and $\sigma_{\pi a}^2 > 0$. If we let the support of ε_U and ε_π be unbounded, nothing in the analysis below would be affected by a shift to this alternative specification. However, that change would bring us outside the compact Y setting study in the rest of the paper, so we opted for preserving the consistency.

a severe concern. Therefore, the policy cannot stabilize, consistently with the cyclical behavior of monetary policies documented in Sargent (1999), Clarida et al. (2000), and Sargent (2008). We also have some natural comparative statics in the extremal robust equilibria, as a more significant exploitable trade-off between inflation and unemployment induces more time spent using an aggressive monetary policy.[25]

In this application, we purposefully chose one of the most straightforward macroeconomic frameworks to isolate and illustrate the effect of an endogenous concern for misspecification. However, incorporating an endogenous concern for misspecification in more elaborate models is a valuable enterprise. For example, the fact that evidence impacts the trust in the model may be used to explain the observed pattern of initial underreaction to information when only beliefs within a model are adjusted and medium-run overreaction when the belief adjustment compounds with a change in model trust (see Angeletos et al., 2021 and the references therein for a discussion of this pattern).

5 Representation

We next move to characterize the average robust control model in terms of observable choices in an Anscombe-Aumann framework. In line with the literature on decision theory under uncertainty, our goal is to associate the decision criterion in equation (1) with axioms on a binary preference relation over acts.

Before jumping into the details of the axiomatization, we provide a high-level description of the steps involved and the intuitive meaning of the axioms we link to the representation. In terms of observability requirements, we allow the analyst to elicit preferences for bets both on the data-generating process, e.g., the urn composition, and on the actual realization, e.g., the color of the drawn ball.[26] The analysis then has two nested levels: 1) An axiomatization of the static decision criterion, 2) An axiomatization of the changes in the preference parameters, and in particular, the speed of adjustment of the concern for misspecification.

The static decision criterion belongs to the variational class of Maccheroni et al. (2006a). More importantly, within this class, it is identified by a relaxed Sure-Thing

---

[25]It is well-known that non-extremal equilibria are less well-behaved in terms of comparative statics. See Diamond (1982) for a very early example. A supermodularity condition between the concern for misspecification and the conservative policy payoff guarantees equilibrium uniqueness.

[26]This is standard when dealing with multiple sources of uncertainty, see for example Klibanoff et al. (2005) and Gul and Pesendorfer (2014), and Cerreia-Vioglio et al. (2013a) for a general framework and results. We discuss how to relax this requirement in Section 6.4.

Principle: the agent satisfies it for bets that involve the identity of the model (e.g., bets on an urn composition) and for bets on events conditional on the model (e.g., bets on the color after having revealed the urn composition). However, failures of the Sure-Thing principle can happen for acts that involve the realization of the outcome without conditioning on the model (e.g., bets on the color without knowing the urn composition, which are the ones involved in the classical Ellsberg's paradox). The final conceptual axiom involved in the representation of equation (1) is a notion of uniform conditional misspecification concern. It requires that conditional on being told the identity of their best-fitting model, the agent is equally concerned about it not being exact regardless of which one it is.

We consider a collection of binary relations indexed by the observed history to characterize the agent's dynamic preferences. Three other axioms identify the qualitative changes of the preference parameters u, λ, μ. Constant Preference Invariance guarantees that the taste u for uncertain alternatives is stable over time. Dynamic Consistency over Models guarantees that the probability distribution over models is updated in a Bayesian fashion. Finally, we axiomatize the asymptotic speed of adjustment of the misspecification concern. To do so, we need a quantitative notion of how similar two preference relations are, which is defined using an event E and two deterministic and strictly ranked outcomes, x and y, as measuring rods. Loosely speaking, two relations are (x, y, E, ε) similar if their certain equivalents for the binary act that pays x if E realizes and y otherwise are ε close. With this, an Asymptotic Frequentism axiom singles out the statistically sophisticated type: for every (x, y, E, ε), the conditional preferences after sufficiently long sequences of outcomes sharing the same empirical frequency must be (x, y, E, ε)-similar. Conversely, a lenient type asymptotically becomes similar to those SEU preferences that are less misspecification concerned than the initial preference. The demanding type must approach the preferences of a maxmin agent, thus confirming in a decision-theoretic setting the insights of Theorem 1.

5.1 Notation and Preliminaries

The agent evaluates simple acts, i.e., measurable and finite ranged maps from a state space S into a convex set of outcomes X, where S is endowed with a σ-algebra of events Σ. The set of those acts is denoted as $\mathcal{F}$. Given any $x \in X$, $x \in \mathcal{F}$ is the act that delivers x in every state, and in this way, we identify X as the subset of constant acts in $\mathcal{F}$. If $f, g \in \mathcal{F}$, and $E \in \Sigma$, we denote as gEf the simple act that yields $g(s)$ if $s \in E$ and $f(s)$ if $s \notin E$. Since X is convex, for every $f, g \in \mathcal{F}$, and $\gamma \in (0, 1)$, we denote as $\gamma f + (1 - \gamma) g \in \mathcal{F}$ the

simple act that pays $\gamma f(s) + (1-\gamma) g(s)$ for all $s \in S$.

We model the agent's preference with a binary relation $\succsim$ on $\mathcal{F}$. As usual, $\succ$ and $\sim$ denote the asymmetric and symmetric parts of $\succsim$. An event E is *null* if $fEh \sim gEh$ for every $f, g, h \in \mathcal{F}$. An event is *nonnull* if it is not null. For every $E \in \Sigma$, the conditional preference relation $\succsim_E$ is defined by $f \succsim_E g$ if $fEh \succsim gEh$ for some $h \in \mathcal{F}$.

A key concept to understand the concern for misspecification evolution is a notion of being more misspecification concerned from Ghirardato and Marinacci (2002).

Definition 8. Given two preferences $\succsim_1$ and $\succsim_2$ on $\mathcal{F}$, we say that $\succsim_1$ is more concerned with misspecification than $\succsim_2$ if, for each $f \in \mathcal{F}$ and each $x \in X$, $f \succsim_1 x$ implies $f \succsim_2 x$.

5.2 Decision Criterion

When formalized in terms of a binary relation, the average robust control decision criterion reads as follows.

Definition 9. A tuple (u, Q, μ, λ) is an *average robust control representation* of the preference relation $\succsim$ if $u : X \to \mathbb{R}$ is a nonconstant affine function, $Q \subseteq \Delta(S)$ is a nonempty set, $\mu \in \Delta(Q)$, $\lambda \geq 0$, and for all $f, g \in \mathcal{F}$

$$f \succsim g \iff \mathbb{E}_\mu \left[\min_{p \in \Delta(S)} \left(\int_S u(f)\, dp + \frac{R(p\|q)}{\lambda} \right) \right] \geq \mathbb{E}_\mu \left[\min_{p \in \Delta(S)} \left(\int_S u(g)\, dp + \frac{R(p\|q)}{\lambda} \right) \right]. \tag{7}$$

The average robust control representation is the counterpart of (1) when expressed over acts. An apparent difference is that u here takes as input only outcomes instead of pair of actions and consequences. However, this discrepancy is inconsequential, as in Section 2 we can define a larger space of consequences $\hat{Y} = A \times Y$ that includes both actions and outcomes and transforming each model $p \in \Delta(Y)^A$ into an element of $\hat{p} \in \Delta\left(\hat{Y}\right)^A$ such that $\hat{p}_a(a', y) = 0$ if $a' \neq a$ and $\hat{p}_a(a, y) = p_a(y)$ for all $y \in Y$. Still, this embedding of actions into outcomes muddles the interpretation of the learning results significantly. Therefore we opted to maintain the distinction explicit at the cost of some visual discrepancy between equations (1) and (7).[27]

---

[27]See Fishburn (1970) Chapter 12.1 for a more detailed discussion of the equivalence of a formulation with exogenously given states and one where states are maps from actions into consequences.

5.3 Static Axioms

Our first axiomatic step is a static one. We characterize in terms of behavioral axioms an agent that evaluates accordingly to equation (7) the acts whose consequences are obtained in the same period and before any new information is received.

Axiom 1 (Variational Axiom). ***Weak Order.***

*** Weak Certainty Independence.*** *If $f, g \in \mathcal{F}$, $x, x' \in X$, $\gamma \in (0,1)$, and $\gamma f + (1 - \gamma) x \succsim \gamma g + (1 - \gamma) x$, then $\gamma f + (1 - \gamma) x' \succsim \gamma g + (1 - \gamma) x'$.*

*** Continuity.*** *If $f, g, h \in \mathcal{F}$ the sets $\{\gamma \in [0,1] : \gamma f + (1 - \gamma) g \succsim h\}$ and $\{\gamma \in [0,1] : h \succsim \gamma f + (1 - \gamma) g\}$ are closed.*

*** Monotonicity.*** *If $f, g \in \mathcal{F}$, and $f(s) \succsim g(s)$ for all $s \in S$, then $f \succsim g$.*

*** Uncertainty Aversion.*** *If $f, g \in \mathcal{F}$, $\gamma \in (0,1)$, and $f \sim g$, then $g + \gamma(f - g) \succsim f$.*

*** Nondegeneracy.*** *$f \succ g$ for some $f, g \in \mathcal{F}$.*

*** Weak Monotone Continuity.*** *If $f, g \in \mathcal{F}$, $x \in X$, $(E_n)_{n \in \mathbb{N}} \in \Sigma^{\mathbb{N}}$ with $f \succ g$, $E_1 \supseteq E_2 \supseteq ...$ and $\cap_{n \in \mathbb{N}} E_n = \emptyset$, then there exists $n_0 \in \mathbb{N}$ such that $x E_{n_0} f \succ g$.*

Maccheroni et al. (2006a) shows that Axiom 1 characterizes the class of variational preferences. Weak Order, Continuity, and Nondegeneracy are standard technical requirements. Weak Monotone Continuity guarantees that the probabilistic scenarios considered by the agent are countably additive. Weak Certainty Independence allows the agent to perceive some advantage in hedging, but this cannot come from mixing with different constants using the same weights. Monotonicity requires that the preference over acts is minimally consistent with the preference over the outcomes they induce. Uncertainty Aversion leads to aversion for the acts that perform well for a postulated model but poorly for its perturbations.

5.3.1 Structured Preferences

We are considering agents who face two levels of uncertainty: the uncertainty on the best structured description of the data-generating process and whether each description is exact. A representation is structured if it allows separating these two layers. In particular, to achieve this separation, we consider a state space S that admits the decomposition $S = \Omega \times \Delta(\Omega)$ for some finite Ω endowed with its Borel sigma-algebra.

Definition 10. An average robust control representation (u, Q, μ, λ) is *structured* if Q is finite and there exists $(\rho_q)_{q \in Q}$ such that for every $q \in Q$, and $\omega \in \Omega$, $\rho_q \in \Delta(\Omega)$ and $q(\{\omega, \rho_q\}) = \rho_q(\omega)$.

The interpretation of a structured representation is that the state space can be factored in two components, the realization of the single period consequence $\omega \in \Omega$ and a component $\rho \in \Delta(\Omega)$ that pins down the distribution over states each period. An event E is *structured* if $E = \Omega \times B$ for some $B \in \mathcal{B}(\Delta(\Omega))$. The sigma-algebra generated by the structured events is denoted as Σ_s.[28]

We say that an event $E \subseteq S$ satisfies the sure-thing principle if, for all $f, g, h, h' \in \mathcal{F}$ we have that $fEh \succsim gEh$ implies $fEh' \succsim gEh'$. We denote by Σ_{st} the set of events that satisfy the sure-thing principle.

Axiom 2 (Structured Savage). *(i) There is a finite set $E \subseteq S$ such that $S \setminus E$ is null. (ii) **P2.** $\Sigma_s \subseteq \Sigma_{st}$. (iii) **P4.** If $E, E' \in \Sigma_s$ and $x, y, w, z \in X$ are such that $x \succ y$ and $w \succ z$, then*

$$xEy \succ xE'y \Rightarrow wEz \succ wE'z.$$

Structured Savage requires that (i) the agent posits a finite number of models and (ii) guarantees that when evaluating acts that only depend on the identity of the structured model, the agent satisfies the Sure-Thing Principle.[29] It also (iii) guarantees that when an agent faces alternatives whose outcomes depend only on whether the DGP belongs to two sets of models, their choices consistently reveal the one deemed more likely.

Axiom 3 (Intramodel Sure-Thing Principle). *For every $f, g, h, h' \in \mathcal{F}$,*

$$fWh \succsim_\rho gWh \implies fWh' \succsim_\rho gWh' \qquad \forall W \subseteq \Omega, \forall \rho \in \Delta(\Omega).$$

Structured Savage's P2 and the Intramodel STP imply that bets *between* models and preference over acts *within* a model satisfy the STP. However, they admit violations of the STP for acts whose payoff depends on both the model's identity and the outcome realization within the model, as the ones of the original Ellsberg's paradox.

The case we study is when the relative likelihood of the structured models is only captured by the belief μ. In particular, the agent is equally concerned about how much each model departs from the actual data-generating process.

---

[28]With a slight abuse of notation for every $B \in \mathcal{B}(\Delta(\Omega))$ and $W \subseteq \Omega$ we denote as $\succsim_B$ and $\succsim_W$ the binary relations $\succsim_{\Omega \times B}$ and $\succsim_{W \times \Delta(\Omega)}$ and we write fBg and fWg for $f(\Omega \times B)g$ and $f(W \times \Delta(\Omega))g$.

[29]The extension to infinitely many models does not provide additional conceptual difficulties but makes the conditioning involved in the dynamic axioms much more cumbersome. Gul and Pesendorfer (2014) introduces the idea of sources of uncertainty for which the decision maker can quantify uncertainty and connects it with the Sure-Thing Principle.

Axiom 4 (Uniform Misspecification Concern). *For every $\rho, \rho' \in \Delta(\Omega)$ and $f, g \in \mathcal{F}$ such that*

$$\rho\left(\{\omega : f(\omega, \rho) = y\}\right) = \rho'\left(\{\omega : g(\omega, \rho') = y\}\right) \quad \forall y \in X$$

and $\Omega \times \{\rho\}$, $\Omega \times \{\rho'\}$ are nonnull we have

$$f \succsim_\rho x \Longleftrightarrow g \succsim_{\rho'} x \qquad \forall x \in X.$$

This axiom requires that if acts f and g induce identical outcome distributions under ρ and ρ', they are compared with a safe alternative in the same way conditional on the best fitting model being revealed to be ρ or ρ'.

Definition 11. The state space is adequate if: (i) there exist $k \in (0,1)$ and $(W_\rho)_{\rho \in \Delta(\Omega)} \in (2^\Omega)^{\Delta(\Omega)}$ such that for all $\rho \in \Delta(\Omega)$ with $\Omega \times \{\rho\}$ nonnull, $\rho(W_\rho) = k$, (ii) for every $\omega, \omega' \in \Omega$, and $\rho \in \Delta(\Omega)$ such that $\{\omega\} \times \{\rho\}$ and $\{\omega'\} \times \{\rho\}$ are nonnull, $\rho(\omega) = \rho(\omega')$.

All the agent's structured models have an event with the same probability and are uniform over a model-specific set of outcomes. It is well-known that equal probability requirements are essential for probabilistic sophistication with respect to a finite measure over states to have a bite (see, e.g., Chew and Sagi, 2006). They can be relaxed if we allow for a continuum Ω. The only role of (i) for us is to obtain a concern for misspecification λ that is not model dependent (i.e., not to have $(\lambda_q)_{q \in Q}$ in the representation) from Uniform Misspecification Concern, an axiom with an unmistakable flavor of probabilistic sophistication.

Axiom 5 (Uncertainty Neutrality Over Models). *Let $x, y, w, z \in X$, $\rho \in \Delta(\Omega)$, and $\gamma \in (0,1)$. Then $[\gamma x + (1 - \gamma) y]_\rho w \sim y_\rho z$ if and only if $x_\rho w \sim [(1 - \gamma) x + \gamma y]_\rho z$.*

Uncertainty Neutrality over Models guarantees that at the level of bets over models, the agent is "risk-neutral", as changing the performance under ρ by $(x - y)\gamma$ has an impact that does not depend on the level of utility under that model. It is immediate from the proof of Theorem 3 that if dropped, it leads to a more general representation with a nonlinear utility index U over the performance of each robust control model.

Theorem 3. *Suppose that S is adequate, there at least three disjoint nonnull events in Σ_s, and every nonnull $E \in \Sigma_s$ contains at least three disjoint nonnull events. The following are equivalent:*

1. *$\succsim$ admits a structured average robust control representation (u, Q, μ, λ);*

2. $\succsim$ satisfies Variational Axiom, Structured Savage, Uniform Misspecification Concern, Intramodel Sure-Thing Principle, and Uncertainty Neutrality over Models.

Moreover, in this case, every two structured average robust control representations share the same μ.

The theorem characterizes the representation (u, Q, μ, λ) with probabilistic uncertainty *about* the model (Structured Savage), probabilistic sophistication *given* a model (Intramodel Sure-Thing Principle), and *incomplete trust* in any model (Uncertainty Aversion).

Corollary 2. *Suppose that $\succsim$ admits a structured average robust control representation (u, Q, μ, λ). Then $\succsim$ is more misspecification concerned than the subjective utility preference with utility index u and belief $\int_Q q\,d\mu(q)$.*

5.4 Dynamic Axioms

We next provide axioms that characterize the dynamic adjustment of preferences in the face of information. In particular, we look at joint axioms on a collection of history-dependent binary relations $\left(\succsim^h\right)_{h \in \mathcal{H}}$ indexed by the realized history. Recall that the relevant set of length $t \in \mathbb{N}$ histories for structured preferences is Ω^t.

Axiom 6 (Constant Preference Invariance)**.** *For every $x, x' \in X$ and $h \in \mathcal{H}$,*

$$x \succsim^h x' \Leftrightarrow x \succsim^\emptyset x'.$$

This axiom captures the fact that we are not considering the problem of an agent discovering their taste. The preferences over uncertain alternatives are fixed and do not react to new information.

Axiom 7 (Dynamic Consistency over Models)**.** *Let $f, g \in \mathcal{F}$ be Σ_s-measurable, $t \in \mathbb{N}$, $\omega^t \in \Omega^t$ and $\bar{z}, \underline{z} \in X$ be such that $\bar{z} \succsim f(s) \succsim \underline{z}$ and $\bar{z} \succsim g(s) \succsim \underline{z}$ for all $s \in S$. Let*

$$h^{\omega^t}(\omega, \rho) = \gamma_{h(\omega, \rho)} \prod_{i=1}^{t} \rho(\omega_i)\, \bar{z} + \left(1 - \gamma_{h(\omega, \rho)} \prod_{i=1}^{t} \rho(\omega_i)\right) \underline{z} \qquad \forall (\omega, \rho) \in S, \forall h \in \{f, g\}$$

where $\gamma_{h(\omega, \rho)}$ satisfies $h(\omega, \rho) \sim \bar{z}\gamma_{h(\omega, \rho)} + \left(1 - \gamma_{h(\omega, \rho)}\right) \underline{z}$. Then, we have

$$f \succsim^{\omega^t} g \Longleftrightarrow f^{\omega^t} \succsim g^{\omega^t}.$$

27

The second dynamic axiom requires Bayesian rationality when considering acts whose consequences only depend on the structured model. Formally, it requires that when comparing acts that only bet on the identity of the model, at a given history, we can reduce the comparison to acts evaluated ex-ante. To do so, the payoff conditional to each model must be scaled proportionally to the amount of evidence that has been generated in favor of that model.[30]

To single out the *quantitative* speed at which the concern for misspecification is adjusted, we need a quantitative measure of similarity. For every $x, y \in X$ with $x \succ y$ and $E \in \Sigma$ let $\boldsymbol{\gamma}_{\succsim}^{xEy}$ be defined by

$$\boldsymbol{\gamma}_{\succsim}^{xEy} x + \left(1 - \boldsymbol{\gamma}_{\succsim}^{xEy}\right) y \sim xEy.$$

That is, $\boldsymbol{\gamma}_{\succsim}^{xEy}$ is the weight to alternative x in the certain equivalent to act xEy. It captures both the confidence in event E and the attitudes towards uncertainty. It is easy to see that under the Variational Axiom $\boldsymbol{\gamma}_{\succsim}^{xEy}$ always exists and is unique.

For every $x, y \in X$, $E \in \Sigma$, $\varepsilon \in (0, 1)$, and $\succsim$ and $\succsim'$ that satisfy the Variational Axiom, we say that $\succsim$ is (x, y, E, ε)-*similar* to $\succsim'$ if $\left| \boldsymbol{\gamma}_{\succsim}^{xEy} - \boldsymbol{\gamma}_{\succsim'}^{xEy} \right| \leq \varepsilon$. That is, the certain equivalent of the binary act xEy is ε close under preferences $\succsim$ and $\succsim'$.

Axiom 8 (Asymptotic Frequentism). *For every $\rho \in \Delta(\Omega)$, $x, y \in X$ with $x \succ^{\emptyset} y$, $\varepsilon > 0$, and $E \in \Sigma$ there is $\tau \in \mathbb{N}$ such that if $\min\{t, t'\} \geq \tau$ and $h_t, h_{t'}$ have outcome frequency ρ then $\succsim^{h_t}$ is (x, y, E, ε)-similar to $\succsim^{h_{t'}}$.*

The axiom requires that for every binary act xEy, a sufficiently long sequence of outcomes with the same empirical frequency stabilize the certain equivalent.

Proposition 4. *Suppose that: (i) For every $h \in \mathcal{H}$, $\succsim^h$ satisfies the axioms of Theorem 3 and (ii) $\left(\succsim^h\right)_{h \in \mathcal{H}}$ satisfies Constant Preference Invariance, Dynamic Consistency over Models, and Asymptotic Frequentism. Then for every $h \in \mathcal{H}$, $\succsim^h$ admits an average robust control representation $(u, Q, \mu(\cdot|h), \lambda_h)$ and for every sequence $(h_{t_n})_{n \in \mathbb{N}}$ with a constant*

---

[30]This axiom can lead to fruitful implications beyond our average robust control decision criterion, as it implies Bayesian updating for each decision criterion that performs an average of model-specific evaluations (that could, for example, take the form of other divergence preferences or rank-dependent utility evaluations). In this way, it would complement the elegant theory of subjective learning developed in Dillenberger et al. (2014), which does not require that the analyst observes the same information as the agent.

outcome frequency not in $\{\rho_q : q \in Q\}$,

$$\lim_{n \to \infty} \lambda_{h_{t_n}} \Big/ \left(\frac{LLR\left(h_{t_n}, Q\right)}{t_n} \right) \tag{8}$$

exists. Moreover, if for some $q \in Q$, $x \succ^\emptyset y$, and $E \subseteq \Omega$ with $\rho_q(E) > 0$

$$\liminf_{n \to \infty} \gamma_{\succsim h_{t_n}}^{x\left(E \times \{\rho_q\}\right)y} > 0,$$

the limit is finite, and if

$$\limsup_{n \to \infty} \gamma_{\succsim h_{t_n}}^{x\left(E \times \{\rho_q\}\right)y} < \rho_q(E)\, \mu(q),$$

it is strictly positive.

The proof has three main steps. First, we show that Constant Preference Invariance and Dynamic Consistency over Models imply that the preference for constant acts is time-invariant and that the belief over structured models is updated according to Bayesian updating. Second, we show that the likelihood ratio statistic of the models Q is growing linearly in t_n along the sequence of histories $(h_{t_n})_{n \in \mathbb{N}}$, so that the denominator in equation (8) converges.[31] Because the outcome frequency does not correspond to a model in Q, this limit is not 0. With this, the proof amounts to showing that the revealed concern for misspecification also converges. The third step rules out the existence of different finite limit points for $(\lambda_{t_n})_{n \in \mathbb{N}}$ by contradiction. If these points exist, then for every pair of strictly ranked outcomes $x \succ y$, we construct an event E for which the DM does not satisfy the Sure-Thing Principle such that the preference with the high concern for misspecification has a strictly lower certain equivalent than the one with the low concern.

Axiom 9 (Asymptotic Concern)**.** *Let $f \in \mathcal{F}$, $x \in X$, and $\hat\rho, \rho \in \Delta(\Omega)$ be such that $\Omega \times \{\hat\rho\}$ is $\succsim^\emptyset$-null, $\Omega \times \{\rho\}$ is $\succsim^\emptyset$-nonnull, and $\rho \gg \hat\rho$. If $\rho\left(\{\omega \in \Omega : x \succ f(\omega, \rho)\}\right) > 0$, then there exists $\tau \in \mathbb{N}$ such that for all $t \geq \tau$ and all h_t with outcome frequency $\hat\rho$, $x \succsim_\rho^{h_t} f$.*

Asymptotic Concern requires that long-run failures in explaining the data (i.e., an empirical frequency $\hat\rho$ that is not among the agent's structured models) increase the concern

---

[31] Given the finiteness of Ω, we can focus on the case in which the alternative set of models $N(Q) = \Delta(\Omega)$, discussed in Section 2, i.e., $LLR(h_t, Q) = -\log\left(\frac{\max_{q \in Q} \prod_{\tau=1}^{t} \rho_q(\omega_\tau)}{\max_{\rho \in \Delta(\Omega)} \prod_{\tau=1}^{t} \rho(\omega_\tau)} \right)$. The extension to general sets of alternative models is straightforward.

so that every certain outcome is preferable to an act with worse payoffs under a relevant model ρ.

Proposition 5. *Suppose that: (i) For every $h \in \mathcal{H}$, $\succsim^h$ satisfies the axioms of Theorem 3 and (ii) $\left(\succsim^h\right)_{h\in\mathcal{H}}$ satisfies Constant Preference Invariance, Dynamic Consistency over Models, and Asymptotic Concern. Then for every $h \in \mathcal{H}$, $\succsim^h$ admits an average robust control representation $(u, Q, \mu(\cdot|h), \lambda_h)$ and for every sequence $(h_{t_n})_{n\in\mathbb{N}}$ with a constant outcome frequency that is not in $\{\rho_q : q \in Q\}$ we have*

$$\lim_{t\to\infty} \frac{LLR(h_{t_n}, Q)}{\lambda_{h_{t_n}} t_n} = 0.$$

This result shows that Asymptotic Concern characterizes agents who apply an excessively demanding time discount to the log-likelihood ratio statistic (see equation (2)). Indeed, the elicited ratio between the LLR and the concern for misspecification revealed by the choices grows sublinearly over time, the condition that defines demanding agents.

Axiom 10 (Asymptotic Leniency). *Let $x, y \in X$, $E \in \Sigma$, $\rho \in \Delta(\Omega)$ and $\varepsilon > 0$, be such that $x \succ y$. For every Bayesian SEU preferences $\left(\succeq^h\right)_{h\in\mathcal{H}}$ such that $\succeq^\emptyset$ is less misspecification averse than $\succsim^\emptyset$, there exists $\tau \in \mathbb{N}$ such that for every $t \geq \tau$ and h_t with outcome frequency ρ, $\succsim^{h_t}$ is (x, y, E, ε)-similar to $\succeq^{h_t}$.*

Asymptotic Leniency requires that if the empirical distribution converges to some $\rho \in \Delta(\Omega)$, the preferences of the agents approximate, i.e., are eventually (x, y, E, ε)-similar to the updated preferences of an SEU whose model contingent preferences were initially less misspecification averse than the agent.

Proposition 6. *Suppose that: (i) For every $h \in \mathcal{H}$, $\succsim^h$ satisfies the axioms of Theorem 3 and (ii) $\left(\succsim^h\right)_{h\in\mathcal{H}}$ satisfies Constant Preference Invariance, Dynamic Consistency over Models, and Asymptotic Leniency. Then for every $h \in \mathcal{H}$, $\succsim^h$ admits an average robust control representation $(u, Q, \mu(\cdot|h), \lambda_h)$ and for every sequence $(h_{t_n})_{n\in\mathbb{N}}$ with constant outcome frequency not in $\{\rho_q : q \in Q\}$*

$$\lim_{t\to\infty} \frac{LLR(h_{t_n}, Q)}{t_n \lambda_{h_{t_n}}} = \infty.$$

This proposition shows that convergence to subjective expected utility maximization (in the form of Asymptotic Leniency) characterizes excessively lenient model evaluations.

6 Discussion

6.1 Related Literature

A few papers allow the agents to realize that they are misspecified. In particular, in He and Libgober (2022), Ba (2022), Fudenberg and Lanzani (2023), Gagnon-Bartsch et al. (2022), and Cho and Libgober (2022) misspecification can be eliminated by "light bulb realizations" evolutionary pressure, or the use of a machine-learning inspired algorithm. The key difference with our approach is that in these papers, as well in the earlier Cho and Kasa (2015) and Giacomini et al. (2020), where agents switch between models on the basis of a specification test, the agents act as if they have complete trust in the set of models currently entertained and are never concerned about being misspecified. Still, there is a tight connection between the robust control decision criterion and a maxmin decision criterion where the set of models expands as the penalization term in the robust control increases (see Hansen and Sargent, 2011, for a textbook treatment). In light of this, compared to the previous set of papers, our work can additionally be interpreted as providing the first *smooth* framework for expanding (or restricting) the set of considered models as a function of the evidence.

Although not dealing with misspecification, Epstein and Schneider (2007) also studies a learning problem with nonSEU preferences and where the likelihood ratio between different DGPs plays an important role. In the myopic case, and under the notation of our paper, they consider a maxmin decision criterion

$$\min_{\mu \in C_t} \int_\Theta \mathbb{E}_{q_a^\theta} \left[u\left(a, y\right) \right] \mathrm{d}\mu\left(\theta | h_t \right)$$

where

$$C_t = \left\{ \nu \in C_0 : \frac{\int_\Theta \prod_{i=1}^t q_{a_t}^\theta\left(y_t\right) \mathrm{d}\mu\left(\theta\right)}{\max_{\mu \in C_0} \int_\Theta \prod_{i=1}^t q_{a_t}^\theta\left(y_t\right) \mathrm{d}\mu\left(\theta\right)} \geq \beta \right\},$$

$\beta \in (0,1)$ and $C_0 \subseteq \Delta\left(\Theta\right)$. That is, the look at maxmin preferences with a set of priors that only contains the priors that did not perform excessively worse than the best rationalizing one. Beyond considering maxmin instead of robust control preferences, the papers have a completely different set of results, as Epstein and Schneider (2007) only deals with the question of beliefs stabilization. Crucially, by always considering models that are only supported on Θ, Epstein and Schneider (2007) cannot capture the concern for misspecification that is at the core of the motivation for this paper, but it rather

31

captures a gradually increasing confidence in having understood which of the priors in Θ is the most accurate. Importantly, they do not restrict the agent to be myopic.

Farther afield, Ortoleva (2012) proposes and axiomatizes a model where a decision maker can reject their model in favor of a backup one when faced with events with sufficiently low probability. Karni and Vierø (2013) proposes and axiomatizes a model where the agent becomes progressively aware of more states and acts. However, their decision maker trusts their probability over states completely when making decisions. Banerjee et al. (2020) studies a Wald problem where the agent trade-offs between robustness and the subjective expected utility performance of the experiment. Differently from us, the concern in this model does not evolve, and the agent makes a single decision. Epstein and Ji (2022) characterizes optimal stopping with a concern for robustness captured by maxmin preferences, showing that the robustness concern, in general, induces earlier stopping.

There is fast-growing literature on learning under misspecification with subjective expected utility preferences. Arrow and Green (1973) gives the first general framework for this problem, and Nyarko (1991) points out that the combination of misspecification and endogenous data can lead to cycles. This literature has been revived by the more recent Esponda and Pouzo (2016); see Bohren and Hauser (2021), Esponda et al. (2021a), Fudenberg et al. (2021), and Frick et al. (2023) for analyses of more closely related settings.

The identification of an agent who is disappointed with minor discrepancies between the empirical and the theoretical distributions as a believer in the Law of Small Numbers follows the formalization of this bias proposed by Rabin (2002). The normative role of the likelihood ratio that makes it proportional to the relative entropy (cf. Proposition 1 and Theorem 1) is somewhat reminiscent of the normative role of (absolute) entropy as a measure of informativeness found by Cabrales et al. (2013).

Hansen and Sargent (2007) mentions a time-varying penalization parameter as a way to maintain dynamic consistency in the robust control model. Maenhout (2004) also uses a time-varying penalization parameter in a portfolio selection problem to keep the recursive discounted preferences homothetic at any history. See Pathak (2002) for a critical perspective on the latter paper and the subsequent literature. In both cases, the parameter evolution does not capture the fit of the models to the observed data. Anderson et al. (2003) and Barillas et al. (2009) pioneer a literature that calibrates the (time-invariant) concern for misspecification from the acceptable error probability in likelihood ratio test between the unperturbed model and the worst-case model (that does not depend on the action there). See Hansen and Sargent (2011) for a textbook treatment.

In the evolutionary game theory literature studying preferences formation, the closest papers are Dekel et al. (2007) and Robatto and Szentes (2017). In the former, players may be "misspecified" in the sense of having a vN-M utility different from the one determining reproductive fit. In the latter, the evolutionary pressure determines the risk attitudes of the players.

On the axiomatic side, the static decision criterion considered here is due to Cerreia-Vioglio et al. (2022).[32] The explicit use of a state space where every state describes both the single-period outcome realization and the probability distribution over outcomes follows the approach introduced in Cerreia-Vioglio et al. (2013a) as a two-stage "statistical" interpretation and axiomatization of some of the decision criteria under ambiguity, in particular the smooth ambiguity one. For this criterion, this approach has been recently extended by Denti and Pomatto (2022). They allow for a fully revealed-preference elicitation of the relevant probability distributions, viewed as subjective statistical models. See also Dean and Ortoleva (2017) for a less related decision criterion where the agent has a prior over multiple data-generating processes and evaluates each of them with rank-dependent utility and Gilboa et al. (2020) for a different quasi-Bayesian criterion that combines Bayesian updating with case-reasoning rather than misspecification considerations.

6.2 Experimental Evidence

We are unaware of experiments that explicitly test the positive relation between misspecification and the belief in the Law of Small Numbers with uncertainty aversion. However, the findings in Esponda et al. (2022) suggest that a mechanism similar to the one outlined in this paper is actually at play. The paper studies the repeated behavior of two groups of agents, one with an agnostic (full support) belief about the possible data-generating process faced and one that is misspecified because of base rate neglect. The long-run average play of the misspecified agents is in between the best reply to the misspecified model and the uniform distribution over outcomes. Notably, this behavior is not the best reply to the observed empirical frequency, which suggests that, as in our model, even in the medium run (200 repetitions in the experiment above), the agents do not altogether drop their models; they rely less on them to make their choices. Instead, the correctly specified agents converge to making choices that are optimal only under the actual data-generating

---

[32]Given the Donsker–Varadhan variational formula, our decision criterion can also be seen as the average of CARA certain equivalents, an object studied and characterized from a statistical perspective in Mu et al. (2021).

process, i.e., they behave as a subjective expected utility maximizer with a belief concentrated on the true DGP. More indirectly De Filippis et al. (2022) shows that there is an overreaction of beliefs to consistent signal than to inconsistent signal, suggesting that agents who find a model consistently validated may rely more on it to make decisions. Another important avenue to bring the literature on misspecification is to try to bridge the decision criterion proposed here and in Cerreia-Vioglio et al. (2022) with the distinction between uncertainty about and within the experimental design advocated in Shmaya and Yariv (2016).

6.3 Forward-looking Agents

One key generalization to our model would be to allow for forward-looking agents. Of course, as for many decision criteria that depart from SEU, the main complication is dealing with the fact that the most immediate extension of the criterion to forward-looking agents would induce dynamic inconsistencies under some information structures (see Appendix A.3.3 for a simple explicit example). One approach would be to directly impose a recursive formulation for the preferences, as in Maccheroni et al. (2006b) and Klibanoff et al. (2009). Since the decision criterion belongs to the variational class, we know from the first reference that a recursive formulation can be obtained. A complementary approach does not impose recursivity and allows for dynamic inconsistency. Preliminary analysis suggests that if we consider agents who do not anticipate their future taste variations, little is changed. However, analyzing an uncommitted, forward-looking, and sophisticated agent playing an intra-personal equilibrium with their future selves would require combining the insights of this paper with the approach developed in Battigalli et al. (2019). Analogously, to extend the axiomatic exercise to sophisticated agents, the techniques of this paper should be combined with the consistent planning approach of Siniscalchi (2011).

6.4 Endogenous Structured Models

The more natural extension for the decision-theoretic part of the paper involves using axioms that do not explicitly allow the agent to bet on the identity of the structured model. Allowing such bets is relatively standard when dealing with two levels of uncertainty for which the agent has different attitudes (Klibanoff et al., 2005 being the most prominent example). However, Denti and Pomatto (2022) proposed an identifiability condition that avoids the need for explicit bets on the structured models. Identifiability requires a way

to partition S that singles out the probabilistic model. In particular, each probabilistic model assigns probability one to its corresponding partition element.

When considering a structured environment, we required that this identification is spelled out in the description of the states, with the second component being the distribution ρ. Although in light of the results of Denti and Pomatto (2022), the axiomatization of this static criterion without this restriction does not generate conceptual complications, the dynamic characterization becomes significantly more involved. In particular, the challenge is created by the conditioning with respect to the endogenously identified model. This substantial extension is left for future work.

7 Conclusion

In this work, we propose a novel model of agents actively learning about the environment and dynamically adjusting their concern for misspecification on the basis of the evidence they face. We show that the agents develop different long-run uncertainty attitudes depending on their understanding of how quickly evidence in favor or against a model is accumulated. Statistically sophisticated agents converge to robust control preferences a la Hansen and Sargent (2001), with the misspecification concern endogenously determined by their models fit with the true DGP at the equilibrium action. In contrast, an agent who is too demanding in evaluating their model converges to behave as a maxmin agent a la Gilboa and Schmeidler (1989), while a lenient agent eventually becomes a standard subjective expected utility maximizer. These results provide the first learning foundation for nonstandard decision criteria.

We then point out that in natural environments, the behavior of the statistically sophisticated type need not converge, and we characterize the limit frequency of time spent playing each action. We apply this result to a simple macroeconomic model and obtain a new rationale for the periodic switches in monetary policies.

We also provide an axiomatization of the proposed decision criterion and its evolution in the face of evidence. We introduce a new axiom type, Asymptotic Frequentism, requiring long streams of outcomes with the same empirical frequency to induce similar preferences. We prove that this axiom induces the statistically sophisticated behavior studied in the learning part of the paper.

A   Appendix

A.1   Learning Results

A.1.1   Preliminaries

By Assumption 1, there exists $K \in \mathbb{R}_{++}$ such that for all $a \in A$ and $q \in Q$

$$-\ln \tilde{q}_a \leq K \qquad \forall q \in Q$$

holds p_a^*-a.s. To see this, observe that the assumption directly implies that there exists $K \in \mathbb{R}_{++}$ such that for all $a \in A$ and $q \in Q$,

$$-\ln \tilde{q}_a \leq 2K, \qquad p_a^*\text{-a.s.} \tag{9}$$

Suppose by contradiction that there exists a measurable $C \subseteq \operatorname{supp} p_a^*$ such that $\max_{q \in Q} -\ln \tilde{q}_a(y) > K$ for all $y \in C$. Take $y^* \in C$ and $q' \in Q$ such that $-\ln \tilde{q}'_a(y^*) > 2K$. By the continuity of q', there exists $\varepsilon > 0$ such that $-\ln \tilde{q}'_a(y) > K$ for every $y \in B_\varepsilon(y^*)$. But since $y* \in C \subseteq \operatorname{supp} p_a^*$, $p_a^*(B_\varepsilon(y^*)) > 0$, a contradiction with equation (9).

Throughout Appendix A.1, the symbol K will denote such a strictly positive real number. For an arbitrary Borel measurable subset C of a metric space, we endow the space $C^{\mathbb{N}}$ with the Borel σ-algebra, $\mathcal{B}\left(C^{\mathbb{N}}\right)$, corresponding to the product topology on $C^{\mathbb{N}}$. For $(k_1, ..., k_t) \in C^t$, $t \in \mathbb{N}$, we denote by $k^t = (k_1, ..., k_t)$ both the finite sequence in C^t and the elementary cylinder in $C^{\mathbb{N}}$ that it identifies. For every policy $\Pi \in A^{\mathcal{H}}$, the density of the objective probability distribution over infinite histories is defined over a finite number of periods $I \subseteq \mathbb{N}$ with $t_I = \max I$ as

$$\tilde{\mathbb{P}}_\Pi\left((a_\tau, y_\tau)_{\tau \in I}\right) = \begin{cases} 1 & \text{if } \exists \, (\hat{a}_\tau, \hat{y}_\tau)_{\tau=1}^{t_I} \in (A \times Y)^{t_I} : \\ & \quad \hat{a}_{\tau+1} = \Pi\left(\hat{a}^\tau, \hat{y}^\tau\right), \forall \tau \in \{0, ..., t_I - 1\} \\ & \quad \text{and } (\hat{a}_\tau, \hat{y}_\tau) = (a_\tau, y_\tau), \forall \tau \in I, \\ 0 & \text{otherwise,} \end{cases} \tag{10}$$

and

$$\mathbb{P}_\Pi\left((a_\tau)_{\tau \in I}, C\right) = \int_C \tilde{\mathbb{P}}_\Pi\left((a_\tau, \cdot)_{\tau \in I}\right) d\left(\prod_{\tau \in I} p_{a_\tau}^*\right) \qquad \forall C \in \mathcal{B}\left(Y^I\right).$$

Since the corresponding set of finite-dimensional probability measures is consistent, there is a unique probability measure over infinite sequences of action-outcome pairs with these

marginals, defined through the Kolmogorov extension theorem (see Theorem V.5.1 in Parthasarathy, 2005 for the version for standard separable Borel spaces used here).

For every $t \in \mathbb{N}$ and history $h_t = (a^t, y^t) \in \mathcal{H}$ let $p^{h_t} \in \Delta(Y)^A$ be the action contingent (finite support) probability measure over outcomes corresponding to the empirical frequency: for all $a \in A$ such that $\sum_{\tau=1}^t \mathbb{I}_{\{a\}}(a_\tau) > 0$,

$$p_a^{h_t}(C) = \frac{\sum_{\tau=1}^t \mathbb{I}_{\{(a,y):y\in C\}}(a_\tau, y_\tau)}{\sum_{\tau=1}^t \mathbb{I}_{\{a\}}(a_\tau)} \qquad \forall C \subseteq \mathcal{B}(Y)$$

and $p_a^{h_t} = \delta_{\bar{y}}$ for some arbitrary fixed $\bar{y} \in Y$ if $\sum_{\tau=1}^t \mathbb{I}_{\{a\}}(a_\tau) = 0$. For every two histories $h_t, h_\tau \in \mathcal{H}$ we write $h_t \succ h_\tau$ if there is $n \in \mathbb{N}$ and $(a_i, y_i)_{i=1}^n$ such that $h_t = (h_\tau, (a_i, y_i)_{i=1}^n)$. For all $b \in A$ let Π^b the policy that prescribes b at every period. Define the set $Q^\varepsilon(a)$ as the models at most ε away from a relative entropy minimizer given action a, $Q^\varepsilon(a) = \{q \in Q : \exists q' \in Q(a) \cap B_\varepsilon(q)\}$.

A.1.2   Results

Our first lemma justifies the repeated use of min and max rather than inf and sup in the definitions of the average robust criterion and LLR.

Lemma 1.   1. For every $a \in A$, $\lambda \in \mathbb{R}_{++}$, and $q \in Q$, $\emptyset \neq \operatorname{argmin}_{p_a \in \Delta(Y)} \left(\mathbb{E}_{p_a}[u(a,y)] + \frac{R(p_a||q_a)}{\lambda} \right)$.

2. For every $t \in \mathbb{N}$, $h_t \in \mathcal{H}_t$, $\emptyset \neq \operatorname{argmax}_{p \in N(Q)} \prod_{\tau=1}^t \tilde{p}_{a_\tau}(y_\tau)$.

Proof. 1) Fix $a \in A$, $\lambda \in \mathbb{R}_{++}$, and $q \in Q$. Since u is continuous and Y is compact, $u(a, \cdot)$ is bounded, $\mathbb{E}_{p_a}[u(a,y)] \geq \min_{y \in Y} u(a,y) \in \mathbb{R}$ for all $p_a \in \Delta(Y)$, and $p_a \mapsto \mathbb{E}_{p_a}[u(a,y)]$ is continuous. Since Y is a compact metric space, by, e.g., Royden and Fitzpatrick (1988), it is a Polish space and so $p_a \mapsto R(p_a||q_a)$ is lower semicontinuous by Lemma 1.4.3 in Dupuis and Ellis (2011). Therefore, the set

$$E := \left\{ p_a \in \Delta(Y) : \mathbb{E}_{p_a}[u(a,y)] + \frac{1}{\lambda} R(p_a||q_a) \leq \mathbb{E}_{q_a}[u(a,y)] \right\}$$

is closed, and as $R(q_a||q_a) = 0$ we clearly have

$$\operatorname{argmin}_{p_a \in \Delta(Y)} \mathbb{E}_{p_a}[u(a,y)] + \frac{1}{\lambda} R(p_a||q_a) = \operatorname{argmin}_{p_a \in E} \mathbb{E}_{p_a}[u(a,y)] + \frac{1}{\lambda} R(p_a||q_a).$$

Since Y is a compact metric space, by Theorem 15.11 in Aliprantis and Border (2013) and Proposition 11.15 in Royden and Fitzpatrick (1988) so are $\Delta(Y)$ and E endowed with the

37

topology of weak convergence of measures. Since

$$p_a \mapsto \mathbb{E}_{p_a}\left[u\left(a, y\right)\right] + \frac{1}{\lambda} R\left(p_a || q_a\right)$$

is real-valued (with values in $[\min_{y \in Y} u\left(a, y\right), \mathbb{E}_{q_a}\left[u\left(a, y\right)\right]]$) and lower semicontinuous on the compact E, it admits a minimizer by the generalized Weierstrass' theorem (see, e.g., Theorem 2.43 in Aliprantis and Border, 2013).

2) Let $t \in \mathbb{N}$, $h_t \in \mathcal{H}_t$. The map

$$
\begin{array}{ccc}
\{\tilde{p} : p \in N(Q)\} & \to & \mathbb{R} \\
\tilde{p} & \mapsto & \prod_{\tau=1}^{t} \tilde{p}_{a_\tau}(y_\tau)
\end{array}
$$

is continuous. Then, the maximum is attained since $\{\tilde{p} : p \in N(Q)\}$ is compact by Assumption 2 (ii). ∎

The next lemma provides a useful rewriting of the LLR as a weighted average of the empirical log-likelihood ratio when playing the different actions, with weights proportional to how frequently each action has been used in the past.

Lemma 2. *For every $t \in \mathbb{N}$ and $h_t = (a^t, y^t) \in \mathcal{H}_t$, if $q' \in \mathrm{argmax}_{q \in Q} \prod_{\tau=1}^{t} \tilde{q}_{a_\tau}(y_\tau)$, and $p \in \mathrm{argmax}_{r \in N(Q)} \prod_{\tau=1}^{t} \tilde{r}_{a_\tau}(y_\tau)$ then $LLR\left(h_t, Q\right) = \sum_{a \in A} \sum_{\tau=1}^{t} \mathbb{I}_{\{a\}}\left(a_\tau\right) \int_Y \log\left(\frac{dp_a(y)}{dq'_a(y)}\right) dp_a^{h_t}(y)$.*

Proof. We have

$$LLR\left(h_t, Q\right) = -\log\left(\frac{\max_{q \in Q} \prod_{\tau=1}^{t} \tilde{q}_{a_\tau}(y_\tau)}{\max_{r \in N(Q)} \prod_{\tau=1}^{t} \tilde{r}_{a_\tau}(y_\tau)}\right) = \log\left(\frac{\max_{r \in N(Q)} \prod_{\tau=1}^{t} \tilde{r}_{a_\tau}(y_\tau)}{\max_{q \in Q} \prod_{\tau=1}^{t} \tilde{q}_{a_\tau}(y_\tau)}\right)$$

$$= \log\left(\frac{\prod_{\tau=1}^{t} \tilde{p}_{a_\tau}(y_\tau)}{\prod_{\tau=1}^{t} \tilde{q}'_{a_\tau}(y_\tau)}\right) = \log\left(\prod_{\tau=1}^{t} \tilde{p}_{a_\tau}(y_\tau)\right) - \log\left(\prod_{\tau=1}^{t} \tilde{q}'_{a_\tau}(y_\tau)\right)$$

$$= \log\left(\prod_{y \in Y} \prod_{a \in A} \tilde{p}_a(y)^{\sum_{\tau=1}^{t} \mathbb{I}_{\{a\}}(a_\tau) p_a^{h_t}(\{y\})}\right) - \log\left(\prod_{y \in Y} \prod_{a \in A} \tilde{q}'_a(y)^{\sum_{\tau=1}^{t} \mathbb{I}_{\{a\}}(a_\tau) p_a^{h_t}(\{y\})}\right)$$

$$= \sum_{a \in A} \sum_{\tau=1}^{t} \mathbb{I}_{\{a\}}\left(a_\tau\right) \sum_{y \in Y} p_a^{h_t}\left(\{y\}\right) \log\left(\tilde{p}_a\left(y\right)\right) - \sum_{a \in A} \sum_{\tau=1}^{t} \mathbb{I}_{\{a\}}\left(a_\tau\right) \sum_{y \in Y} p_a^{h_t}\left(\{y\}\right) \log\left(\tilde{q}'_a\left(y\right)\right)$$

$$= \sum_{a \in A} \sum_{\tau=1}^{t} \mathbb{I}_{\{a\}}\left(a_\tau\right) \sum_{y \in Y} p_a^{h_t}\left(\{y\}\right) \left(\log\left(\tilde{p}_a\left(y\right)\right) - \log\left(\tilde{q}'_a\left(y\right)\right)\right)$$

$$= \sum_{a \in A} \sum_{\tau=1}^{t} \mathbb{I}_{\{a\}}\left(a_\tau\right) \int_Y \log\left(\frac{\tilde{p}_a\left(y\right)}{\tilde{q}'_a\left(y\right)}\right) dp_a^{h_t}\left(y\right).$$

$\blacksquare$

The next lemma shows that a robust control evaluation with respect to a structured model $q \in Q$ converges to a subjective expected utility evaluation as λ tends to 0, generalizing previous results in the decision-theoretic literature, where the function evaluated was a finite range one, to continuous utility functions.

Lemma 3. *For every $a \in A$, $q \in Q$, and $(q_n, \lambda_n)_{n \in \mathbb{N}} \in (Q \times \mathbb{R}_{++})^{\mathbb{N}}$ with $\lim_{n \to \infty} (q_n, \lambda_n) = (q, 0)$ we have*

$$\lim_{n \to \infty} \min_{p_a \in \Delta(Y)} \left(\mathbb{E}_{p_a} \left[u(a, y) \right] + \frac{R(p_a || q_{a,n})}{\lambda_n} \right) = \mathbb{E}_{q_a} \left[u(a, y) \right].$$

Proof. Fix $a \in A$ and define $\bar{u} = \max_{y \in Y} u(a, y) - \min_{y \in Y} u(a, y)$. For every $n \in \mathbb{N}$,

$$\min_{p_a \in \Delta(Y)} \left(\mathbb{E}_{p_a} \left[u(a, y) \right] + \frac{R(p_a || q_{a,n})}{\lambda_n} \right) \in \left[\min_{y \in Y} u(a, y), \mathbb{E}_{q_{a,n}} \left[u(a, y) \right] \right] \subseteq \left[\min_{y \in Y} u(a, y), \max_{y \in Y} u(a, y) \right],$$

so possibly restricting to a subsequence, we can assume that the limit in the LHS of the statement is well defined. The statement is then proved by showing that any such subsequence converges to the RHS. In particular, we show that we cannot have

$$\lim_{n \to \infty} \min_{p_a \in \Delta(Y)} \left(\mathbb{E}_{p_a} \left[u(a, y) \right] + \frac{R(p_a || q_{a,n})}{\lambda_n} \right) < \mathbb{E}_{q_a} \left[u(a, y) \right]. \tag{11}$$

This is sufficient as $\lim_{n \to \infty} \mathbb{E}_{q_{a,n}} \left[u(a, y) \right] = \mathbb{E}_{q_a} \left[u(a, y) \right]$ and therefore we know by the lower semicontinuity of R (see Lemma 1.4.3 in Dupuis and Ellis (2011)) that

$$\lim_{n \to \infty} \min_{p_a \in \Delta(Y)} \left(\mathbb{E}_{p_a} \left[u(a, y) \right] + \frac{R(p_a || q_{a,n})}{\lambda_n} \right) \leq \mathbb{E}_{q_a} \left[u(a, y) \right].$$

If equation (11) held, there would be an $\varepsilon \in \mathbb{R}_{++}$ with

$$\lim_{n \to \infty} \min_{p_a \in \Delta(Y)} \left(\mathbb{E}_{p_a} \left[u(a, y) \right] + \frac{R(p_a || q_{a,n})}{\lambda_n} \right) = \mathbb{E}_{q_a} \left[u(a, y) \right] - \varepsilon. \tag{12}$$

For every $n \in \mathbb{N}$, let $p_a^n \in \Delta(Y)$ be an arbitrary element of $\operatorname{argmin}_{p_a \in \Delta(Y)} \left(\mathbb{E}_{p_a} \left[u(a, y) \right] + \frac{R(p_a || q_{a,n})}{\lambda_n} \right)$. Since Y is a compact metric space, by Theorem 15.11 in Aliprantis and Border (2013) so is $\Delta(Y)$, and therefore, we can assume (by restricting to a subsequence) that p_a^n converges

to some $\hat{p}_a \in \Delta(Y)$. By equation (12) and the fact that $\lim_{n\to\infty} p_a^n = \hat{p}_a$, we have

$$\mathbb{E}_{\hat{p}_a}[u(a,y)] \leq \mathbb{E}_{q_a}[u(a,y)] - \varepsilon.$$

Therefore,

$$\int_0^{\bar{u}} 1 - \hat{p}_a\left(\left\{y \in Y : u(a,y) - \min_{\bar{y} \in Y} u(a,\bar{y}) \leq x\right\}\right) dx + \frac{3}{4}\varepsilon = \mathbb{E}_{\hat{p}_a}[u(a,y)] + \frac{3}{4}\varepsilon - \min_{\bar{y} \in Y} u(a,\bar{y})$$

$$\leq \mathbb{E}_{q_a}[u(a,y)] - \min_{\bar{y} \in Y} u(a,\bar{y}) = \int_0^{\bar{u}} 1 - q_a\left(\left\{y \in Y : u(a,y) - \min_{\bar{y} \in Y} u(a,\bar{y}) \leq x\right\}\right) dx. \qquad (13)$$

Claim 1. *There exist $M \in \mathbb{R}$ and $L \in \mathbb{R}_{++}$ such that*

$$\hat{p}_a\left(\{y \in Y : u(a,y) \leq M - L\}\right) - q_a\left(\{y \in Y : u(a,y) \leq M\}\right) \geq \frac{\varepsilon}{2\bar{u}}. \qquad (14)$$

Proof of the Claim. Suppose that for every $M \in \mathbb{R}$ and $L \in \mathbb{R}_{++}$ equation (14) does not hold. Then for every $L \in \mathbb{R}_{++}$

$$\int_0^{\bar{u}} 1 - \hat{p}_a\left(\left\{y \in Y : u(a,y) - \min_{\bar{y} \in Y} u(a,\bar{y}) \leq x\right\}\right) dx$$

$$= \int_0^{\bar{u}+L} 1 - \hat{p}_a\left(\left\{y \in Y : u(a,y) - \min_{\bar{y} \in Y} u(a,\bar{y}) + L \leq x\right\}\right) dx - L$$

$$\geq \int_0^{\bar{u}+L} 1 - q_a\left(\left\{y \in Y : u(a,y) - \min_{\bar{y} \in Y} u(a,\bar{y}) \leq x\right\}\right) - \frac{\varepsilon}{2\bar{u}}dx - L$$

$$= \int_0^{\bar{u}} 1 - q_a\left(\left\{y \in Y : u(a,y) - \min_{\bar{y} \in Y} u(a,\bar{y}) \leq x\right\}\right) dx - \varepsilon/2 - L\frac{\varepsilon}{2\bar{u}} - L.$$

Since L can be chosen to be arbitrarily small, we have

$$\int_0^{\bar{u}} 1 - \hat{p}_a\left(\left\{y \in Y : u(a,y) - \min_{\bar{y} \in Y} u(a,\bar{y}) \leq x\right\}\right) dx \geq \int_0^{\bar{u}} 1 - q_a\left(\left\{y \in Y : u(a,y) - \min_{\bar{y} \in Y} u(a,\bar{y}) \leq x\right\}\right)$$

a contradiction with equation (13). $\qquad \square$

The claim, in turn, implies that there exists $N \in \mathbb{N}$ such that for all $n \geq N$

$$p_a^n\left(\left\{y \in Y : u(a,y) \leq M - \frac{L}{2}\right\}\right) - q_{a,n}\left(\left\{y \in Y : u(a,y) \leq M - \frac{L}{2}\right\}\right) \geq \frac{\varepsilon}{4\bar{u}}.$$

But then

$$\min_{p_a \in \Delta(Y)} \mathbb{E}_{p_a}\left[u\left(a,y\right)\right] + \frac{1}{\lambda_n} R\left(p_a || q_{a,n}\right)$$

$$= \mathbb{E}_{p_a^n}\left[u\left(a,y\right)\right] + \frac{1}{\lambda_n} R\left(p_a^n || q_{a,n}\right)$$

$$\geq \min_{y \in Y} u\left(a,y\right) + \left(p_a^n\left(\left\{y \in Y : u\left(a,y\right) \leq M - \frac{L}{2}\right\}\right) \log \frac{p_a^n\left(\left\{y \in Y : u\left(a,y\right) \leq M - \frac{L}{2}\right\}\right)}{q_{a,n}\left(\left\{y \in Y : u\left(a,y\right) \leq M - \frac{L}{2}\right\}\right)}\right) / \lambda_n$$

$$+ \left(p_a^n\left(\left\{y \in Y : u\left(a,y\right) > M - \frac{L}{2}\right\}\right) \log \frac{p_a^n\left(\left\{y \in Y : u\left(a,y\right) > M - \frac{L}{2}\right\}\right)}{q_{a,n}\left(\left\{y \in Y : u\left(a,y\right) > M - \frac{L}{2}\right\}\right)}\right) / \lambda_n$$

where the inequality follows from Theorem 1.24 in Liese and Vajda (1987). But, the last term diverges to $+\infty$ as n goes to infinity, a contradiction with $\min_{p_a \in \Delta(Y)} \mathbb{E}_{p_a}\left[u\left(a,y\right)\right] + \frac{R(p_a || q_{a,n})}{\lambda_n} \leq \max_{y \in Y} u\left(a,y\right) < \infty$. ∎

Lemma 4. *1. For every $a \in A$, the function $G : \Delta(Q) \times \mathbb{R}_+ \to \mathbb{R}$ defined by*

$$G\left(\nu, \lambda\right) : = \int_Q \min_{p_a \in \Delta(Y)} \left(\mathbb{E}_{p_a}\left[u\left(a,y\right)\right] + \frac{R\left(p_a || q_a\right)}{\lambda}\right) d\nu\left(q\right) \quad \forall \nu \in \Delta(Q), \forall \lambda \in \mathbb{R}_{++}$$

and

$$G\left(\nu, 0\right) : = \int_Q \mathbb{E}_{q_a}\left[u\left(a,y\right)\right] d\nu\left(q\right) \quad \forall \nu \in \Delta(Q)$$

is continuous.

2. The correspondence $BR\left(\cdot\right) : \mathbb{R}_+ \times \Delta(Q) \rightrightarrows A$ where

$$BR^0\left(\nu\right) : = BR^{Seu}\left(\nu\right) \quad \forall \nu \in \Delta(Q)$$

is upper hemicontinuous.

Proof. (1) Fix $a \in A$. For every $q \in Q$, let $F\left(q,0\right) := \mathbb{E}_{q_a}\left[u\left(a,y\right)\right]$ and observe that for each $\lambda \in \mathbb{R}_{++}$, by Proposition 1.4.2 in Dupuis and Ellis (2011) we have

$$F\left(q, \lambda\right) : = \min_{p_a \in \Delta(Y)} \left(\mathbb{E}_{p_a}\left[u\left(a,y\right)\right] + \frac{R\left(p_a || q_a\right)}{\lambda}\right) = \frac{-\log\left(\int_Y \exp\left(-\lambda u\left(a,y\right)\right) dq_a\left(y\right)\right)}{\lambda}.$$

Since Y is compact and u is continuous, for all $\lambda \in \mathbb{R}_{++}$ and $q \in Q$, the RHS belongs to $[\min_{y \in Y} u\left(a,y\right), \max_{y \in Y} u\left(a,y\right)]$. For every $\lambda \in \mathbb{R}_{++}$, $\exp\left(-\lambda u\left(a,\cdot\right)\right)$ is a continuous and

41

bounded function that is bounded away from 0. Recall that weak and vague convergence for probability measures are equivalent when the state space is compact, and that Y is indeed compact. Therefore,

$$q \mapsto \int_Y \exp\left(-\lambda u\left(a, y\right)\right) \mathrm{d}q_a\left(y\right)$$

is continuous by definition of the weak convergence of measures, and F is continuous by Lemma 3 (at $\lambda = 0$) and Theorem 15.7.3 in Kallenberg (1973) (at $\lambda \neq 0$).

Let $(\nu_n, \lambda_n)_{n \in \mathbb{N}} \in \Delta(Q) \times \mathbb{R}_{++}$ be a convergent sequence with limit (ν, λ). Suppose first that $\lambda > 0$. Then

$$\lim_{n \to \infty} \int_Q \frac{\log\left(\int_Y \exp\left(-\lambda_n u\left(a, y\right)\right) \mathrm{d}q_a\left(y\right)\right)}{-\lambda_n} \mathrm{d}\nu_n\left(q\right) = \int_Q \frac{\log\left(\int_Y \exp\left(-\lambda u\left(a, y\right)\right) \mathrm{d}q_a\left(y\right)\right)}{-\lambda} \mathrm{d}\nu\left(q\right)$$

by Theorem 15.7.3 in Kallenberg (1973) and the joint continuity of F established above. Next, suppose that $\lambda = 0$. Then $\lim_{n \to \infty} \int_Q \frac{\log(\mathbb{E}_{qa}[\exp(-\lambda_n u(a,y))])}{-\lambda_n} \mathrm{d}\nu_n\left(q\right) = \int_Q \mathbb{E}_{q_a}\left[u\left(a, y\right)\right] \mathrm{d}\nu\left(q\right)$ again by Theorem 15.7.3 in Kallenberg (1973) and the joint continuity of F established above. This proves (i).

(2) Follows by (1) and Theorem 17.31 in Aliprantis and Border (2013). ∎

Lemma 5. *For every $a \in A$, if $(q_n, p_a^n)_{n \in \mathbb{N}} \in \left(Q \times \Delta\left(Y\right)\right)^{\mathbb{N}}$ is such that $\lim_{n \to \infty} \left(q_n, p_a^n\right)_{n \in \mathbb{N}} = \left(q', \bar{p}_a\right)$ and $\mathrm{supp}\, p_a^n \subseteq \{y \in Y : \max_{q \in Q} -\ln \tilde{q}_a\left(y\right) \leq K\}$ then*

$$\lim_{n \to \infty} -\int_Y \log\left(\tilde{q}_{a,n}\left(y\right)\right) dp_a^n\left(y\right) = -\int_Y \log\left(\tilde{q}_a'\left(y\right)\right) d\bar{p}_a\left(y\right).$$

Proof. Recall that weak and vague convergence for probability measures are equivalent when the state space is compact, and that Y is indeed compact. Therefore, by Assumptions 1-2 the assumptions of Theorem 15.7.3 in Kallenberg (1973) are satisfied for the sequence of integrand functions and probability measures $\left(\log\left(\tilde{q}_{a,n}\right), p_a^n\right)_{n \in \mathbb{N}}$. ∎

The following lemma shows that under every policy, almost surely the infinite sequence of observations do not contain a realization that provides an arbitrarily large evidence against a structured model.

Lemma 6. *For every $\Pi \in A^{\mathcal{H}}$,*

$$\mathbb{P}_{\Pi}\left(\left\{(a_i, y_i)_{i \in \mathbb{N}} \in \left(A \times Y\right)^{\mathbb{N}} : \forall t \in \mathbb{N}, \forall q \in Q, -\ln \tilde{q}_{a_t}\left(y_t\right) \leq K\right\}\right) = 1.$$

Proof. By the definition of Radon-Nykodim derivative and equation (10), for every $t \in \mathbb{N}$,

$$\mathbb{P}_\Pi \left(\left\{ (a_i, y_i)_{i \in \mathbb{N}} \in (A \times Y)^\mathbb{N} : \exists q \in Q, -\ln \tilde{q}_{a_t}(y_t) > K \right\} \right) = 0.$$

Since $\mathbb{P}_\Pi$ is a measure, it is countably subadditive and so

$$\mathbb{P}_\Pi \left(\left\{ (a_i, y_i)_{i \in \mathbb{N}} \in (A \times Y)^\mathbb{N} : \forall t \in \mathbb{N}, \forall q \in Q, -\ln \tilde{q}_{a_t}(y_t) \leq K \right\} \right)$$

$$= 1 - \mathbb{P}_\Pi \left(\left\{ (a_i, y_i)_{i \in \mathbb{N}} \in (A \times Y)^\mathbb{N} : \exists t \in \mathbb{N}, \exists q \in Q, -\ln \tilde{q}_{a_t}(y_t) > K \right\} \right)$$

$$\geq 1 - \sum_{t=1}^\infty \mathbb{P}_\Pi \left(\left\{ (a_i, y_i)_{i \in \mathbb{N}} \in (A \times Y)^\mathbb{N} : \exists q \in Q, -\ln \tilde{q}_{a_t}(y_t) > K \right\} \right) = 1,$$

proving the statement. $\blacksquare$

The following lemma shows that, on every history where the empirical action process stabilizes on α^*, and the empirical outcome distribution contingent on the actions played infinitely often converges to the true distribution, the limit LLR can be rewritten as the minimum of an α^* weighted average of the relative entropy from the true DGP.

Lemma 7. *Let $\alpha^* \in \Delta(A)$ and $(a_t, y_t)_{t \in \mathbb{N}} \in (A \times Y)^\mathbb{N}$ be such that $\max_{q \in Q, t \in \mathbb{N}} -\ln \tilde{q}_{a_t}(y_t) \leq K$. For every $t \in \mathbb{N}$, set $h_t = (a^t, y^t)$, and let $q(h_t)$ and $r(h_t)$ be two arbitrary elements of $\operatorname{argmax}_{q \in Q} \prod_{\tau=1}^t \tilde{q}_{a_\tau}(y_\tau)$ and $\operatorname{argmax}_{p \in N(Q)} \prod_{\tau=1}^t \tilde{p}_{a_\tau}(y_\tau)$, respectively. If*

$$\lim_{t \to \infty} \left(\alpha_t(h_t), \left(p_a^{h_t} \right)_{a \in \operatorname{supp} \alpha^*} \right) = \left(\alpha^*, (p_a^*)_{a \in \operatorname{supp} \alpha^*} \right)$$

then

$$\frac{LLR(h_t, Q)}{t} = \lim_{t \to \infty} \frac{\sum_{a \in A} \sum_{\tau=1}^t \mathbb{I}_{\{a\}}(a_\tau) \int_Y \log \left(\frac{\tilde{r}_a(h_t)(y)}{\tilde{q}_a(h_t)(y)} \right) dp_a^{h_t}(y)}{t} = \min_{q \in Q} \sum_{a \in A} \alpha^*(a) R(p_a^* || q_a).$$

Proof. By assumption of the lemma, for all $t \in \mathbb{N}$, we have

$$\sum_{a \in A} \sum_{\tau=1}^t \mathbb{I}_{\{a\}}(a_\tau) \int_Y \log \left(\frac{\tilde{r}_a(h_t)(y)}{\tilde{q}_a(h_t)(y)} \right) dp_a^{h_t}(y) / t$$

$$= \left(\sum_{a \in A} \sum_{\tau=1}^t \mathbb{I}_{\{a\}}(a_\tau) \left(\int_Y \log (\tilde{r}_a(h_t)(y)) dp_a^{h_t}(y) - \int_Y \log (\tilde{q}_a(h_t)(y)) dp_a^{h_t}(y) \right) \right) / t.$$

By Assumption 2 (i), we have

$$\sum_{a \in A} \sum_{\tau=1}^{t} \mathbb{I}_{\{a\}} (a_\tau) \int_Y - \log \left(\tilde{r}_a (h_t) (y) \right) \mathrm{d}p_a^{h_t} (y) / t \leq 0 \qquad \forall t \in \mathbb{N}.$$

By Assumption 2 (ii) and the Arzela-Ascoli Theorem, there exists $K' \in \mathbb{R}_{++}$ such that for all $a \in A$,

$$- \log \left(\tilde{r}_a (h_t) (y) \right) \geq -K', \tag{15}$$

p_a^*-almost surely. Therefore, for all $t \in \mathbb{N}$

$$\sum_{a \in A} \sum_{\tau=1}^{t} \mathbb{I}_{\{a\}} (a_\tau) \int_Y - \log \left(\tilde{r}_a (h_t) (y) \right) \mathrm{d}p_a^{h_t} (y) / t \subseteq [0, K].$$

We know show that every convergent subsequence converges to 0. Indeed, take any such subsequence of periods $(t_n)_{n \in \mathbb{N}}$ and, possibly restricting to a further subsequence, suppose $\tilde{r}_a (h_{t_n})$ converges to some $\tilde{r}_a$ for every $a \in \mathrm{supp}\alpha^*$. Then

$$0 \leq \sum_{a \in A} \alpha^* (a) \int_Y - \log \left(\tilde{r}_a (y) \right) \mathrm{d}p_a^* (y) \leq \sum_{a \in A} \alpha^* (a) \liminf_{n \to \infty} \int_Y - \log \left(\tilde{r}_a (h_{t_n}) (y) \right) \mathrm{d}p_a^{h_{t_n}} (y)$$

$$= \liminf_{n \to \infty} \sum_{a \in A} \sum_{\tau=1}^{t_n} \mathbb{I}_{\{a\}} (a_\tau) \int_Y - \log \left(\tilde{r}_a (h_{t_n}) (y) \right) \mathrm{d}p_a^{h_{t_n}} (y) / t_n$$

where the first inequality follows from Gibbs inequality and the second since by equation (15) we can apply Lemma 3.2 in Serfozo (1982). Therefore, we have

$$\lim_{n \to \infty} \sum_{a \in A} \sum_{\tau=1}^{t} \mathbb{I}_{\{a\}} (a_\tau) \int_Y - \log \left(\tilde{r}_a (h_{t_n}) (y) \right) \mathrm{d}p_a^{h_{t_n}} (y) / t = 0.$$

Since this holds for every converging subsequence, it holds for the original sequence. So

$$\lim_{t \to \infty} \sum_{a \in A} \sum_{\tau=1}^{t} \mathbb{I}_{\{a\}} (a_\tau) \int_Y \log \left(\frac{\tilde{r}_a (h_t) (y)}{\tilde{q}_a (h_t) (y)} \right) \mathrm{d}p_a^{h_t} (y) / t$$

$$= - \lim_{t \to \infty} \sum_{a \in A} \sum_{\tau=1}^{t} \mathbb{I}_{\{a\}} (a_\tau) \int_Y \log \left(\tilde{q}_a (h_t) (y) \right) \mathrm{d}p_a^{h_t} (y) / t$$

$$= - \lim_{t \to \infty} \min_{q \in Q} \sum_{a \in A} \sum_{\tau=1}^{t} \mathbb{I}_{\{a\}} (a_\tau) \int_Y \log \left(\tilde{q}_a (y) \right) \mathrm{d}p_a^{h_t} (y) / t.$$

44

Therefore the result follows from Lemma 5 and Theorem 17.31 in Aliprantis and Border (2013). ∎

Lemma 8. *For every $a \in A$,*

$$\limsup_{k \to \infty} \min_{q \in Q} \min_{p_a \in \Delta(Y)} \int_Y u(a,y)\, dp_a(y) + \frac{R(p_a||q_a)}{k} = \min_{y \in \cup_{q \in Q} \mathrm{supp} q_a} u(a,y).$$

Proof. Let $\hat{y} \in \operatorname{argmin}_{y \in \cup_{q \in Q} \mathrm{supp} q_a} u(a,y)$. If $\max_{y \in Y} u(a,y) = u(a,\hat{y})$ the statement is trivially true, so suppose that $\max_{y \in Y} u(a,y) > u(a,\hat{y})$. By Assumption 1 (i) we have that

$$\inf_{q \in Q} q_a(B_\varepsilon(\hat{y})) > 0 \qquad \forall \varepsilon \in \mathbb{R}_{++}.$$

Otherwise, by the compactness of Q the portmanteau theorem (see, e.g., Theorem 11.1.1 Dudley, 2018) would imply that there exists $\hat{q} \in Q$ with $\hat{q}_a(B_{\varepsilon/2}(\hat{y})) = 0$. But then, since there exists $\bar{q} \in Q$ with $\hat{y} \in \mathrm{supp} \bar{q}_a = \mathrm{supp} p_a^*$, and so $p_a^*(B_{\varepsilon/2}(\hat{y})) > 0$, we would obtain a contradiction with $p_a^* \sim \hat{q}_a$. Fix $\bar{\varepsilon} \in \left(0, \frac{\max_{y \in Y} u(a,y) - u(a,\hat{y})}{2}\right)$. Since $u(a,\cdot)$ is continuous, there exists ε such that

$$y \in B_\varepsilon(\hat{y}) \implies u(a,y) \le u(a,\hat{y}) + \bar{\varepsilon}.$$

Then, for all $q \in Q$

$$
\begin{aligned}
u(a,\hat{y}) &\le \min_{p_a \in \Delta(Y)} \int_Y u(a,y)\, dp_a + \frac{R(p_a||q_a)}{k} = -\frac{1}{k} \log\left(\int_Y \exp(-ku(a,y))\, dq_a(y)\right) \\
&\le -\frac{1}{k} \log\left(\begin{array}{l} \exp(-k(u(a,\hat{y}) + \bar{\varepsilon})) \inf_{\hat{q} \in Q} \hat{q}_a(B_\varepsilon(\hat{y})) \\ + (1 - \inf_{\hat{q} \in Q} \hat{q}_a(B_\varepsilon(\hat{y}))) \exp(-k \max_{y \in Y} u(a,y)) \end{array}\right)
\end{aligned}
$$

where the equality follows from Proposition 1.4.2 in Dupuis and Ellis (2011). Moreover, the last term converges to $u(a,\hat{y}) + \bar{\varepsilon}$ as k goes to infinity by a simple application of L'Hôpital's rule. Since $\bar{\varepsilon} < \frac{\max_{y \in Y} u(a,y) - u(a,\hat{y})}{2}$ was arbitrarily chosen, and the last term does not depend on q this proves the desired uniformity of the convergence. ∎

Proof of Proposition 1. Let (u,a,Y) be a decision problem where (i) and (ii) are satisfied and $\varepsilon \in \mathbb{R}_{++}$. We start by showing that there exists $c \in \mathbb{R}_{++}$ such that

$$\Lambda(h_t) = \frac{LLR(h_t, Q)}{ct} \qquad \forall t \in \mathbb{N}, \forall h_t \in \mathcal{H}_t, \tag{16}$$

is ε-safe and ε-consistent under correct specification. This is done by first deriving a $c \in \mathbb{R}_{++}$ such that ε-safety is satisfied, and then showing that there exists a δ that delivers ε-consistency under correct specification. Safety is trivially satisfied by every policy if $\max_{a \in A} \min_{y \in Y} u(a, y) = \min_{a \in A, y \in Y} u(a, y)$, so in that case pick an arbitrary $c \in \mathbb{R}_{++}$. Suppose instead that we have $\max_{a \in A} \min_{y \in Y} u(a, y) > \min_{a \in A, y \in Y} u(a, y)$. Let $\hat{P} \subseteq \Delta(Y)^A$ be the set of p^* that satisfy Assumption 1, (i) and (ii) jointly with Q, and define

$$\underline{A}(p^*) := \left\{ a' \in A : \max_{a \in A} \min_{y \in Y} u(a, y) > \mathbb{E}_{p^*_{a'}}[u(a', y)] + \frac{\varepsilon}{2} \right\}.$$

Claim 2. *There exists $\varphi^* > 0$ such that for every $\Pi \in A^{\mathcal{H}}$ and $p^* \in \hat{P}$,*

$$\mathbb{P}_\Pi \left(\begin{array}{c} \left\{ \liminf_{t \to \infty} \frac{\sum_{i=1}^t u(\mathbf{a}_i, \mathbf{y}_i)}{t} - \max_{a \in A} \min_{y \in Y} u(a, y) - \varepsilon < 0 \right\} \\ \cap \left\{ \limsup_{t \to \infty} \alpha_t(\mathbf{h}_t)(a') < \varphi^*, \forall a' \in \underline{A}(p^*) \right\} \end{array} \right) = 0.$$

That is, almost surely the payoff is at most ε-lower than the safe guarantee if the actions whose objective expected performance is lower than the guarantee are played sufficiently rarely (i.e., each of them has an average frequency lower than φ^*).

Proof of the Claim. Consider the stochastic process defined by

$$\mathbf{X}_t = u(\Pi(\mathbf{h}_{t-1}), \mathbf{y}_t) - \mathbb{E}_{p^*_{\Pi(\mathbf{h}_{t-1})}}[u(\Pi(\mathbf{h}_{t-1}), y)] \quad \forall t \in \mathbb{N}$$

with the sequence of sigma-algebras $(\mathcal{F}_t)_{t \in \mathbb{N}}$ generated by the stochastic process of histories $(\mathbf{h}_t)_{t \in \mathbb{N}}$. The stochastic process is not i.i.d., as previous utility realizations affect current period choices. Nevertheless it is a martingale difference sequence, as u is continuous in y on the compact Y, so $\mathbb{E}[|\mathbf{X}_t|] \leq 2 \max_{a \in A, y \in Y} |u(a, y)| < \infty$ and $\mathbb{E}[\mathbf{X}_t | \mathcal{F}_{t-1}] = 0$ by equation (10). A fortiori, $(\mathbf{X}_t)_{t \in \mathbb{N}}$ is a mixingale difference sequence, and by the strong law of large numbers for mixingale sequences (see Theorem 2.7 in Hall and Heyde, 2014 for the version that applies here), we have

$$\lim_{n \to \infty} \frac{\sum_{t=1}^n \mathbf{X}_t}{n} = 0 \qquad \mathbb{P}_\Pi\text{-a.s.}$$

so that

$$\liminf_{t\to\infty} \frac{\sum_{i=1}^{t} u\left(\mathbf{a}_i, \mathbf{y}_i\right)}{t} = \liminf_{t\to\infty} \sum_{i=1}^{t} \frac{\mathbf{X}_i + \mathbb{E}_{p^*_{\mathbf{a}_t}}\left[u\left(\mathbf{a}_t, \cdot\right)\right]}{t}$$

$$\geq \left(1 - \limsup_{t\to\infty} \alpha_t\left(\mathbf{h}_t\right)\left(\underline{A}\left(p^*\right)\right)\right)\left(\max_{\bar{a}\in A}\min_{y\in Y} u\left(\bar{a}, y\right) - \frac{\varepsilon}{2}\right) + \limsup_{t\to\infty} \alpha_t\left(\mathbf{h}_t\right)\left(\underline{A}\left(p^*\right)\right) \min_{a\in a, y\in Y} u\left(a, y\right)$$

$$\geq \left(1 - \sum_{a\in\underline{A}(p^*)} \limsup_{t\to\infty} \alpha_t\left(\mathbf{h}_t\right)\left(a\right)\right) \max_{\bar{a}\in A}\min_{y\in Y} u\left(\bar{a}, y\right) - \frac{\varepsilon}{2} + \sum_{a\in\underline{A}(p^*)} \limsup_{t\to\infty} \alpha_t\left(\mathbf{h}_t\right)\left(a\right) \min_{a\in a, y\in Y} u\left(a, y\right)$$

$$\geq \left(1 - |A| \max_{a\in\underline{A}(p^*)} \limsup_{t\to\infty} \alpha_t\left(\mathbf{h}_t\right)\left(a\right)\right) \max_{\bar{a}\in A}\min_{y\in Y} u\left(\bar{a}, y\right) - \frac{\varepsilon}{2}$$

$$+ \left(|A| \max_{a\in\underline{A}(p^*)} \limsup_{t\to\infty} \alpha_t\left(\mathbf{h}_t\right)\left(a\right)\right) \min_{a\in a, y\in Y} u\left(a, y\right)$$

and therefore the claim follows from setting

$$\frac{\varepsilon}{2\left(\max_{\bar{a}\in A}\min_{y\in Y} u\left(\bar{a}, y\right) - \min_{a\in a, y\in Y} u\left(a, y\right)\right)|A|} = \varphi^*.$$

□

Claim 3. *There exists $\bar{\lambda} \in \mathbb{R}_{++}$ such that if $\lambda \geq \bar{\lambda}$ then for every $p^* \in \hat{P}$, $a' \in \underline{A}\left(p^*\right)$, $\nu \in \Delta\left(Q\right)$, we have $a' \notin BR^{\lambda}\left(\nu\right)$.*

That is, if the agent is sufficiently misspecification concerned, they do not play actions that can perform worse than the safe guarantee.

Proof of the Claim. Observe that if $\underline{A}\left(p^*\right) \neq \emptyset$, then by Assumption 1 (i) for all $q \in Q$, there is $y \in \operatorname{supp} q_{a'}$ with $u\left(a', y\right) \leq \max_{\bar{a}\in A}\min_{y\in Y} u\left(\bar{a}, y\right) - \frac{\varepsilon}{2}$. But then the claim follows from Lemma 8. □

Claim 4. *There exists $J \in \left(0, 1\right)$ such that for every $p^* \in \hat{P}$, $a' \in \underline{A}\left(p^*\right)$, $\mu \in \Delta\left(Q\right)$, and $\lambda \in \mathbb{R}_+$,*

$$\mu\left(\left\{q \in Q : R\left(p^*_{a'}||q_{a'}\right) > J\right\}\right) \leq J \implies a' \notin BR^{\lambda}\left(\mu\right). \tag{17}$$

That is, if the beliefs are sufficiently concentrated on the parameters that are close to the true DGP, and under the true DGP a' performs worse than the safe guarantee, a' cannot be chosen regardless of the level of misspecification concern.

Proof of the Claim. Observe that given Claim 3, the statement immediately holds for $\lambda > \bar{\lambda}$. Suppose by contradiction that equation (17) does not hold true. This means that there exist a convergent $\left(p^*_n, \mu_n, \lambda_n\right)_{n\in\mathbb{N}} \in \hat{P} \times \Delta\left(Q\right) \times \left[0, \bar{\lambda}\right]$ and $a' \in A$ with

$\mu_n \left(\left\{ q \in Q : R \left(p_{a',n}^* || q_{a'} \right) > \frac{1}{n} \right\} \right) \leq \frac{1}{n}$, $a' \in \underline{A} (p_n^*)$, and $a' \in BR^{\lambda_n} (\mu_n)$. By the lower semicontinuity of R and the fact that $R \left(p_{a'} || q_{a'} \right) = 0$ if and only if $p_{a'} = q_{a'}$, (see, e.g., Lemma 1.4.3 in Dupuis and Ellis (2011)), as well as Lemma 4 this in turn implies that there exists $q \in Q$ with $a' \in \underline{A} (q)$ and $\mathbb{E}_{q_{a'}} [u (a', y)] \geq \max_{\bar{a} \in A} \min_{y \in Y} u (\bar{a}, y)$, a contradiction. $\square$

Let $c = \frac{J \varphi^*}{4 \lambda}$, and take an arbitrary $p^* \in \hat{P}$. If for all $a' \in \underline{A} (p^*) \limsup_{t \to \infty} \alpha_t (\mathbf{h}_t) (a') < \varphi^*$, $\mathbb{P}_\Pi$-a.s., ε-safety follows by Claim 2. Suppose by contradiction that there is an action $a' \in \underline{A} (p^*)$ with $\limsup_{t \to \infty} \alpha_t (\mathbf{h}_t) (a') \geq \varphi^*$. By Claim 4, it must be the case that $\min_{q \in Q} R \left(p_{a'}^* || q_{a'} \right) = R \left(p_{a'}^* || q_{a'}^* \right) \geq J$.

But then, by Theorem 11.4.1 in Dudley (2018) as well as Lemmas 6 and 7 we have that

$$\liminf_{t \to \infty} \Lambda (\mathbf{h}_t) \geq \frac{\min_{q \in Q} R (p^* || q)}{c} \geq \frac{J}{c} = 2 \bar{\lambda} \qquad \mathbb{P}_\Pi\text{-a.s.}$$

Then by Claim 3, we have that for all $a' \in \underline{A} (p^*)$

$$\limsup_{t \to \infty} \alpha_t (\mathbf{h}_t) (a') = 0 \qquad \mathbb{P}_\Pi\text{-a.s.}$$

a contradiction.

Since Q is compact, for every $\epsilon \in (0, 1)$ we can pick $\delta < 0$ such that for all $p^* \in \hat{P}$, and $a \in A$, $\min_{q \in Q} R \left(p_a^* || q_a \right) < \delta$ implies that for $q \in Q^\epsilon (a)$, by Lemma 3,

$$\mathbb{E}_{p_{a'}^*} [u (a', y)] - \min_{p \in \Delta(Y)} \left(\mathbb{E}_p [u (a', y)] + \frac{R (p_{a'} || q_{a'})}{\lambda} \right) \leq \frac{\varepsilon}{4} \qquad \forall a' \in A, \forall \lambda \in [0, 2\delta].$$

But then Λ is ε-consistent with this δ by Lemmas 6, 7, and Berk (1966), page 54.

We show that there is a decision problem (u, a, Y) such that if the concern for misspecification of the agent is such that

$$\Lambda (\mathbf{h}_t) = o \left(\frac{LLR (\mathbf{h}_t, Q)}{t} \right) \qquad \mathbb{P}_\Pi\text{-a.s.}$$

then the decision rule is not $\frac{1}{10}$-safe. Suppose that $A = \{1, -1, 0\}$ and $Y = \{-1, 1\}$. The utility function is $u (a, y) = ay$. Each model q considered by the agent is described by $q_a (1)$ for some arbitrary $a \in A$. With this, let $Q = \{0.9, 0.4\}$, $p_a^* (1) = 0.6$, and $\mu (0.9) = \frac{1}{2} = \mu (0.4)$. Let $N (Q) = [0, 1]$, i.e., the unstructured models include all the action-independent data-generating processes. We have $\max_{\bar{a} \in A} \min_{y \in Y} u (\bar{a}, y) = \min_{y \in Y} u (0, y) = 0$. However, by the Strong Law of Large Numbers it follows that $\mathbb{P}_\Pi$-

almost surely $\lim_{t\to\infty} \left(\sum_{\tau=1}^{t} \mathbb{I}_{\{1\}}(\mathbf{y}_\tau)\right)/t = 0.6$. Therefore, by Lemma 2 we have that

$$\lim_{t\to\infty} \frac{LLR(\mathbf{h}_t, Q)}{t} = R(0.6||0.4) \qquad \mathbb{P}_\Pi\text{-a.s.}$$

and so

$$\lim_{t\to\infty} \Lambda(\mathbf{h}_t) = 0 \qquad \mathbb{P}_\Pi\text{-a.s.}$$

Moreover, for the constant function $\phi(\varepsilon) = \frac{1}{2}$ for all $\varepsilon \in \mathbb{R}_{++}$ the prior is ϕ-positive on Q in the sense of Fudenberg et al. (2022a), and by their Lemma 1, $\mu(0.4|\mathbf{h}_t) \to 1$, $\mathbb{P}_\Pi$-a.s. But then by the upper hemicontinuity of $BR^{(\cdot)}(\cdot)$ established in Lemma 4

$$\liminf_{t\to\infty} \frac{\sum_{i=1}^{t} u(\mathbf{a}_i, \mathbf{y}_i)}{t} = -0.2 < 0 = \max_{\bar{a}\in A} \min_{y\in Y} u(\bar{a}, y) \qquad \mathbb{P}_\Pi\text{-a.s.}$$

proving the desired result.

Finally, we show that there is a decision problem (u, a, Y) such that if the concern for misspecification of the agent is such that

$$o(\Lambda(\mathbf{h}_t)) = \frac{LLR(\mathbf{h}_t, Q)}{t} \qquad \mathbb{P}_\Pi\text{-a.s.}$$

then the decision rule is not $\frac{1}{10}$-consistent. Let Π be a Λ-optimal policy.

Let $\delta \in (0, 0.4)$ and suppose $A = \{1, -1, 0\}$ and $Y = \{-1, 1\}$. The utility function is $u(a, y) = ay$. Again, each model q considered by the agent is described by $q_a(1)$ for some arbitrary $a \in A$. With this, let $Q = \{0.6, 0.4\}$ and $p_a^*(1) = 0.6 + \delta$. Let $N(Q) = [0, 1]$, i.e., the unstructured models include all the action-independent data-generating processes.

Let $\bar{\lambda}$ be such that $\{0\} = BR^\lambda(\mu)$ for all $\lambda \geq \bar{\lambda}$ and $\mu \in \Delta(Q)$. Such a $\bar{\lambda}$ exists because for $a \in \{-1, 1\}$ and $q \in Q$, $\lim_{\lambda\to\infty} \min_{p_a\in\Delta(Y)} \mathbb{E}_{p_a}[u(a, y)] + \frac{R(p_a||q_a)}{\lambda} = -1$. Let

$$C_t = \left\{ h_t \in \mathcal{H}_t : \left(\sum_{\tau=1}^{t} \mathbb{I}_{\{1\}}(y_\tau)\right)/t \geq 0.6, R\left(\left(\sum_{\tau=1}^{t} \mathbb{I}_{\{1\}}(y_\tau)\right)/t||0.6\right) \geq R(0.6 + \delta||0.6)/2 \right\}.$$

For every $h_t \in C_t$, by Lemmas 2, 6, and 7, $\lim_{t\to\infty} LLR(h_t, Q)/t \geq R(0.6 + \delta||0.6)/2$ so that $\Lambda(h_t)$ is diverging to $+\infty$ and it is eventually larger than $\bar{\lambda}$. But by Sanov's Theorem (see, e.g., Theorem 2.2.1 in Dupuis and Ellis, 2011), $\lim_{t\to\infty} \mathbb{P}_\Pi[C_t] = 1$, so the result follows. ∎

Proof of Proposition 2. Suppose that a^* is a Λ-limit action. Thus, since for every

policy $\Pi \in A^{\mathcal{H}}$

$$\mathbb{P}_{\Pi}\left[\sup\{\mathbf{t}\colon \mathbf{a_t} \neq a^*\} < \infty\right] \leq \sum_{t=0}^{\infty} \sum_{h_t \in \mathcal{H}_t} \mathbb{P}_{\Pi}\left[a^* \in BR^{\Lambda(\mathbf{h}_\tau)}\left(\mu\left(\cdot|\mathbf{h}_\tau\right)\right), \forall \tau \geq t | h_t\right] \mathbb{P}_{\Pi}[h_t],$$

there are a Λ-optimal policy $\tilde{\Pi} \in A^{\mathcal{H}}$, $t \in \mathbb{N}_0$, and $(a^t, y^t) \in \mathcal{H}_t$ with $\mathbb{P}_{\tilde{\Pi}}[(a^t, y^t)] > 0$ such that with positive probability $\tilde{\Pi}$ prescribes a^* after (a^t, y^t) in every future period. Define $\nu = \mu(\cdot|(a^t, y^t))$, and notice that by Assumption 1 (i) $\operatorname{supp}\nu = \operatorname{supp}\mu = Q$. As the evolution of beliefs and misspecification concern under Π^{a^*}, i.e., the policy that plays a^* in every period, is the same as under $\tilde{\Pi}$ for every history where the agent continues to play a^*, we have that

$$\mathbb{P}_{\tilde{\Pi}}[\forall \tau > t, a^* = \tilde{\Pi}(\mathbf{h}_\tau)|(a^t, y^t)] > 0 \implies \mathbb{P}_{\Pi^{a^*}}[\forall \tau > 0, a^* \in BR^{\Lambda\left((a^t, y^t), \mathbf{h}_\tau\right)}\left(\nu\left(\cdot|\mathbf{h}_\tau\right)\right)] > 0.$$

We now show that if a^* is not a self-confirming equilibrium, the latter equals zero, which establishes that a^* cannot be a Λ-limit action. By the strong law of large numbers (see, e.g., Theorem 8.3.5 in Dudley, 2018),

$$\lim_{\tau \to \infty} p_{a^*}^{\mathbf{h}_\tau} = p_{a^*}^* \qquad \mathbb{P}_{\Pi^{a^*}}\text{-a.s.}$$

Therefore, since by Lemma 6, $\max_{q \in Q, t \in \mathbb{N}} - \ln \tilde{q}_{\mathbf{a}_t}(\mathbf{y}_t) \leq K$, $\mathbb{P}_{\Pi^{a^*}}$-almost surely, by Lemma 7 we have $\lim_{\tau \to \infty} \frac{LLR\left((a^t, y^t), \mathbf{h}_\tau, Q\right)}{\tau} = 0 = \lim_{\tau \to \infty} \frac{\Lambda\left((a^t, y^t), \mathbf{h}_\tau\right)}{\tau}$, $\mathbb{P}_{\Pi^{a^*}}$ almost surely. With this, as by Assumption 1 (ii), the assumptions of Berk (1966), page 54, are satisfied, for every $\varepsilon \in \mathbb{R}_{++}$ we have

$$\lim_{\tau \to \infty} \nu\left(B_\varepsilon\left(\{q \in Q : q_{a^*} = p_{a^*}^*\}\right)|\mathbf{h}_\tau\right) = 1, \qquad \mathbb{P}_{\Pi^{a^*}}\text{-a.s.}$$

and the desired conclusion follows from Lemma 4. ∎

Proof of Theorem 1. We start with the preliminary observation that by Lemma 6, $\max_{q \in Q, t \in \mathbb{N}} - \ln \tilde{q}_{\mathbf{a}_t}(\mathbf{y}_t) \leq K$, $\mathbb{P}_{\Pi^{a^*}}$-almost surely. This will allow us to invoke Lemma 7 in all the various cases.

1) Suppose by contradiction that a^* is a Λ-limit action but is not a Berk-Nash equilibrium. Thus, since for every policy $\Pi \in A^{\mathcal{H}}$

$$\mathbb{P}_{\Pi}\left[\sup\{\mathbf{t}\colon \mathbf{a_t} \neq a^*\} < \infty\right] \leq \sum_{t=0}^{\infty} \sum_{h_t \in \mathcal{H}_t} \mathbb{P}_{\Pi}\left[a^* \in BR^{\Lambda(\mathbf{h}_\tau)}\left(\mu\left(\cdot|\mathbf{h}_\tau\right)\right), \forall \tau \geq t | h_t\right] \mathbb{P}_{\Pi}[h_t],$$

there are a Λ-optimal policy $\tilde{\Pi} \in A^{\mathcal{H}}$, $t \in \mathbb{N}_0$, and $(a^t, y^t) \in \mathcal{H}_t$ with $\mathbb{P}_{\tilde{\Pi}}[(a^t, y^t)] > 0$ such that with positive probability $\tilde{\Pi}$ prescribes a^* after (a^t, y^t) in every future period. Define $\nu = \mu(\cdot|(a^t, y^t))$, and notice that by Assumption 1 (i) supp $\nu = $ supp $\mu = Q$. As the evolution of beliefs and misspecification concern under Π^{a^*}, i.e., the policy that plays a^* in every period, is the same as under $\tilde{\Pi}$ for every history where the agent continues to play a^*, we have that

$$\mathbb{P}_{\tilde{\Pi}}[\forall \tau > t, a^* = \tilde{\Pi}(\mathbf{h}_\tau)|(a^t, y^t)] > 0 \implies \mathbb{P}_{\Pi^{a^*}}[\forall \tau > 0, a^* \in BR^{\Lambda\left((a^t,y^t),\mathbf{h}_\tau\right)}\left(\nu\left(\cdot|\mathbf{h}_\tau\right)\right)] > 0.$$

We now show that the latter equals zero, which establishes that a^* cannot be a Λ-limit action.

Since Y is a compact metric space, it is separable (see, e.g., Proposition 9.24 in Royden and Fitzpatrick, 1988). Thus, by Theorem 11.4.1 in Dudley (2018), $\lim_{\tau \to \infty} p^{\mathbf{h}_\tau}_{a^*} = p^*_{a^*}$, $\mathbb{P}_{\Pi^{a^*}}$-a.s. Then, by Lemma 7 and equation (5) we have $\lim_{\tau \to \infty} \Lambda\left((a^t, y^t), \mathbf{h}_\tau\right) = 0$, $\mathbb{P}_{\Pi^{a^*}}$-a.s. By Assumption 1 (ii), the assumptions of Berk (1966), page 54, are satisfied, and we have that for every $\varepsilon \in \mathbb{R}_{++}$, $\nu\left(Q^\varepsilon\left(a^*\right)|\mathbf{h}_\tau\right) \to 1$, $\mathbb{P}_{\Pi^{a^*}}$-a.s. Therefore, since Q is compact, $\left(\Lambda\left((a^t, y^t), \mathbf{h}_\tau\right), \nu\left(\cdot|\mathbf{h}_\tau\right)\right)_{\tau \in \mathbb{N}}$ admits $\mathbb{P}_{\Pi^{a^*}}$ almost surely a subsequence convergent to $(0, \nu^*)$ for some $\nu^* \in \Delta\left(Q\left(a^*\right)\right)$. With this, the result follows from Lemma 4.

2) Suppose by contradiction that $a^* \notin BR^{Meu}\left(\left\{p \in \Delta\left(Y\right)^A : \exists q \in Q, \forall a \in A, q_a \gg p_a\right\}\right)$ and that a^* is a Λ-limit action. Thus, since for every policy $\Pi \in A^{\mathcal{H}}$

$$\mathbb{P}_\Pi\left[\sup\{\mathbf{t}\colon \mathbf{a_t} \neq a^*\} < \infty\right] \leq \sum_{t=0}^{\infty} \sum_{h_t \in \mathcal{H}_t} \mathbb{P}_\Pi\left[a^* \in BR^{\Lambda(\mathbf{h}_\tau)}\left(\mu\left(\cdot|\mathbf{h}_\tau\right)\right), \forall \tau \geq t|h_t\right] \mathbb{P}_\Pi[h_t],$$

there are a Λ-optimal policy $\tilde{\Pi} \in A^{\mathcal{H}}$, $t \in \mathbb{N}_0$, and $(a^t, y^t) \in \mathcal{H}_t$ with $\mathbb{P}_{\tilde{\Pi}}[(a^t, y^t)] > 0$ such that with positive probability $\tilde{\Pi}$ prescribes a^* after (a^t, y^t) in every future period. Define $\nu = \mu(\cdot|(a^t, y^t))$, and notice that by Assumption 1 (i) supp $\nu = $ supp $\mu = Q$. As the evolution of beliefs and misspecification concern under Π^{a^*}, i.e., the policy that plays a^* in every period, is the same as under $\tilde{\Pi}$ for every history where the agent continues to play a^*, we have that

$$\mathbb{P}_{\tilde{\Pi}}[a^* = \tilde{\Pi}(\mathbf{h}_\tau), \forall \tau > t|(a^t, y^t)] > 0 \implies \mathbb{P}_{\Pi^{a^*}}[a^* \in BR^{\Lambda\left((a^t,y^t),\mathbf{h}_\tau\right)}\left(\nu\left(\cdot|\mathbf{h}_\tau\right)\right), \forall \tau > t] > 0.$$

We now show that the latter equals zero, which establishes that a^* cannot be a Λ-limit action.

By Theorem 11.4.1 in Dudley (2018),

$$\lim_{\tau \to \infty} p_{a*}^{\mathbf{h}_\tau} = p_{a*}^* \qquad \mathbb{P}_{\Pi^{a*}}\text{-a.s.}$$

Then, by Lemmas 2 and 7, and equation (6), we have

$$\lim_{\tau \to \infty} \Lambda\left((a^t, y^t), \mathbf{h}_\tau\right) = \infty \qquad \mathbb{P}_{\Pi^{a*}}\text{-a.s.}$$

By Assumption 1 (i) for all $q, q' \in Q$ and $a \in A$ we have $q_a \sim q'_a$. So we obtain

$$\left\{p \in \Delta\left(Y\right)^A : \exists q \in Q, \forall a \in A, q_a \gg p_a\right\} = \left\{p \in \Delta\left(Y\right)^A : \forall q \in Q, \forall a \in A, q_a \gg p_a\right\}.$$

Therefore, by Lemma 8 for all $a \in A$ we have that $\mathbb{P}_{\Pi^{a*}}$ almost surely

$$\limsup_{\tau \to \infty} \min_{q \in Q} \min_{p_a \in \Delta(Y)} \int_Y u\left(a, y\right) \mathrm{d}p_a + \frac{R\left(p_a || q_a\right)}{\Lambda\left((a^t, y^t), \mathbf{h}_\tau\right)} = \min_{y \in \cup_{q \in Q} \mathrm{supp} q_a} u\left(a, y\right).$$

But since by Assumption 1 (i) for all $\tau \in \mathbb{N}$, $\mu\left(\cdot | \mathbf{h}_\tau\right) \subseteq Q$, $\mathbb{P}_{\Pi^{a*}}$ almost surely, we have

$$\lim_{\tau \to \infty} \mathbb{E}_{\mu(\cdot|\mathbf{h}_\tau)} \left[\min_{p_a \in \Delta(Y)} \int_Y u\left(a, y\right) \mathrm{d}p_a + \frac{R\left(p_a || q_a\right)}{\Lambda\left((a^t, y^t), \mathbf{h}_\tau\right)}\right] = \min_{y \in \cup_{q \in Q} \mathrm{supp} q_a} u\left(a, y\right) \qquad \mathbb{P}_{\Pi^{a*}}\text{-a.s.}$$

With this, the result follows from the finiteness of the action space.

3) It follows from the more general Theorem 2. ∎

Lemma 9. *For every $c \in \mathbb{R}_{++}$ the function $\alpha \mapsto \min_{q \in Q} \sum_{a \in A} \alpha\left(a\right) R\left(p_a^* || q_a\right) / c$ is continuous and the correspondence $Q\left(\cdot\right) : \Delta\left(A\right) \to 2^Q$ is upper hemicontinuous.*

Proof. We first show that the function

$$\begin{aligned}
\Delta\left(A\right) \times Q &\to \mathbb{R} \\
\left(\alpha, q\right) &\mapsto \sum_{a \in A} \alpha\left(a\right) R\left(p_a^* || q_a\right) / c
\end{aligned} \tag{18}$$

is continuous. Fix an $a \in A$ and let $\left(q_n\right)_{n \in \mathbb{N}} \in Q^{\mathbb{N}}$ be a sequence that converges to $q \in Q$. By Assumption 1 (ii), $\tilde{q}_{a,n}(y)$ is converging to $\tilde{q}_a(y)$ for p_a^* almost every y. Then

$$|R\left(p_a^* || q_{a,n}\right) - R\left(p_a^* || q_a\right)| = \left|\int_Y \log\left(\frac{\tilde{q}_a\left(y\right)}{\tilde{q}_{a,n}\left(y\right)}\right) \mathrm{d}p_a^*\left(y\right)\right|$$

and observe that the integrand on the right-hand side is dominated by a constant by As-

sumption 1 (i). Therefore, by the dominated convergence theorem $|R\left(p_a^*||q_{a,n}\right)-R\left(p_a^*||q_a\right)|$ converges to 0. Since A is finite and the function in equation (18) is linear in α, we have obtained the desired continuity. With this, the statement follows from Theorem 17.31 in Aliprantis and Border (2013).

Proof of Proposition 3. Consider the following three-player game. The action sets are $A_1 = \Delta\left(A\right)$, $A_2 = \Delta\left(Q\right)$, $A_3 = \mathbb{R}_+$ with arbitrary elements denoted as α, ν, λ. The utility functions are

$$U_1\left(\alpha,\nu,\lambda\right) = \begin{cases} \sum_{a\in A}\alpha\left(a\right)\int_Q \min_{p_a\in\Delta(Y)}\left(\mathbb{E}_{p_a}\left[u\left(a,y\right)\right]+\frac{R(p_a||q_a)}{\lambda}\right)\mathrm{d}\nu\left(q\right) & \lambda \neq 0 \\ \sum_{a\in A}\alpha\left(a\right)\int_Q \mathbb{E}_{q_a}\left[u\left(a,y\right)\right]\mathrm{d}\nu\left(q\right) & \lambda = 0, \end{cases}$$

$$U_2\left(\alpha,\nu,\lambda\right) = -\int_Q \sum_{a\in A}\alpha\left(a\right)R\left(p_a^*||q_a\right)\mathrm{d}\nu\left(q\right),$$

$$U_3\left(\alpha,\nu,\lambda\right) = -\left(\lambda - \min_{q\in Q}\sum_{a\in A}\alpha\left(a\right)R\left(p_a^*||q_a\right)/c\right)^2.$$

Observe that for the purpose of finding the equilibria of this game, it is without loss of generality to limit the actions of player 3 to $\left[0,\bar\lambda\right]$ with

$$\bar\lambda: = \frac{\max_{\alpha\in\Delta(A)}\min_{q\in Q}\sum_{a\in A}\alpha\left(a\right)R\left(p_a^*||q_a\right)}{c}$$

$$= \frac{\max_{\alpha\in\Delta(A)}\min_{q\in Q}\sum_{a\in A}\alpha\left(a\right)\int_Y -\log\left(\tilde q_a\left(y\right)\right)\mathrm{d}p_a^*\left(y\right)}{c} < \infty,$$

where the inequality holds by Assumption 1 (i). Therefore, since the compactness of Q implies that also $\Delta\left(Q\right)$ is compact by Theorem 15.11 in Aliprantis and Border (2013) all the three action sets are compact. Moreover, they are clearly convex.

The utility function U_1 is jointly continuous in its second and third argument by Lemma 4. Moreover, U_2 is trivially continuous in its first and third argument and U_3 is continuous in its first and second argument by Lemma 9. Therefore the game is better-reply secure (see Reny, 1999, page 1033). Moreover, U_1 and U_2 are respectively linear in A_1 and A_2 while U_3 is concave in A_3.

Therefore, by Theorem 3.1 and Footnote 8 in Reny (1999) this game admits a pure-strategy equilibrium $\left(\alpha^*,\nu^*,\lambda^*\right)$. But observe that

$$\lambda^* \in \mathrm{argmax}_{\lambda\in\mathbb{R}_+} -\left(\lambda - \min_{q\in Q}\sum_{a\in A}\alpha^*\left(a\right)R\left(p_a^*||q_a\right)/c\right)^2 \implies \lambda^* = \frac{\min_{q\in Q}\sum_{a\in A}\alpha^*\left(a\right)R\left(p_a^*||q_a\right)}{c},$$

$$\alpha^* \in \operatorname{argmax}_{\alpha \in \Delta(A)} U_1\left(\alpha, \nu^*, \lambda^*\right) \implies \alpha^* \in \Delta\left(BR^{\lambda^*}\left(\nu^*\right)\right),$$

and

$$\nu^* \in \operatorname{argmax}_{\nu \in \Delta(Q)} - \int_Q \sum_{a \in A} \alpha^*\left(a\right) R\left(p_a^* || q_a\right) \mathrm{d}\nu\left(q\right) \implies \nu^* \in \Delta\left(Q\left(\alpha^*\right)\right).$$

Therefore, α^* is a mixed c-robust equilibrium sustained by the ν^* and λ^*. $\blacksquare$

Proof of Theorem 2. We start with the preliminary observation that by Lemma 6, $\max_{q \in Q, t \in \mathbb{N}} - \ln \tilde{q}_{\mathbf{a}_t}\left(\mathbf{y}_t\right) \leq K$, $\mathbb{P}_\Pi$-almost surely. This will allow us to invoke Lemma 7.

Let $\Upsilon = \left\{\alpha - \alpha' : \alpha, \alpha' \in \Delta(A)\right\}$ and for all $\varepsilon \in \mathbb{R}_+$ and $\alpha' \in \Delta(A)$, $M_\varepsilon\left(\alpha'\right) = \left\{\nu \in \Delta\left(Q\right) : \int_Q \sum_{a \in A} \alpha'\left(a\right) R\left(p_a^* || q_a\right) \mathrm{d}\nu\left(q\right) \leq \varepsilon + \min_{q \in Q} \sum_{a \in A} \alpha'\left(a\right) R\left(p_a^* || q_a\right)\right\}$. By Esponda et al., 2021a, Part 1a of the proof of Theorem 2,[33] $M_{(\cdot)}\left(\cdot\right)$ is upper hemicontinuous. We define $F : \mathbb{R}_+ \times \mathbb{R}_+ \times \Delta\left(A\right) \rightrightarrows \Upsilon$ by

$$F\left(\varepsilon, \varepsilon', \alpha\right) = \left\{\iota \in \left[\Delta\left(\cup_{\lambda' \in B_{\varepsilon'}\left(\min_{q \in Q} \sum_{a \in A} \alpha(a) R(p_a^* || q_a) / c\right) \cap \left[0, \frac{2K}{c}\right]} BR^{\lambda'}\left(\Delta\left(M_\varepsilon\left(\alpha\right)\right)\right)\right) - \alpha\right]\right\}$$

and $\chi_\alpha = F\left(0, 0, \alpha\right) + \alpha$.

Claim 5. *F and χ_α are upper hemicontinuous, compact-valued, and convex-valued.*

Proof of the Claim. We show that F has a closed graph to conclude that it is upper hemicontinuous. Since Υ is compact, this is enough by, e.g., Proposition E.3 in Ok (2011). The remaining claims are immediateLet

$$\left(\iota_n, \varepsilon_n, \varepsilon_n', \alpha_n\right)_{n \in \mathbb{N}} \in \left(\Upsilon \times \mathbb{R}_+ \times \mathbb{R}_+ \times \Delta\left(A\right)\right)^{\mathbb{N}}$$

be such that $\iota_n \in F\left(\varepsilon_n, \varepsilon_n', \alpha_n\right)$ for all $n \in \mathbb{N}$ and convergent to $\left(\iota, \varepsilon, \varepsilon', \alpha\right)$. Since A is finite, it is without loss of generality (possibly truncating some initial elements of the sequence) to take $\iota_n\left(a\right) > -\alpha_n\left(a\right)$ for all $n \in \mathbb{N}$ and for all a for which $\iota\left(a\right) > -\alpha\left(a\right)$. Then for all $\hat{a}$ such that $\iota\left(\hat{a}\right) > -\alpha\left(\hat{a}\right)$, there is a sequence $\left(\nu_n^{\hat{a}}, \lambda_n^{\hat{a}}\right)_{n \in \mathbb{N}} \in \left(\Delta\left(Q\right) \times \left[0, 2K/c\right]\right)^{\mathbb{N}}$ such that $\nu_n^{\hat{a}} \in M_{\varepsilon_n}\left(\alpha_n'\right)$, $\lambda_n^{\hat{a}} \in B_{\varepsilon_n'}\left(\min_{q \in Q} \sum_{a \in A} \alpha_n\left(a\right) R\left(p_a^* || q_a\right) / c\right)$, and $\hat{a} \in BR^{\lambda_n^{\hat{a}}}\left(\nu_n^{\hat{a}}\right)$. Since $\Delta\left(Q\right)$ and $\left[0, 2K/c\right]$ are compact (by Theorem 15.11 in Aliprantis and Border, 2013) by restricting to a subsequence we can take $\left(\nu_n^{\hat{a}}, \lambda_n^{\hat{a}}\right)_{n \in \mathbb{N}}$ to be convergent to some $\left(\nu^{\hat{a}}, \lambda^{\hat{a}}\right) \in \Delta\left(Q\right) \times \left[0, 2K/c\right]$. Since $M_{(\cdot)}\left(\cdot\right)$ is upper hemicontinuous $\nu^{\hat{a}} \in M_\varepsilon\left(\alpha\right)$. Since by Lemmas

---

[33]Observe that Assumption 1 (i-ii) implies their Assumption 2 (ii-iii), except for the fact that we do not require finite dimensionality of Y and Q. It is readily checked that since they are still assumed to be compact this does not create any issues in the proof of their Theorem 2.

9 and 4 $\alpha \mapsto \min_{q \in Q} \sum_{a \in A} \alpha(a) R(p_a^* || q_a) / c$ is continuous and $Q(\cdot)$, $BR^{(\cdot)}(\cdot)$ are upper hemicontinuous,

$$\lambda^{\hat{a}} \in B_{\varepsilon'} \left(\min_{q \in Q} \sum_{a \in A} \alpha(a) R(p_a^* || q_a) / c \right).$$

and,

$$\hat{a} \in BR^{\lambda^{\hat{a}}}\left(\nu^{\hat{a}}\right) \subseteq \left\{ BR^{\hat{\lambda}}(\hat{\nu}) : \hat{\nu} \in M_\varepsilon(\alpha), \hat{\lambda} \in B_{\varepsilon'} \left(\min_{q \in Q} \sum_{a \in A} \alpha(a) R(p_a^* || q_a) / c \right) \cap \left[0, \frac{2K}{c} \right] \right\}$$

showing that $(\iota, \varepsilon, \varepsilon', \alpha)$ belongs to the graph of the correspondence. $\qquad \square$

Observe that $(\alpha_t)_{t \in \mathbb{N}}$ satisfies the following differential inclusion: for all $a \in A$, $t \in \mathbb{N}$, $h_t \in \mathcal{H}_t$, and $h_{t+1} \in \mathcal{H}_{t+1}$ such that $h_{t+1} \succ h_t$

$$\alpha_{t+1}(h_{t+1})(a) \in \left\{ \alpha_t(h_t)(a) + \frac{1}{t+1} \left(\mathbb{I}_{\{a'\}}(a) - \alpha_t(h_t)(a) \right) : a' \in BR^{\Lambda(h_t)}(\mu(\cdot|h_t)) \right\}.$$

Set $\tau_0 = 0$ and $\tau_t = \sum_{i=1}^{t} \frac{1}{i}$ for all $t \in \mathbb{N}$. The continuous-time interpolation of α_t is the function $w : \mathbb{R}_+ \to \Delta(A)$

$$w(\tau_t + l) = \begin{cases} \alpha_t + l \frac{\alpha_{t+1} - \alpha_t}{\tau_{t+1} - \tau_t}, & \forall t \in \mathbb{N}, \forall l \in \left[0, \frac{1}{t+1} \right] \\ \alpha_1 & t = 0, \forall l \in [0, 1]. \end{cases} \tag{19}$$

We use the theory of stochastic approximation for differential inclusions (Benaim et al., 2005 and Esponda et al., 2021a) to show that (19) can be approximated by a solution to

$$\dot{\underline{\alpha}}_t \in \chi_{\underline{\alpha}_t} - \underline{\alpha}_t. \tag{20}$$

A solution over $[0, T]$, $T \in \mathbb{R}_{++}$, to the differential inclusion (20) with initial point $\hat{\alpha} \in \Delta(A)$ is a mapping $\underline{\alpha}_{(\cdot)} : [0, T] \to \Delta(A)$ that is absolutely continuous over compact intervals such that $\underline{\alpha}_0 = \hat{\alpha}$ and (20) is satisfied for almost every t. Let $S_{\hat{\alpha}}^T$ be the set of the solutions to (20) over $[0, T]$, $T \in \mathbb{R}_{++}$, with initial conditions $\hat{\alpha} \in \Delta(A)$. A solution to (20) exists by Claim 5 and Theorem 2.1.4 in Aubin and Cellina (2012), i.e., $S_{\hat{\alpha}}^T$ is nonempty for every $T \in \mathbb{R}_{++}$ and $\hat{\alpha} \in \Delta(A)$. Let $S^T = \cup_{\hat{\alpha} \in \Delta(A)} S_{\hat{\alpha}}^T$.

Observe that w is Lipschitz continuous of order 1 as for all $(h_t)_{t \in \mathbb{N}} \in \times_{t \in \mathbb{N}} \mathcal{H}_t$,

$$h_{t+1} \succ h_t \quad \forall t \in \mathbb{N} \implies \frac{|| \alpha_{t+1}(h_{t+1}) - \alpha_t(h_t) ||_\infty}{\tau_{t+1} - \tau_t} \leq \frac{1/(t+1)}{1/(t+1)} = 1 \quad \forall t \in \mathbb{N}. \tag{21}$$

Therefore w is absolutely continuous, as for all $n \in \mathbb{N}$ and $\left(l_i, \tilde{l}_i\right)_{i=1}^n$, $\sum_{i=1}^n ||w(l_i) - w\left(\tilde{l}_i\right)||_\infty \leq \sum_{i=1}^n |l_i - \tilde{l}_i|$. Moreover, α_t is uniformly bounded because it takes values in $\Delta(A)$.

By Claim 5, χ_α satisfies Hypothesis 1.1 in Benaim et al. (2005). Moreover, by Theorem 1 in Esponda et al. (2021a) and Lemma 7 we have that $\mathbb{P}_\Pi$-almost surely, if $\lim_{t\to\infty} \alpha_t(\mathbf{h}_t) = \alpha^*$, we eventually have $\mu(\cdot|\mathbf{h}_t) \in M_\varepsilon(\alpha^*)$ and

$$\Lambda(\mathbf{h}_t) \in B_{\varepsilon'}\left(\min_{q \in Q} \sum_{a \in A} \alpha^*(a) R(p_a^*||q_a)/c\right)$$

for all $(\varepsilon, \varepsilon') \in \mathbb{R}_{++}^2$. Thus, there is a sequence $(\hat{\varepsilon}_t)_{t \in \mathbb{N}} \in \mathbb{R}_{++}^\mathbb{N}$ converging to 0 with $\chi_{\alpha_t} - \alpha_t \in F(\hat{\varepsilon}_t, \hat{\varepsilon}_t, \alpha_t)$.

Fix $T \in \mathbb{N}$ and define the flow operator $G: C(\mathbb{R}, \Delta(A)) \times \mathbb{R} \to C(\mathbb{R}, \Delta(A))$ as

$$G^t(f)(s) = f(s+t) \qquad \forall f \in C(\mathbb{R}, \Delta(A)), \forall s \in \mathbb{R}, \forall t \in \mathbb{R}.$$

We now show that every limit point of $(G^t(w))_{t \in \mathbb{N}}$ is in S^T. This argument borrows extensively from the proofs Theorem 4.2 in Benaim et al. (2005) and Theorem 2 in Esponda et al. (2021a). However they cannot be directly applied, because the interpolated process w we consider is not a perturbed solution in the sense of Benaim et al. (2005). Indeed, it may not be possible to find an α that *jointly* justifies a_t as a best reply to beliefs in $Q(\alpha)$ and the concern for misspecification $\min_{q \in Q} \sum_{a \in A} \alpha(a) R(p_a^*||q_a)/c$, as perturbations of the empirical frequency α_{t-1} in different directions may be needed for the concern and the belief. Nevertheless, the core of their arguments can be adapted by leveraging the upper hemicontinuity of F established above.

Since w is uniformly continuous by equation (21), the family $(G^t(w))_{t \in \mathbb{N}}$ is equicontinuous, and thus it is relatively compact in the topology of uniform convergence over compact sets by the Arzela-Ascoli theorem (see Willard, 2012 Theorem 43.15 for the version with a noncompact domain). The topology of uniform convergence over compact sets is metrizable since $\Delta(A)$ is metrizable and $\mathbb{R}$ is open (see Theorem 1.14b in Simon, 2020), and so there exists a limit point $z = \lim_{t_n} G^{t_n}(w)$. Define

$$m(t) = \max\{k \in \mathbb{N}: \tau_k \leq t\}$$

and for all $s \in \mathbb{R}$, $v(s) = w'(s_+) = \alpha_{m(s)+1} - \alpha_{m(s)} \in F(\hat{\varepsilon}_{m(s)}, \hat{\varepsilon}_{m(s)}, \alpha_{m(s)})$, and $v_n(s) =$

56

$v\left(t_n + s\right)$ so

$$z\left(T\right) - z\left(0\right) = \lim_{t_n} \left(G^{t_n}\left(w\right)\left(T\right) - G^{t_n}\left(w\right)\left(0\right)\right)$$

$$= \lim_{t_n}\left(w\left(T + t_n\right) - w\left(t_n\right)\right) = \lim_{n\to\infty}\int_0^T v_n\left(s\right)\mathrm{d}s.$$

Since $(v_n)_{n\in\mathbb{N}}$ is uniformly bounded, it is bounded in $L^2\left([0,T], \mathbb{R}^A, Leb\right)$. By the Banach-Alaoglu theorem (see Theorem 6.21 in Aliprantis and Border, 2013), (by restricting to a subsequence) we can take $(v_n)_{n\in\mathbb{N}}$ to be a weakly-convergent subsequence with limit $v^* \in L^2\left([0,T], \mathbb{R}^A, Leb\right)$. By Mazur's lemma (see Corollary V.3.14 in Dunford and Schwartz (1988)), there exist a function $N : \mathbb{N} \to \mathbb{N}$ and a sequence of positive weights $\left(\rho_n\left(n\right), ..., \rho_{N(n)}\left(n\right)\right)_{n\in\mathbb{N}}$ with $\sum_{i=n}^{N(n)} \rho_i\left(n\right) = 1$ for all $n \in \mathbb{N}$ such that if we define

$$\bar{v}_n = \sum_{i=n}^{N(n)} \rho_i\left(n\right) v_i,$$

then $\bar{v}_n$ converges with respect to the $L^2\left([0,T], \mathbb{R}^A, Leb\right)$ norm, and thus almost surely, to v^*.

Let $\tau \in [0,T]$ be such that $\lim_{n\to\infty}\bar{v}_n\left(\tau\right) = v^*\left(\tau\right)$. For every $t \in [0,T]$ and $n \in \mathbb{N}$, define

$$\gamma_n\left(t\right) = \hat{\varepsilon}_{m(t_n+t)} + \left|\left|w\left(t_n + t\right) - \alpha_{m(t_n+t)}\right|\right|$$

and

$$w_n\left(t\right) = w\left(t_n + t\right).$$

Observe that by definition of w, $(\hat{\varepsilon}_t)_{t\in\mathbb{N}}$, and z,

$$\lim_{n\to\infty}\gamma_n\left(t\right) = 0 \text{ and } \lim_{n\to\infty}w_n\left(t\right) = z\left(t\right).$$

But then, by the upper hemicontinuity of F, for every $\varepsilon \in \mathbb{R}_{++}$ there exists N_ε such that for $n \geq N_\varepsilon$, $F\left(\gamma_n\left(t\right), \gamma_n\left(t\right), w_n\left(t\right)\right) \subseteq B_\varepsilon\left(F\left(0, 0, z\left(t\right)\right)\right)$, where the latter set is closed and convex. But since $v_n\left(t\right) \in B_\varepsilon\left(F\left(0, 0, z\left(t\right)\right)\right)$, for all $n \geq N_\varepsilon$ also $\bar{v}_n\left(t\right) \in B_\varepsilon\left(F\left(0, 0, z\left(t\right)\right)\right)$. Therefore, $v^* \in \left(F\left(0, 0, z\left(\tau\right)\right)\right)$. Since the fact that v_n is weakly convergent to v^* implies by definition that $\lim_{n\to\infty}\int_0^T v_n\left(s\right)\mathrm{d}s = \int_0^T v^*\left(s\right)\mathrm{d}s$ we have that $z \in S^T$.

Therefore, by (ii) $\implies$ (i) of Theorem 4.1 in Benaim et al. (2005) (see Esponda et al.,

for the slightly corrected version used here)

$$\lim_{t \to \infty} \inf_{\tilde{\boldsymbol{\alpha}} \in S^T} \sup_{0 \le s \le T} ||w(t+s) - \tilde{\boldsymbol{\alpha}}_s|| = 0 \quad \mathbb{P}_\Pi\text{-a.s. for all } T \in \mathbb{N}. \tag{22}$$

With this, we can replicate an argument from Fudenberg et al. (2022b) to rule out convergence to nonequilibria. If $\alpha^* \in \Delta(A)$ is not a mixed c-robust equilibrium, there is $a \in A$ with $\alpha^*(a) > 0$ and $\delta_a \notin \chi_{\alpha^*}$. Since $\chi_{(\cdot)}$ has a closed graph and maps into the compact $\Delta(A)$, there exists $D \in \mathbb{R}_{++}$ such that for all $\alpha' \in B_D(\alpha^*)$, $\alpha'(a) - \max_{\hat{\alpha} \in \chi_{\alpha'}} \hat{\alpha}(a) > \alpha^*(a)/2$. Therefore, for every initial condition $\bar{\alpha} \in B_D(\alpha^*)$ and every solution of (20), $\underline{\alpha}(a)$ decreases at rate at least $\alpha^*(a)/4$ until it leaves $B_D(\alpha^*)$. So for every initial condition $\bar{\alpha} \in B_D(\alpha^*)$ and every solution, the differential inclusion leaves $B_D(\alpha^*)$ before time $T^* := 4(D + \alpha^*(a))/\alpha^*(a)$.

With this, we can prove that $(\alpha_t(\mathbf{h}_t))_{t \in \mathbb{N}}$ does not converge to α^* on a sample path on which the convergence of equation (22) happens. Since the set of such sample paths has probability 1 under policy Π, this fact concludes the proof. Suppose by contradiction that on one of such paths $(\alpha_t(h_t))_{t \in \mathbb{N}}$ converges to α^*. Therefore, we can choose $\hat{T} \in \mathbb{N}$ such that on that sample path $\alpha_t(h_t) \in B_{D/2}(\alpha^*)$ for all $t \ge \hat{T}$ and

$$\inf_{\underline{\alpha} \in S^{T^*}} \sup_{0 \le s \le T^*} ||w(\hat{T}+s) - \underline{\alpha}_s|| \le D/4. \tag{23}$$

Take any $\underline{\alpha} \in S^{T^*}$ with $\sup_{0 \le s \le T^*} ||w(\hat{T}+s) - \tilde{\boldsymbol{\alpha}}_s|| \le D/2$. Since $w(\hat{T}) \in B_{D/2}(\alpha^*)$, $\underline{\alpha} \in S_{\bar{\alpha}}^{T^*}$ for some initial condition $\bar{\alpha} \in B_D(\alpha^*)$. But then by definition of T^* the differential inclusion leaves $B_D(\alpha^*)$ by time $T^* + \hat{T}$, and by (23), $(\alpha_t(\mathbf{h}_t))_{t \in \mathbb{N}}$ does not stay in $B_{D/2}(\alpha^*)$, a contradiction. ∎

Proof of Corollary 1. We first show that for a sufficiently low c there is no c-robust equilibrium. Observe that by Assumption 3 (i) and Proposition 8 in Battigalli et al. (2022) for every $\alpha \in \Delta(A)$, we have

$$Q(\alpha) = \left\{ q^{(\theta_0^*, \theta_{1\pi}^*, \theta_{1a}^*, \theta_2^*, \theta_3^*)} \right\}. \tag{24}$$

Moreover, since θ^* perfectly predicts the consequences under policy 0, we have

$$\min_{\theta \in \Theta} R\left(p_0^* || q_0^\theta\right) = 0.$$

By Assumption 3 (i) and Lemma 3 in Battigalli et al. (2022), $BR^{Seu}(\Delta(Q(0))) = \{1\}$,

and therefore 0 is not a c-robust equilibrium for any $c \in \mathbb{R}_{++}$. Since f_1 is strictly concave on $\mathbb{R}_{++}$, by Assumption 3 (iii) it follows that $\min_{\theta \in \Theta} R\left(p_1^* || q_1^\theta\right) = R\left(p_1^* || q_1^{\theta^*}\right) > 0$. By Assumption 3 (ii) and Lemma 8 there exists a sufficiently small $\bar{c}$ such that for all $c \leq \bar{c}$,

$$BR^{\frac{\min_{\theta \in \Theta} R\left(p_1^* || q_1^\theta\right)}{c}} \left(\delta_{\theta^*}\right) = \{0\}$$

proving that there is no c-robust equilibrium if $c \leq \bar{c}$. That a mixed c-robust equilibrium exists follows from Proposition 3.[34]

In particular, the maximal (resp. the minimal) equilibrium is defined as the α such that $\sum_{a \in A} \alpha\left(a\right) R\left(p_a^* || q_a^{\theta^*}\right) / c$ is equal to the maximal (resp. minimal) misspecification concern λ such that $1 \in BR^\lambda\left(\delta_{\theta^*}\right)$ (resp. $0 \in BR^\lambda\left(\delta_{\theta^*}\right)$). Since a larger $\theta_{1\pi}^* + \theta_{1a}^*$ makes action 0 more favorable, the comparative statics follows. ∎

A.2 Representation

A.2.1 Preliminaries

Let $B_0\left(\Sigma\right)$ denote the set of all real-valued Σ-measurable simple functions endowed with the supnorm. The subset of functions in $B_0\left(\Sigma\right)$ that take values in $C \subseteq \mathbb{R}$ is denoted as $B_0\left(\Sigma, C\right)$. A functional $I : \Phi \rightarrow \mathbb{R}$ defined on a nonempty subset Φ of $B_0\left(\Sigma\right)$ is a *niveloid* if for every $\varphi, \psi \in \Phi$, $I\left(\varphi\right) - I\left(\psi\right) \leq \sup\left(\varphi - \psi\right)$. It is *translation invariant* if $I\left(\alpha\varphi + \left(1 - \alpha\right) k \mathbb{I}_S\right) = I\left(\alpha\varphi\right) + \left(1 - \alpha\right) k$ for all $\alpha \in [0, 1]$, $\varphi \in \Phi$, and $k \in \mathbb{R}$ such that $\alpha\varphi + \left(1 - \alpha\right) k \mathbb{I}_S$ and $\alpha\varphi$ are in Φ. A niveloid is *normalized* if $I\left(k\mathbb{I}_S\right) = k$ for all $k \in \mathbb{R}$ such that $k\mathbb{I}_S \in \Phi$. A function $c : \Delta(S) \rightarrow \mathbb{R}_+$ is *grounded* if $c^{-1}(0) \neq \emptyset$. An event is *strongly nonnull* if for every $x, x' \in X$ with $x \succ x'$, we have $x \succ x' E x$.

A.2.2 Results

Our first lemma shows that the average robust control representation falls in the variational class.

Lemma 10. *Suppose that there exist a nonconstant affine function* $u : X \rightarrow \mathbb{R}$, *a*

---

[34]To formally invoke Proposition 3, that requires absolute continuity with respect to the true data generating process for all $\theta \in \Theta$, restrict the parameter space to $\{\theta^*\}$. Given equation (24) every mixed c-robust equilibrium with the reduced parameter space remains a mixed c-robust equilibrium with the original Θ.

nonempty and finite $Q \subseteq \Delta(S)$, $\mu \in \Delta(Q)$, and $(\lambda_q)_{q \in Q} \in \mathbb{R}_+^Q$ such that for all $f, g \in \mathcal{F}$

$$f \succsim g \iff \mathbb{E}_\mu \left[\min_{p \in \Delta(S)} \mathbb{E}_p \left[u(f) \right] + \frac{R(p||q)}{\lambda_q} \right] \geq \mathbb{E}_\mu \left[\min_{p \in \Delta(S)} \mathbb{E}_p \left[u(g) \right] + \frac{R(p||q)}{\lambda_q} \right]. \qquad (25)$$

Then $\succsim$ satisfies Weak Order, Weak Certainty Independence, Continuity, Monotonicity, Uncertainty Aversion, Nondegeneracy, Weak Monotone Continuity, and admits the representation

$$f \succsim g \iff \min_{p \in \Delta(S)} \int_S \hat{u}(f) \, dp + \hat{c}(p) \geq \min_{p \in \Delta(S)} \int_S \hat{u}(g) \, dp + \hat{c}(p) \qquad (26)$$

for some nonconstant affine $\hat{u} : X \to \mathbb{R}$ and a grounded, convex, and lower semicontinuous function $\hat{c} : \Delta(S) \to [0, \infty]$. Moreover, we can choose $\hat{u} = u$ and $\hat{c}$ is such that $\hat{c}^{-1}(0) = \mathbb{E}_\mu[q]$.

Proof. We first observe that without loss of generality we can take u to be such that $0 \in \text{int}\, u(X)$ in the representation of equation (25). Indeed, since u is nonconstant and affine, there exists $x \in X$ with $u(x) \in \text{int}\, u(X)$. Define $u'(y) = u(y) - u(x)$ for all $y \in X$. Then, we have

$$
\begin{aligned}
f \quad \succsim \quad & g \iff \mathbb{E}_\mu \left[\min_{p \in \Delta(S)} \mathbb{E}_p \left[u(f) \right] + \frac{R(p||q)}{\lambda_q} \right] \geq \mathbb{E}_\mu \left[\min_{p \in \Delta(S)} \mathbb{E}_p \left[u(g) \right] + \frac{R(p||q)}{\lambda_q} \right] \\
\iff \quad & \mathbb{E}_\mu \left[\min_{p \in \Delta(S)} \mathbb{E}_p \left[u'(f) \right] + \frac{R(p||q)}{\lambda_q} \right] \geq \mathbb{E}_\mu \left[\min_{p \in \Delta(S)} \mathbb{E}_p \left[u'(g) \right] + \frac{R(p||q)}{\lambda_q} \right]
\end{aligned}
$$

and $0 \in \text{int}\, u'(X)$.

Fix $q \in Q$. The functional $I_q : B_0(\Sigma, \mathbb{R}) \to \mathbb{R}$ defined as

$$I_q(\varphi) : = \min_{p \in \Delta(S)} \int_S \varphi(s) \, dp + \frac{1}{\lambda_q} R(p||q) \qquad \forall \varphi \in B_0(\Sigma, \mathbb{R})$$

is easily seen to be monotone, translation invariant, and concave by Theorem 11.13 in Aliprantis and Border (2013) and the concavity of the minimum. Since Q is finite,

$$\hat{I}(\varphi) : = \int_Q I_q(\varphi) \, d\mu(q) \qquad \forall \varphi \in B_0(\Sigma, \mathbb{R})$$

is well-defined and $\hat{I}$ is monotone, concave, and represents $\succsim$. Let $\varphi \in B_0(\Sigma, u(X))$, $k \in u(X)$, and $\gamma \in (0, 1)$. Since u is affine, X is convex, and $0 \in \text{int}\, u(X)$, we have

$\gamma \varphi + (1 - \gamma) k \in B_0 (\Sigma, u (X))$, $\gamma \varphi \in B_0 (\Sigma, u (X))$, and

$$\hat{I} (\gamma \varphi + (1 - \gamma) k) = \int_Q I_q (\gamma \varphi + (1 - \gamma) k) \, \mathrm{d}\mu (q) = \int_Q I_q (\gamma \varphi) + (1 - \gamma) k \mathrm{d}\mu (q) = \hat{I} (\gamma \varphi) + (1 - \gamma) k.$$

But then, notice that

$$\int_Q \left(\min_{p \in \Delta(S)} \int_S u (f) \, \mathrm{d}p + \frac{1}{\lambda_q} R (p||q) \right) \mathrm{d}\mu (q) = \int_Q I_q (u (f)) \, \mathrm{d}\mu (q) = \hat{I} (u (f))$$

where $\hat{I}$ is monotone and translation invariant. Therefore, by Lemma 25 in Maccheroni et al. (2006a), $\hat{I}$ is a concave niveloid, and it is clearly normalized. With this, by Lemma 28 and Footnote 15 in Maccheroni et al. (2006a) $\succsim$ satisfies Weak Order, Weak Certainty Independence, Continuity, Monotonicity, Uncertainty Aversion, and Nondegeneracy.

Fix $f, g \in \mathcal{F}$, $x \in X$, and $(E_i)_{i \in \mathbb{N}} \in \Sigma^{\mathbb{N}}$ with $E_1 \supseteq E_2 \supseteq ...$, $\cap_{i \geq 1} E_i = \emptyset$, and $f \succ g$. Then, for all $q \in Q$, $\lim_{i \to \infty} q (E_i) = 0$ for all $i \in \mathbb{N}$ and by Proposition 1.4.2 in Dupuis and Ellis (2011)

$$-e^{-\lambda_q \left(I_q \left(x \mathbb{I}_{E_i} + u(f) \mathbb{I}_{S \setminus E_i} \right) \right)} = - \int_{S \setminus E_i} \exp (-\lambda_q u (f (s))) \, \mathrm{d}q (s) - \int_{E_i} \exp (-\lambda_q u (x)) \, \mathrm{d}q (s).$$

But then

$$\lim_{i \to \infty} - \exp \left(-\lambda_q \left(I_q \left(x \mathbb{I}_{E_i} + u (f) \mathbb{I}_{S \setminus E_i} \right) \right) \right) = \int_S -e^{-\lambda_q u(f(s))} \mathrm{d}q (s) > \int_S -e^{-\lambda_q u(g(s))} \mathrm{d}q (s)$$

that is

$$\lim_{i \to \infty} I_q \left(x \mathbb{I}_{E_i} + u (f) \mathbb{I}_{S \setminus E_i} \right) > \frac{- \log \left(\int_S \exp (-\lambda_q u (g (s))) \, \mathrm{d}q (s) \right)}{\lambda_q}$$

proving that there exists $i \in \mathbb{N}$ such that $I_q \left(u (x) \mathbb{I}_{E_i} + u (f) \mathbb{I}_{S \setminus E_i} \right) > I_q (u (g))$. Since the statement holds for every $q \in Q$ and Q is finite, there exists $i \in \mathbb{N}$ such that $\hat{I} \left(u (x) \mathbb{I}_{E_i} + u (f) \mathbb{I}_{S \setminus E_i} \right) > \hat{I} (u (g))$ proving that $\succsim$ satisfies Weak Monotone Continuity. Thus, by Theorem 3 and Lemma 30 in Maccheroni et al. (2006a) it admits the representation in equation (26).

By the first part of the lemma we have

$$u (x) \geq u (x') \iff x \succsim x' \iff \hat{u} (x) \geq \hat{u} (x')$$

and therefore by the uniqueness up to a positive affine transformation of $\hat{u}$ guaranteed by

Corollary 5 in Maccheroni et al. (2006a) and the fact that every two affine functions that represent $\succsim$ on X are positive affine transformations of each other (see, e.g., Theorem 5.11 in Kreps, 1988), we can choose $u = \hat{u}$. Finally, by (ii) $\Longrightarrow$ (iii) of Lemma 32 in Maccheroni et al. (2006a) for every $q \in Q$, and $k \in u(X)$, $\partial I_q(k) = \{q\}$. Let $\bar{k} \in \text{int } u(X) \neq \emptyset$ and observe that since Q is finite,

$$
\begin{aligned}
\lim_{\alpha \downarrow 0} \frac{\hat{I}(\bar{k} + \alpha\varphi) - \hat{I}(\bar{k})}{\alpha} &= \lim_{\alpha \downarrow 0} \frac{\mathbb{E}_\mu\left[I_q(\bar{k} + \alpha\varphi)\right] - \mathbb{E}_\mu\left[I_q(\bar{k})\right]}{\alpha} = \lim_{\alpha \downarrow 0} \mathbb{E}_\mu\left[\frac{I_q(\bar{k} + \alpha\varphi) - I_q(\bar{k})}{\alpha}\right] \\
&= \mathbb{E}_\mu\left[\lim_{\alpha \downarrow 0} \frac{I_q(\bar{k} + \alpha\varphi) - I_q(\bar{k})}{\alpha}\right] = \mathbb{E}_\mu\left[\int_S \varphi \mathrm{d}q\right].
\end{aligned}
$$

Now, applying (iii) $\Longrightarrow$ (ii) of Lemma 32 in Maccheroni et al. (2006a), we obtain that the unique $\hat{c}$ identified by the choice of $\hat{u}$ has $\hat{c}^{-1}(0) = \{\mathbb{E}_\mu[q]\}$. ∎

Lemma 11. *If $E \in \Sigma_{st}$ is strongly nonnull and $\succsim$ satisfies Weak Order, Weak Certainty Independence, Continuity, Monotonicity, Uncertainty Aversion, and Weak Monotone Continuity, then $\succsim_E$ satisfies Weak Order, Weak Certainty Independence, Continuity, Monotonicity, Uncertainty Aversion, Nondegeneracy, and Weak Monotone Continuity.*

Proof. Let $f, g, h \in \mathcal{F}$. By Completeness of $\succsim$ at least one between $fEh \succsim gEh$ and $gEh \succsim fEh$ holds. Therefore, by definition of $\succsim_E$ at least one between $f \succsim_E g$ and $g \succsim_E f$ holds.

Let $f, f', f'' \in \mathcal{F}$, with $f \succsim_E f'$ and $f' \succsim_E f''$. By definition of $\succsim_E$, there exist $h', h'' \in \mathcal{F}$ such that $fEh' \succsim f'Eh'$ and $f'Eh'' \succsim f''Eh''$. Since $E \in \Sigma_{st}$, we have $fEh'' \succsim f'Eh''$. By Transitivity of $\succsim$, $fEh'' \succsim f''Eh''$, and so by definition of $\succsim_E$, $f \succsim_E f''$.

Let $f, g \in \mathcal{F}$, $x, x' \in X$, and $\gamma \in (0, 1)$, be such that $\gamma f + (1 - \gamma) x \succsim_E \gamma g + (1 - \gamma) x$. Since $E \in \Sigma_{st}$, we have $(\gamma f + (1 - \gamma) x) Ex \succsim (\gamma g + (1 - \gamma) x) Ex$. By Weak Certainty Independence of $\succsim$ we get $(\gamma f + (1 - \gamma) x') E (\gamma x + (1 - \gamma) x') \succsim (\gamma g + (1 - \gamma) x') E (\gamma x + (1 - \gamma) x')$. But then by definition of $\succsim_E$, we have $\gamma f + (1 - \gamma) x' \succsim_E \gamma g + (1 - \gamma) x'$, proving that $\succsim_E$ satisfies Weak Certainty Independence.

Let $f, g, h, h' \in \mathcal{F}$. Since $E \in \Sigma_{st}$, we have that

$$
\begin{aligned}
\{\gamma \in [0, 1] : \gamma f + (1 - \gamma) g \succsim_E h\} &= \{\gamma \in [0, 1] : (\gamma f + (1 - \gamma) g) Eh' \succsim hEh'\} \\
&= \{\gamma \in [0, 1] : \gamma fEh' + (1 - \gamma) gEh' \succsim hEh'\}
\end{aligned}
$$

and

$$\{\gamma \in [0,1] : h \succsim_E \gamma f + (1-\gamma)g\} = \{\gamma \in [0,1] : hEh' \succsim (\gamma f + (1-\gamma)g)Eh'\}$$
$$= \{\gamma \in [0,1] : hEh' \succsim \gamma fEh' + (1-\gamma)gEh'\}$$

where the sets on the bottom lines are closed by Continuity of $\succsim$, proving that $\succsim_E$ satisfies Continuity.

Let $f, g, h \in \mathcal{F}$ and $f(s) \succsim_E g(s)$ for all $s \in S$. For every $s \in S$, since E is strongly nonnull, we cannot have $g(s) \succ f(s)$, as otherwise we would have $g(s) \succ f(s) Eg(s)$, a contradiction with $f(s) \succsim_E g(s)$. Then, $fEh \succsim gEh$ by Monotonicity of $\succsim$. Therefore, by definition of $\succsim_E$, $f \succsim_E g$ and so $\succsim_E$ satisfies Monotonicity.

Let $f, g, h \in \mathcal{F}$, $\gamma \in (0,1)$ and $f \sim_E g$. Since $E \in \Sigma_{st}$, $fEh \sim gEh$ and by Uncertainty Aversion, $(\gamma f + (1-\gamma)g)Eh = \gamma fEh + (1-\gamma)gEh \succsim fEh$. By definition of $\succsim_E$, this implies that $\gamma f + (1-\gamma)g \succsim_E f$ and so $\succsim_E$ satisfies Uncertainty Aversion.

Since E is nonnull, there exist $f, g, h \in \mathcal{F}$ such that $fEh \succ gEh$. But then, since $E \in \Sigma_{st}$, there is no $h' \in \mathcal{F}$ with $gEh' \succsim fEh'$. Therefore, by definition of $\succsim_E$, $f \succ_E g$ and $\succsim_E$ satisfies Nondegeneracy.

Let $f, g, h \in \mathcal{F}$, $x \in X$, $(E_i)_{i \in \mathbb{N}} \in \Sigma^{\mathbb{N}}$ with $E_1 \supseteq E_2 \supseteq \ldots$ and $\cap_{n \geq 1} E_n = \emptyset$, and $f \succ_E g$. Since $E \in \Sigma_{st}$, $fEh \succ gEh$. Moreover, $(E_i')_{i \in \mathbb{N}}$ where $E_i' = E_i \cap E$ is such that $E_1' \supseteq E_2' \supseteq \ldots$ and $\cap_{n \geq 1} E_n' \subseteq \cap_{n \geq 1} E_n = \emptyset$. Then $(xE_i'f)Eh = xE_i'(fEh)$ for all $i \in \mathbb{N}$ and by Weak Monotone Continuity and the fact that $fEh \succ gEh$ there exists $n_0 \in \mathbb{N}$ such that $(xE_{n_0}'f)Eh \succ gEh$. But notice that $(xE_{n_0}f)Eh = (xE_{n_0}'f)Eh \succ gEh$ and therefore $xE_{n_0}f \succ_E g$, as $E \in \Sigma_{st}$. ∎

Lemma 12. *Let $\Omega \times \{\rho\} \in \Sigma_{st}$ be strongly nonnull and contain at least three disjoint nonnull events, and suppose $\succsim$ satisfies Weak Order, Weak Certainty Independence, Continuity, Monotonicity, Uncertainty Aversion, Nondegeneracy, Weak Monotone Continuity, the Intramodel Sure-Thing Principle, and Structured Savage. For every $f, g \in \mathcal{F}$, we have*

$$f \succsim_\rho g \iff \min_{q \in \Delta(S)} \mathbb{E}_q[u_\rho(f)] + \frac{R(q||p_\rho)}{\lambda_\rho} \geq \min_{q \in \Delta(S)} \mathbb{E}_q[u_\rho(g)] + \frac{R(q||p_\rho)}{\lambda_\rho} \quad (27)$$

where u_ρ is a nonconstant affine function, $\lambda_\rho \in \mathbb{R}_+$, and $p_\rho \in \Delta(S)$. Moreover, u_ρ can be chosen to be the same for all such ρ and $\operatorname{supp} p_\rho \subseteq \Omega \times \{\rho\}$.

Proof. By Lemma 11 $\succsim_\rho$ satisfies Weak Order, Weak Certainty Independence, Continuity, Monotonicity, Uncertainty Aversion, Nondegeneracy, and Weak Monotone Continuity. We

63

now show that for every $f, g, h, \bar{h} \in \mathcal{F}$ and $E \in \Sigma$, we have

$$fEh \succsim_\rho gEh \implies fE\bar{h} \succsim_\rho gE\bar{h}.$$

Observe that by definition of $\succsim_\rho$, $fEh \succsim_\rho gEh$ implies that there exists $\hat{h} \in \mathcal{F}$ such that $(fEh)\rho\hat{h} \succsim (gEh)\rho\hat{h}$. But then, there exists $h' \in \mathcal{F}$ such that

$$
\begin{aligned}
& (fEh)\rho\hat{h} \succsim (gEh)\rho\hat{h} \\
\implies & (f\{(\omega,\rho) : (\omega,\rho) \in E\}h)\rho\hat{h} \succsim (g\{(\omega,\rho) : (\omega,\rho) \in E\}h)\rho\hat{h} \\
\implies & (f\{(\omega,\rho') : \rho' \in \Delta(\Omega), (\omega,\rho) \in E\}h)\rho\hat{h} \succsim (g\{(\omega,\rho') : \rho' \in \Delta(\Omega), (\omega,\rho) \in E\}h)\rho\hat{h} \\
\implies & (f\{(\omega,\rho') : \rho' \in \Delta(\Omega), (\omega,\rho) \in E\}h) \succsim_\rho (g\{(\omega,\rho') : \rho' \in \Delta(\Omega), (\omega,\rho) \in E\}h) \\
\implies & \left(f\{(\omega,\rho') : \rho' \in \Delta(\Omega), (\omega,\rho) \in E\}\bar{h}\right) \succsim_\rho \left(g\{(\omega,\rho') : \rho' \in \Delta(\Omega), (\omega,\rho) \in E\}\bar{h}\right) \\
\implies & \left(f\{(\omega,\rho') : \rho' \in \Delta(\Omega), (\omega,\rho) \in E\}\bar{h}\right)\rho h' \succsim \left(g\{(\omega,\rho') : \rho' \in \Delta(\Omega), (\omega,\rho) \in E\}\bar{h}\right)\rho h' \\
\implies & \left(f\{(\omega,\rho) : (\omega,\rho) \in E\}\bar{h}\right)\rho h' \succsim \left(g\{(\omega,\rho) : (\omega,\rho) \in E\}\bar{h}\right)\rho h' \\
\implies & \left(fE\bar{h}\right)\rho h' \succsim \left(gE\bar{h}\right)\rho h' \implies fE\bar{h} \succsim_\rho gE\bar{h}
\end{aligned}
$$

where the third, fifth, and eighth implications follow from the definition of $\succsim_\rho$, the fourth implication follows from the Intramodel Sure-Thing Principle, and the other implications only rewrite the acts involved.

Next, observe that if $E \subseteq \Omega \times \{\rho\}$ is nonnull, then there exist $f, g, h \in \mathcal{F}$ with $(fEh)\rho h = fEh \succ gEh = (gEh)\rho h$. By Structured Savage P2, this implies that $fEh \succ_\rho gEh$, so that E is nonnull for the preference $\succsim_\rho$. With this, the first part follows from Theorem 1 in Strzalecki (2011). For the second part, notice that by Theorem 3 and Lemma 30 in Maccheroni et al. (2006a), $\succsim$ admits a variational representation:

$$f \succsim g \iff \min_{p \in \Delta(S)} \left(\int u(f)\,\mathrm{d}p + c(p)\right) \geq \min_{p \in \Delta(S)} \left(\int u(g)\,\mathrm{d}p + c(p)\right) \tag{28}$$

for some nonconstant affine $u : X \to \mathbb{R}$ and a lower semicontinuous and grounded function $c : \Delta(S) \to [0, \infty]$.

Next, notice that $\succsim$ and $\succsim_\rho$ coincide on X. Indeed, let $x \succ x'$. Since $\Omega \times \{\rho\}$ is strongly nonnull $x \succ x'\rho x$ and given that $\Omega \times \{\rho\} \in \Sigma_{st}$ it follows that $x \succ_\rho x'$. Conversely, let $x \succsim x'$, then by equation (28) $u(x) \geq u(x')$. Since c is grounded, there exists $q^* \in \Delta(S)$

with $c\left(q^*\right) = 0$. But then

$$
\begin{aligned}
u\left(x\right) &\geq u\left(x'\right) q^*\left(\Omega \times \{\rho\}\right) + \left(1 - q^*\left(\Omega \times \{\rho\}\right)\right) u\left(x\right) \\
&\geq \min_{q \in \Delta(S)} \left(u\left(x'\right) q\left(\Omega \times \{\rho\}\right) + \left(1 - q\left(\Omega \times \{\rho\}\right)\right) u\left(x\right) + c\left(q\right)\right)
\end{aligned}
$$

that is, $x\left(\Omega \times \{\rho\}\right) x \succsim x'\left(\Omega \times \{\rho\}\right) x$, and $x \succsim_\rho x'$. Therefore, by the uniqueness up to a positive affine transformation of u guaranteed by Corollary 5 in Maccheroni et al. (2006a) and the fact that every two affine functions that represent $\succsim$ on X are positive affine transformations of each other (see, e.g., Theorem 5.11 in Kreps, 1988), we can choose $u = u_\rho$. Suppose by way of contradiction that there exists $E \in \Sigma$ such that $E \cap \left(\Omega \times \{\rho\}\right) = \emptyset$ and $p_\rho\left(E\right) > 0$. Let $x, y \in X$ with $x \succ y$. Then,

$$
u\left(x\right) > u\left(y\right) p_\rho\left(E\right) + u\left(x\right)\left(1 - p_\rho\left(E\right)\right) \geq \min_{q \in \Delta(S)} \int u\left(yEx\right) \mathrm{d}q + \frac{1}{\lambda_\rho} R\left(q \| p_\rho\right)
$$

and so by equation (27), $x \succ_\rho yEx$. But since $x = x\left(\Omega \times \{\rho\}\right) x$, $x = \left(yEx\right)\left(\Omega \times \{\rho\}\right) x$ and $\Omega \times \{\rho\} \in \Sigma_{st}$ this would imply $x \succ x$, a contradiction to the Weak Order of $\succsim$. ∎

Lemma 13. *Suppose that the assumptions of Theorem 3 hold. Let $\succsim$ be such that for all $f, g \in \mathcal{F}$*

$$
f \succsim g \iff \mathbb{E}_\mu \left[\min_{p \in \Delta(S)} \mathbb{E}_p\left[u\left(f\right)\right] + \frac{R\left(p \| q\right)}{\lambda_q} \right] \geq \mathbb{E}_\mu \left[\min_{p \in \Delta(S)} \mathbb{E}_p\left[u\left(g\right)\right] + \frac{R\left(p \| q\right)}{\lambda_q} \right]
$$

where $u : X \to \mathbb{R}$ is a nonconstant affine function, $Q \subseteq \Delta\left(S\right)$ is a finite and nonempty set such that

$$
q\left(\{\omega, \rho_q\}\right) = \rho_q\left(\omega\right) \qquad \forall q \in Q, \forall \omega \in \Omega, \tag{29}
$$

for some $\rho_q \in \Delta\left(\Omega\right)$, $\mu \in \Delta\left(Q\right)$, and $\left(\lambda_q\right)_{q \in Q} \in \mathbb{R}_+^Q$. Then:

1. *For every $\Omega \times B \in \Sigma_s$ and $f, h \in \mathcal{F}$*

$$
\begin{aligned}
&\int_Q \min_{p \in \Delta(S)} \int_S u\left(f_{\Omega \times B} h\right) dp + \frac{R\left(p \| q\right)}{\lambda_q} d\mu\left(q\right) \\
&= \int_{\{q \in Q : \rho_q \in B\}} \min_{p \in \Delta(S)} \mathbb{E}_p\left[u\left(f\right)\right] + \frac{R\left(p \| q\right)}{\lambda_q} d\mu\left(q\right) + \int_{Q \backslash \{q \in Q : \rho_q \in B\}} \min_{p \in \Delta(S)} \mathbb{E}_p\left[u\left(h\right)\right] + \frac{R\left(p \| q\right)}{\lambda_q} d\mu\left(q\right).
\end{aligned}
$$

2. *For every $\Omega \times B \in \Sigma_s$, if $\mu\left(\{q \in Q : \rho_q \in B\}\right) = 0$, then $\Omega \times B$ is null.*

65

Proof. 1) Let $\Omega \times B \in \Sigma_s$ and $f, h \in \mathcal{F}$. We have

$$\int_Q \min_{p \in \Delta(S)} \int_S u\left(f_{\Omega \times B} h\right) \mathrm{d}p + \frac{R\left(p||q\right)}{\lambda_q} \mathrm{d}\mu\left(q\right)$$

$$= \int_{\left\{q \in Q : \rho_q \in B\right\}} \min_{p \in \Delta(S)} \int_S u\left(f_{\Omega \times B} h\right) \mathrm{d}p + \frac{R\left(p||q\right)}{\lambda_q} \mathrm{d}\mu\left(q\right)$$

$$+ \int_{Q \setminus \left\{q \in Q : \rho_q \in B\right\}} \min_{p \in \Delta(S)} \int_S u\left(f_{\Omega \times B} h\right) \mathrm{d}p + \frac{R\left(p||q\right)}{\lambda_q} \mathrm{d}\mu\left(q\right)$$

$$= \int_{\left\{q \in Q : \rho_q \in B\right\}} \min_{p \in \Delta(S) : q \gg p} \int_S u\left(f_{\Omega \times B} h\right) \mathrm{d}p + \frac{R\left(p||q\right)}{\lambda_q} \mathrm{d}\mu\left(q\right)$$

$$+ \int_{Q \setminus \left\{q \in Q : \rho_q \in B\right\}} \min_{p \in \Delta(S) : q \gg p} \int_S u\left(f_{\Omega \times B} h\right) \mathrm{d}p + \frac{R\left(p||q\right)}{\lambda_q} \mathrm{d}\mu\left(q\right)$$

$$= \int_{\left\{q \in Q : \rho_q \in B\right\}} \min_{p \in \Delta(S) : q \gg p} \mathbb{E}_p\left[u\left(f\right)\right] + \frac{R\left(p||q\right)}{\lambda_q} \mathrm{d}\mu\left(q\right)$$

$$+ \int_{Q \setminus \left\{q \in Q : \rho_q \in B\right\}} \min_{p \in \Delta(S) : q \gg p} \mathbb{E}_p\left[u\left(h\right)\right] + \frac{R\left(p||q\right)}{\lambda_q} \mathrm{d}\mu\left(q\right)$$

$$= \int_{\left\{q \in Q : \rho_q \in B\right\}} \min_{p \in \Delta(S)} \mathbb{E}_p\left[u\left(f\right)\right] + \frac{R\left(p||q\right)}{\lambda_q} \mathrm{d}\mu\left(q\right) + \int_{Q \setminus \left\{q \in Q : \rho_q \in B\right\}} \min_{p \in \Delta(S)} \mathbb{E}_p\left[u\left(h\right)\right] + \frac{R\left(p||q\right)}{\lambda_q} \mathrm{d}\mu\left(q\right)$$

where the third equality follows from the fact that by equation (29) $q \gg p$ and $\rho_q \in B$ imply $\operatorname{supp} p \subseteq \operatorname{supp} q \subseteq \Omega \times B$ (and conversely $q \gg p$ and $\rho_q \notin B$ imply $\operatorname{supp} p \subseteq \operatorname{supp} q \subseteq S \setminus (\Omega \times B)$).

2) It follows from 1), since in this case for every $f, g, h \in \mathcal{F}$

$$f_{\Omega \times B} h \succsim g_{\Omega \times B} h$$

$$\iff \int_{\left\{q \in Q : \rho_q \notin B\right\}} \min_{p \in \Delta(S)} \mathbb{E}_p\left[u\left(h\right)\right] + \frac{R\left(p||q\right)}{\lambda_q} \mathrm{d}\mu\left(q\right) \geq \int_{\left\{q \in Q : \rho_q \notin B\right\}} \min_{p \in \Delta(S)} \mathbb{E}_p\left[u\left(h\right)\right] + \frac{R\left(p||q\right)}{\lambda_q} \mathrm{d}\mu\left(q\right)$$

and the RHS is always trivially satisfied as an equality. ∎

Lemma 14. *Suppose that the assumptions of Theorem 3 hold and for all $f, g \in \mathcal{F}$*

$$f \succsim g \iff \mathbb{E}_\mu\left[\min_{p \in \Delta(S)} \mathbb{E}_p\left[u\left(f\right)\right] + \frac{R\left(p||q\right)}{\lambda_q}\right] \geq \mathbb{E}_\mu\left[\min_{p \in \Delta(S)} \mathbb{E}_p\left[u\left(g\right)\right] + \frac{R\left(p||q\right)}{\lambda_q}\right]$$

where $u : X \to \mathbb{R}$ is a nonconstant affine function, $Q \subseteq \Delta\left(S\right)$ is finite, nonempty, with

$$q\left(\left\{\omega, \rho_q\right\}\right) = \rho_q\left(\omega\right) \qquad \forall q \in Q, \forall \omega \in \Omega,$$

66

for some $\rho_q \in \Delta(\Omega)$, $\mu \in \Delta(Q)$, and $(\lambda_q)_{q \in Q} \in \mathbb{R}_+^Q$. Then $\succsim$ satisfies Uniform Misspecification Concern if and only if there exists λ^* with $\lambda_q = \lambda^*$ for all $q \in \mathrm{supp}\mu$.

Proof. (If) Let $\rho, \rho' \in \Delta(\Omega)$, $f, g \in \mathcal{F}$, and $x \in X$ be such that $\Omega \times \{\rho\}$ and $\Omega \times \{\rho'\}$ are nonnull,

$$\rho(\{\omega : f(\omega, \rho) = y\}) = \rho'(\{\omega : g(\omega, \rho') = y\}) \quad \forall y \in X, \tag{30}$$

and $f \succsim_{\Omega \times \{\rho\}} x$. Since $\Omega \times \{\rho\}$ and $\Omega \times \{\rho'\}$ are nonnull, by part 2 of Lemma 13 there exist $q, q' \in Q$ with $\mu(\{q\}) > 0$, $\mu(\{q'\}) > 0$, $\rho_q = \rho$, and $\rho_{q'} = \rho'$. Let $\phi(c) = -\exp(-\lambda^* c)$ for all $c \in u(X)$ and $\xi \in \Delta(X)$ be the finite support probability measure such that for all $y \in X$, $\xi(y) = q(\{(\omega, \rho_q) : f(\omega, \rho_q) = y\})$, then $\int_\Omega \phi(u(f))\mathrm{d}q = \int_X \phi(u(y))\mathrm{d}\xi(y)$. Moreover, equation (30) implies $\int_\Omega \phi(u(g))\mathrm{d}q' = \int_X \phi(u(y))\mathrm{d}\xi(y)$. Therefore, by Lemma 13 both $f \succsim_{\Omega \times \{\rho\}} x$ and $g \succsim_{\Omega \times \{\rho'\}} x$ mean that $\int_X \phi(u(y))\mathrm{d}\xi(y) \geq \phi(u(x))$ proving that $\succsim$ satisfies Uniform Misspecification Concern.

(Only if) Suppose by way of contradiction that there exist $q, q' \subseteq Q$ and $k \in \mathbb{R}_{++}$ with $\mu(\{q\}) > 0$, $\mu(\{q'\}) > 0$, and

$$\lambda_q > k > \lambda_{q'}. \tag{31}$$

Since the state space is adequate there exist $W_q \subseteq \Delta(\Omega)$, $W_{q'} \subseteq \Delta(\Omega)$ and $c \in (0,1)$ with $\rho_q(W_q) = \rho_{q'}(W_{q'}) = c$. Moreover, $q(W_q \times \{\rho_q\}) = c = q'(W_{q'} \times \{\rho_{q'}\})$ and $q(W_{q'} \times \{\rho_{q'}\}) = 0 = q'(W_q \times \{\rho_q\})$. Pick $z, y \in X$ with $z \succ y$. We have that

$$\rho_q(\{\omega : z((W_q \times \{\rho_q\}) \cup (W_{q'} \times \{\rho_{q'}\})) y(\omega, \rho_q) = x\})$$
$$= \rho_{q'}(\{\omega : z((W_q \times \{\rho_q\}) \cup (W_{q'} \times \{\rho_{q'}\})) y(\omega, \rho_{q'}) = x\})$$

for all $x \in X$. By the convexity of X and Lemma 13 there exists $\hat{x} \in X$ with $z \succ \hat{x} \succ y$ and $z((W_q \times \{\rho_q\}) \cup (W_{q'} \times \{\rho_{q'}\})) y \sim_{\rho_{q'}} \hat{x}$. But by equation (31) and Lemma 13 we have $\hat{x} \succ_{\rho_q} z((W_q \times \{\rho_q\}) \cup (W_{q'} \times \{\rho_{q'}\})) y$ a violation of Uniform Misspecification Concern. ∎

Proof of Theorem 3. (Only if) That $\succsim$ satisfies Weak Order, Weak Certainty Independence, Continuity, Monotonicity, Uncertainty Aversion, Nondegeneracy, and Weak Monotone Continuity follows from Lemma 10.

Let $\rho \in \Delta(\Omega)$, $W \subseteq \Omega$, $f, g, h, h' \in \mathcal{F}$, and $fWh \succsim_\rho gWh$. If $\Omega \times \{\rho\}$ is null then we trivially have $fWh' \succsim_\rho gWh'$. Therefore, suppose $\Omega \times \{\rho\}$ is nonnull. By Lemma 13, and since $q \mapsto \rho_q$ is injective, there exists $\bar{q} \in \Delta(S)$ with $\rho_{\bar{q}} = \rho$, $\mu(\{\bar{q}\}) > 0$, and $\min_{p \in \Delta(S)} \int_S u(fWh)\mathrm{d}p + \frac{1}{\lambda}R(p||\bar{q}) \geq \min_{p \in \Delta(S)} \int_S u(gWh)\mathrm{d}p + \frac{1}{\lambda}R(p||\bar{q})$. By Proposition

1.4.2 in Dupuis and Ellis (2011) this is equivalent to $\int_S \phi\left(u\left(fWh\right)\right)\mathrm{d}\bar{q} \geq \int_S \phi\left(u\left(gWh\right)\right)\mathrm{d}\bar{q}$ with $\phi\left(\cdot\right) = -\exp\left(-\lambda\left(\cdot\right)\right)$. This is also equivalent to

$$\int_{W\times\Delta(\Omega)} \phi\left(u\left(f\right)\right)\mathrm{d}\bar{q} + \int_{(\Omega\backslash W)\times\Delta(\Omega)} \phi\left(u\left(h\right)\right)\mathrm{d}\bar{q} \geq \int_{W\times\Delta(\Omega)} \phi\left(u\left(g\right)\right)\mathrm{d}\bar{q} + \int_{(\Omega\backslash W)\times\Delta(\Omega)} \phi\left(u\left(h\right)\right)\mathrm{d}\bar{q}$$

or $\int_{W\times\Delta(\Omega)} \phi\left(u\left(f\right)\right)\mathrm{d}\bar{q} \geq \int_{W\times\Delta(\Omega)} \phi\left(u\left(g\right)\right)\mathrm{d}\bar{q}$. But then, by reversing all the steps with h' in place of h we get $fWh' \succsim_\rho gWh'$ and therefore $\succsim$ satisfies Intramodel Sure-Thing Principle.

Moreover, $\succsim$ satisfies Uniform Misspecification Concern by Lemma 14. That there is a finite set $B \subseteq \Delta\left(\Omega\right)$ such that $\Omega \times \left(\Delta\left(\Omega\right) \setminus B\right)$ is null immediately follows from the representation and part 2 of Lemma 13. Let $\Omega \times B \in \Sigma_s$ and $f,g,h,h' \in \mathcal{F}$. If $\Omega \times B$ is null, we clearly have that $\Omega \times B \in \Sigma_{st}$. Suppose $\Omega \times B$ is nonnull, then

$$f\left(\Omega \times B\right)h \quad \succsim \quad g\left(\Omega \times B\right)h$$

$$\Longleftrightarrow$$

$$\int_{\left\{q\in Q:\rho_q\in B\right\}} \min_{p\in\Delta(S)} \mathbb{E}_p\left[u\left(f\right)\right] + \frac{R\left(p||q\right)}{\lambda_q}\mathrm{d}\mu\left(q\right) \quad \geq \quad \int_{\left\{q\in Q:\rho_q\in B\right\}} \min_{p\in\Delta(S)} \mathbb{E}_p\left[u\left(g\right)\right] + \frac{R\left(p||q\right)}{\lambda_q}\mathrm{d}\mu\left(q\right)$$

$$\Longleftrightarrow$$

$$f\left(\Omega \times B\right)h' \quad \succsim \quad g\left(\Omega \times B\right)h'$$

where the two equivalences follow by Lemma 13. This proves that $\Omega \times B \in \Sigma_{st}$. Since B was chosen to be an arbitrary measurable subset of $\Delta\left(\Omega\right)$, $\Sigma_s \subseteq \Sigma_{st}$, and Structured Savage P2 holds.

That $\succsim$ satisfies Structured Savage P4 and Uncertainty Neutrality over Models immediately follows from Lemma 13 and the representation.

(If) By Structured Savage's P2, $\Sigma_s \subseteq \Sigma_{st}$. Suppose $E \in \Sigma_s$ is nonnull, and let $x, x' \in X$ with $x \succ x'$. Then there exist $f,g,h \in \mathcal{F}$ such that $fEh \succ gEh$. Since f and g are simple acts, they assume finitely many values, and by Weak Order, there exist $\bar{x}, \underline{x} \in X$ with

$$\bar{x} \succsim f\left(s\right), \qquad g\left(s\right) \succsim \underline{x}, \qquad \forall s \in E.$$

Since $E \in \Sigma_s \subseteq \Sigma_{st}$, $fE\bar{x} \succ gE\bar{x}$. By the Monotonicity and Weak Order parts of the Variational Axiom, $\underline{x}\emptyset\bar{x} = \bar{x}E\bar{x} \succsim fE\bar{x} \succ gE\bar{x} \succsim \underline{x}E\bar{x}$. Therefore, by Structured Savage P4, $x = x'\emptyset x \succ x'Ex$. Since $E \in \Sigma_s$ and $x, x' \in X$ were arbitrarily chosen, each nonnull $E \in \Sigma_s$ is also strongly nonnull.

Next, fix a finite $B \subseteq \Delta(\Omega)$, such that for each $\rho \in B$, $\Omega \times \{\rho\}$ is nonnull, and such that $S \setminus \{\Omega \times B\}$ is null. Such a set exists by the Structured Savage axiom, and the cardinality of B is at least 3 by assumption of the theorem. For every $\rho \in B$, by the previous part of the proof $\Omega \times \{\rho\}$ is strongly nonnull and so by Lemma 12 we have

$$f \succsim_\rho g \iff \min_{p \in \Delta(S)} \int_S \hat{u}(f) \, dp + \frac{1}{\lambda_\rho} R(p \| q_\rho) \tag{32}$$

for some $q_\rho \in \Delta(S)$ with support contained in $\Omega \times \{\rho\}$ and a nonconstant affine $\hat{u}$.

Claim 6. *We have* $q_\rho = \rho \times \delta_\rho$.

Proof of the Claim. Since the space is adequate, there exists $v_\rho \in (0,1)$ such that $\rho(\omega) \in \{0, v_\rho\}$. In particular, by applying Uniform Misspecification Concern with $\rho = \rho'$, we obtain that $q_\rho(\omega, \rho) = v_\rho \iff \rho(\omega) = v_\rho$, and the desired conclusion follows. $\square$

Let

$$Q = \{q_\rho \in \Delta(S) : \rho \in B\}. \tag{33}$$

Identify each act $f \in \mathcal{F}$ with the real-valued function $\hat{f} : Q \to \hat{u}(X)$ with

$$\hat{f}(q_\rho) = \min_{p \in \Delta(S)} \int_S \hat{u}(f) \, dp + \frac{1}{\lambda_\rho} R(p \| q_\rho) \qquad \forall \rho \in B$$

where λ_ρ is given by equation (32).

We now show that

$$\hat{f} = \hat{g} \implies f \sim g \qquad \forall f, g \in \mathcal{F}.$$

We partition S in $\left\{ \{\Omega \times \rho\}_{\rho \in B}, S \setminus \{\Omega \times B\} \right\}$ and establish the claim by induction on the number of cells of the partition on which f and g are not identical. Let f and g be such that $\hat{f} = \hat{g}$ and they differ on one element of the partition, say E. Then $f = fEg \sim g$ by definition of $\sim_E$ and Structured Savage P2, so $f \sim g$. For the inductive step, suppose that whenever f and g are such that $\hat{f} = \hat{g}$ and they differ at most on $n \in \mathbb{N}$ elements of the partition, we have $f \sim g$. Let f and g be such that $\hat{f} = \hat{g}$ and they differ on $n + 1 \in \mathbb{N}$ elements of the partition. Let E be an element of the partition on which they differ. Then, fEg and g differ on one element of the partition, and fEg and f differ on n elements of the partition. Therefore, by the inductive hypothesis, we have $g \sim fEg \sim f$.

Moreover, it is immediate to see that $\hat{u}(X)^Q \subseteq \left\{ \hat{f} : f \in \mathcal{F} \right\}$. Therefore, with a slight abuse of notation we let $\succsim$ denote also the binary relation on $\hat{u}(X)^Q$ defined by $\hat{f} \succsim \hat{g}$ if and only if $f \succsim g$.

Claim 7. *For every* $v, v', w, z \in \hat{u}(X)$, $\rho \in B$, *and* $\gamma \in (0, 1)$

$$v_\rho w \gtrsim (\gamma v + (1 - \gamma) v')_\rho z \iff ((1 - \gamma) v + \gamma v')_\rho w \gtrsim v'_\rho z.$$

Proof of the Claim. If $v = v'$ the equivalence is obvious. Suppose without loss of generality that $v' > v$.

1. Let $v_\rho w \gtrsim (\gamma v + (1 - \gamma) v')_\rho z$. This implies that $w > z$. Then, by Continuity, Structured Savage, and the fact that $\Omega \times \{\rho\}$ is strongly nonnull there exists $\alpha \in [0, 1]$ with

$$v_\rho (\alpha w + (1 - \alpha) z) \sim (\gamma v + (1 - \gamma) v')_\rho z.$$

By Uncertainty Neutrality over Models, this implies that $((1 - \gamma) v + \gamma v')_\rho (\alpha w + (1 - \alpha) z) \sim v'_\rho z$. By Monotonicity, this implies that $((1 - \gamma) v + \gamma v')_\rho w \gtrsim v'_\rho z$.

2. Let $((1 - \gamma) v + \gamma v')_\rho w \gtrsim v'_\rho z$. This implies that $w > z$. Then, by Continuity, Structured Savage, and the fact that $\Omega \times \{\rho\}$ is strongly nonnull there exists $\alpha \in [0, 1]$ with $((1 - \gamma) v + \gamma v')_\rho (\alpha w + (1 - \alpha) z) \sim v'_\rho z$. By Uncertainty Neutrality over Models, this implies that $v_\rho (\alpha w + (1 - \alpha) z) \sim (\gamma v + (1 - \gamma) v')_\rho z$. By Monotonicity, this implies that $v_\rho w \gtrsim (\gamma v + (1 - \gamma) v')_\rho z$. $\square$

By the previous claim, Continuity, Structured Savage, and Theorem VII.3.5 in Wakker (2013)[35] there exists $\mu \in \Delta(Q)$ such that for all $\psi, \psi' \in \hat{u}(X)^Q$

$$\psi \gtrsim \psi' \iff \sum_{q \in Q} \psi(q) \mu(q) \geq \sum_{q \in Q} \psi'(q) \mu(q).$$

Moreover, by Observation VII.3.5 in Wakker (2013), μ is uniquely identified.

---

[35]Formally, one needs to apply Theorem VII.3.5 in Wakker (2013) twice. The first application gives

$$\psi \gtrsim \psi' \iff \sum_{q \in Q} U_q(\psi(q)) \geq \sum_{q \in Q} U_q(\psi'(q))$$

for some concave and increasing functions $(U_q : \hat{u}(X) \to \mathbb{R})_{q \in Q}$. The second application is to the preference $\succsim^-$ defined over $(\hat{u}(X))^Q$ by $\psi \succsim^- \psi'$ if and only if $\psi' \gtrsim \psi$ for all $\psi, \psi' \in (\hat{u}(X))^Q$. It gives that for all $\psi, \psi' \in \hat{u}(X)^Q$

$$\psi \gtrsim \psi' \iff \psi' \succsim^- \psi \iff \sum_{q \in Q} U_q^-(\psi'(q)) \geq \sum_{q \in Q} U_q^-(\psi(q)) \iff \sum_{q \in Q} -U_q^-(\psi(q)) \geq \sum_{q \in Q} -U_q^-(\psi'(q))$$

for some decreasing and concave functions $(U_q^- : \hat{u}(X) \to \mathbb{R})_{q \in Q}$. But since $-U_q^-(\cdot)$ is an increasing and convex function, then by Observation VII.3.5 in Wakker (2013) $\psi \gtrsim \psi'$ if and only if $\sum_{q \in Q} U_q^L(\psi(q)) \geq \sum_{q \in Q} U_q^L(\psi'(q))$ for some increasing and linear functions $(U_q^L : \hat{u}(X) \to \mathbb{R})_{q \in Q}$, and the result follows by the Riesz Representation theorem.

But then, by definition of $\succsim$, we obtain that for all $f, g \in \mathcal{F}$

$$f \succsim g \iff \mathbb{E}_\mu \left[\min_{p \in \Delta(S)} \mathbb{E}_p \left[u(f) \right] + \frac{R(p || q_\rho)}{\lambda_\rho} \right] \geq \mathbb{E}_\mu \left[\min_{p \in \Delta(S)} \mathbb{E}_p \left[u(g) \right] + \frac{R(p || q_\rho)}{\lambda_\rho} \right].$$

Moreover, by Lemma 14 Uniform Misspecification Concern implies that $\lambda = \lambda_\rho$ for all $\rho \in B$, proving the result. ∎

Proof of Corollary 2. It immediately follows from Lemma 10 and Proposition 8 in Maccheroni et al. (2006a). ∎

Proposition 7. *Suppose: (i) For every $h \in \mathcal{H}$, $\succsim^h$ satisfies the axioms of Theorem 3 and (ii) $\left(\succsim^h \right)_{h \in \mathcal{H}}$ satisfies Constant Preference Invariance and Dynamic Consistency over Models. Then for every $h \in \mathcal{H}$, $\succsim^h$ admits an average robust control representation $\left(u, Q, \mu(\cdot | h), \lambda_h \right)$ where*

$$q \left(\{ \omega, \rho_q \} \right) = \rho_q(\omega) \qquad \forall q \in Q, \forall \omega \in \Omega.$$

Proof. That each $\succsim^h$ admits a average robust control representation $\left(u_h, Q_h, \mu_h, \lambda_h \right)$ where $q \left(\{ \omega, \rho_q \} \right) = \rho_q(\omega)$ for all $q \in Q_h$ and $\omega \in \Omega$ for some $\rho_q \in \Delta(\Omega)$ follows from (the proof of) Theorem 3. That $u_h = u$ for some constant affine u follows from Constant Preference Invariance.

We now prove that Dynamic Consistency over Models implies $\mu(\cdot | h_t) = \mu_{h_t}$ for all $h_t = (\omega_i)_{i=1}^t \in \mathcal{H}_t$ such that $\prod_{i=1}^t \rho_q(\omega_i) > 0$ for some $q \in Q$. By definition, we have $\mu_{h_t} = \mu$ for the empty history. Let f and g be measurable with respect to Σ_s. Then we can suppress the dependence on ω in $f(\omega, \rho)$ and $g(\omega, \rho)$ and we have that $f \succsim^{h_t} g$ if and only if $f^{h_t} \succsim^\emptyset g^{h_t}$.

But by construction, the latter is equivalent to $\mathbb{E}_\mu \left[\gamma_{f(\rho_q)} \prod_{i=1}^t \rho_q(\omega_i) \left(u(\bar{z}) - u(\underline{z}) \right) \right] \geq \mathbb{E}_\mu \left[\gamma_{g(\rho_q)} \prod_{i=1}^t \rho_q(\omega_i) \left(u(\bar{z}) - u(\underline{z}) \right) \right]$. Dividing both sides by the strictly positive ex-ante probability of history h_t, we obtain

$$\frac{\mathbb{E}_\mu \left[\gamma_{f(\rho_q)} \prod_{i=1}^t \rho_q(\omega_i) \left(u(\bar{z}) - u(\underline{z}) \right) \right]}{\int_{\Delta(\Delta(S))} \prod_{i=1}^t \rho_q(\omega_i) \, \mathrm{d}\mu(q)} \geq \frac{\mathbb{E}_\mu \left[\gamma_{g(\rho_q)} \prod_{i=1}^t \rho_q(\omega_i) \left(u(\bar{z}) - u(\underline{z}) \right) \right]}{\int_{\Delta(\Delta(S))} \prod_{i=1}^t \rho_q(\omega_i) \, \mathrm{d}\mu(q)}.$$

71

But then, by the formula for Bayesian updating, this is equivalent to

$$\int_{\Delta(\Delta(S))} \gamma_{f(\rho_q)} \left(u\left(\bar{z}\right) - u\left(\underline{z}\right) \right) d\mu\left(q|h_t\right) \geq \int_{\Delta(\Delta(S))} \gamma_{g(\rho_q)} \left(u\left(\bar{z}\right) - u\left(\underline{z}\right) \right) d\mu\left(q|h_t\right)$$

that is $\int_{\Delta(\Delta(S))} u\left(f\left(\rho_q\right)\right) d\mu\left(q|h_t\right) \geq \int_{\Delta(\Delta(S))} u\left(g\left(\rho_q\right)\right) d\mu\left(q|h_t\right)$. Stated differently, $\succsim^{h_t}$ admits an SEU representation of the acts measurable with respect to Σ_s with Bernoulli utility u and probability measure $\mu\left(\cdot|h_t\right)$. Since for the histories $h_t = (\omega_i)_{i=1}^t \in \mathcal{H}_t$ where $\prod_{i=1}^t \rho_q\left(\omega_i\right) = 0$ for all $q \in Q$ Bayesian updating does not impose any restriction, the result follows. ∎

Proof of Proposition 4. By Proposition 7, $\succsim^h$ admits an average robust control representation $(u, Q, \mu\left(\cdot|h\right), \lambda_h)$ for every $h \in \mathcal{H}$. Observe that since the outcome frequency is constant along the sequence $(h_{t_n})_{n \in \mathbb{N}}$, by Lemma 2, $\frac{LLR(h_{t_n}, Q)}{t_n} = 1/c$ for some $c \in \mathbb{R}_{++}$ and for all $n \in \mathbb{N}$. Suppose by way of contradiction that

$$l := \liminf_{n \to \infty} c\lambda^{h_{t_n}} = \liminf_{n \to \infty} \frac{\lambda_{h_{t_n}}}{\frac{LLR(h_{t_n}, Q)}{t_n}} < \limsup_{n \to \infty} \frac{\lambda_{h_{t_n}}}{\frac{LLR(h_{t_n}, Q)}{t_n}} = \limsup_{n \to \infty} c\lambda_{h_{t_n}} =: L.$$

Let $\bar{q} \in Q$ be such that $\Omega \times \{\rho_{\bar{q}}\}$ is nonnull and so that $\bar{q} \in \min_{q \in Q} R\left(p^{h_{t_1}}||q\right)$. Since $\Omega \times \{\rho_{\bar{q}}\}$ contains at least three nonnull events, by Lemma 13, there is $E \subseteq W$ and $r \in (0, 1)$ with $\rho_{\bar{q}}(E) = r$. Let $x, z \in X$, $\gamma^*, \gamma_* \in (0, 1)$, and $\lambda^*, \lambda_* \in \left(\frac{l}{c}, \frac{L}{c}\right)$ be such that $x \succ^\emptyset z$, $\lambda^* > \lambda_*$,

$$\frac{-\mu\left(\bar{q}\right) \log\left(r \exp\left(-\lambda^*\left(u\left(z\right)\right)\right) + (1-r) \exp\left(-\lambda^*\left(u\left(x\right)\right)\right)\right)}{\mu\left(\min_{q \in Q} R\left(p^{h_{t_1}}||q\right)\right) \lambda^*} + \left(1 - \frac{\mu\left(\bar{q}\right)}{\mu\left(\min_{q \in Q} R\left(p^{h_{t_1}}||q\right)\right)}\right) u\left(z\right)$$
$$= u\left(\gamma^* x + (1 - \gamma^*) z\right),$$

and

$$\frac{-\mu\left(\bar{q}\right) \log\left(r \exp\left(-\lambda_*\left(u\left(z\right)\right)\right) + (1-r) \exp\left(-\lambda_*\left(u\left(x\right)\right)\right)\right)}{\mu\left(\min_{q \in Q} R\left(p^{h_{t_1}}||q\right)\right) \lambda_*} + \left(1 - \frac{\mu\left(\bar{q}\right)}{\mu\left(\min_{q \in Q} R\left(p^{h_{t_1}}||q\right)\right)}\right) u\left(z\right)$$
$$= u\left(\gamma_* x + (1 - \gamma_*) z\right),$$

where the existence of such γ_*, γ^* is guaranteed by u being affine. Moreover, it is easy to see that $\gamma_* > \gamma^*$. Consider a subsequence $(n_m)_{m \in \mathbb{N}}$ such that

$$\lim_{m \to \infty} c\lambda^{h_{t_{n_m}}} = l.$$

Moreover, let $M \in \mathbb{N}$ be such that for all $m \geq M$

$$\lambda^{h_{t_{nm}}} < \frac{\lambda_* + \frac{l}{c}}{2}.$$

Similarly, let $(n_{\tilde{m}})_{\tilde{m} \in \mathbb{N}}$ such that

$$\lim_{\tilde{m} \to \infty} c\lambda^{h_{t_{n\tilde{m}}}} = L.$$

Moreover, let $\tilde{M} \in \mathbb{N}$ be such that for all $\tilde{m} \geq \tilde{M}$

$$\lambda^{h_{t_{n\tilde{m}}}} > \frac{\lambda^* + \frac{L}{c}}{2}.$$

With this, by Proposition 7 and Proposition 1.4.2 in Dupuis and Ellis (2011) we have that for all $m \geq M$ and $\tilde{m} \geq \tilde{M}$

$$\boldsymbol{\gamma}^{x\left(E \times \{\rho_{\bar{q}}\}\right)z}_{\succsim^{h_{t_{nm}}}} > \gamma_* \text{ and } \boldsymbol{\gamma}^{x\left(E \times \{\rho_{\bar{q}}\}\right)z}_{\succsim^{h_{t_{n\tilde{m}}}}} < \gamma^*.$$

But this in turn implies that $\succsim^{h_{t_{nm}}}$ is never $\left(x, y, \left(E \times \{\rho_{\bar{q}}\}\right), (\gamma_* - \gamma^*)\right)$-similar to $\succsim_{h_{t_{n\tilde{m}}}}$ for

$$\min\{m, \tilde{m}\} \geq \max\left\{M, \tilde{M}\right\},$$

a contradiction. This shows that either $\lambda_{h_{t_n}}$ converges or it diverges to plus infinity. The last part of the statement immediately by taking E in the first part of the proof to be equal to the one whose existence is asserted in the statement, and by the construction of γ_* and γ^* above. ∎

Proof of Proposition 5. By Proposition 7 we know that each $\succsim^h$ admits an average robust control representation $(u, Q, \mu(\cdot|h), \lambda_h)$. Without loss of generality suppose that $\mu(\{q\}|\emptyset) > 0$ for all $q \in Q$. Let $(h_{t_n})_{n \in \mathbb{N}} \in \times_{n \in \mathbb{N}}\Omega^{t_n}$ be a sequence of histories with empirical frequency $\hat{\rho} \notin \{\rho_q : q \in Q\}$. Observe that since the outcome frequency is constant along the sequence $(h_{t_n})_{n \in \mathbb{N}}$, by Lemma 2, $\frac{LLR(h_{t_n}, Q)}{t_n} = c$ for some $c \in \mathbb{R}_{++}$ and for all $n \in \mathbb{N}$. Suppose by way of contradiction that

$$L: = \limsup_{n \to \infty} \frac{LLR(h_{t_n}, Q)}{\lambda_{h_{t_n}} t_n} > 0. \tag{34}$$

Since the state space is adequate there exist $k \in (0, 1)$ and $(W_q)_{q \in Q} \in \left(2^\Omega\right)^Q$ such that

$\rho_q (W_q) = k$ for all $q \in Q$. Define $E = \cup_{q \in Q} \left(W_q \times \{\rho_q\} \right)$. Let $x, y \in X$ with $x \succ^{\emptyset} y$ and choose also $z \in X$ such that $x \succ^{\emptyset} z \succ^{\emptyset} y$ and

$$- \exp \left(-2 \frac{c}{L} u(z) \right) = -k \exp \left(-2 \frac{c}{L} u(x) \right) - (1-k) \exp \left(-2 \frac{c}{L} u(y) \right)$$

where the existence of such z is guaranteed by u being affine and X being convex. Let $f \in \mathcal{F}$ be defined as $f = xEy$. But then equation (34) implies that for infinitely many $n \in \mathbb{N}$

$$f \succ_{\rho}^{h_{t_n}} z$$

a contradiction with Asymptotic Concern for every $\rho \in \{\rho_q : q \in Q\}$. ∎

Proof of Proposition 6. By Proposition 7 we know that each $\succsim^h$ admits an average robust control representation $(u, Q, \mu(\cdot|h), \lambda_h)$ where

$$q \left(\{\omega, \rho_q\} \right) = \rho_q (\omega) \qquad \forall q \in Q, \forall \omega \in \Omega,$$

for some $\rho_q \in \Delta(\Omega)$. Without loss of generality suppose that $\mu(\{q\}|\emptyset) > 0$ for all $q \in Q$. Consider a sequence of histories $(h_{t_n})_{n \in \mathbb{N}}$ with outcome frequency constant and not in $\{\rho_q : q \in Q\}$. Observe that since the outcome frequency is constant along the sequence $(h_{t_n})_{n \in \mathbb{N}}$, by Lemma 2, $\frac{LLR(h_{t_n}, Q)}{t_n} = c$ for some $c \in \mathbb{R}_{++}$ and for all $n \in \mathbb{N}$. Suppose by way of contradiction that

$$L := \liminf_{n \to \infty} \frac{LLR(h_{t_n}, Q)}{\lambda_{h_{t_n}} t_n} \in \mathbb{R}. \tag{35}$$

As the state space is adequate there exist $k \in (0,1)$ and $(W_q)_{q \in Q} \in \left(2^{\Omega} \right)^Q$ such that for every $q \in Q$, $\rho_q (W_q) = k$. Let $x, z \in X$ and $\gamma \in (0,1)$ be such that $x \succ^{\emptyset} z$ and

$$- \ln \left(k \exp \left(\frac{-cu(x)}{2 \max\{1, L\}} \right) + (1-k) \exp \left(\frac{-cu(z)}{2 \max\{1, L\}} \right) \right) = \frac{c \left(u \left(\gamma x + (1-\gamma) z \right) \right)}{2 \max\{1, L\}},$$

where the existence of such γ is guaranteed by u being affine, and $\gamma < k$. Define $E = \cup_{q \in Q} \left(W_q \times \{\rho_q\} \right)$. Consider a subsequence $(n_m)_{m \in \mathbb{N}}$ such that

$$\lim_{m \to \infty} \frac{LLR \left(h_{t_{n_m}}, Q \right)}{\lambda_{h_{t_{n_m}}} t_{n_m}} = L.$$

Moreover, let M be such that for all $m \geq M$

$$\lambda^{h_{t_{n_m}}}/c > \frac{1}{2\max\{1, L\}}.$$

With this, by Proposition 7 if we let $\geq$ be the subjective utility preference with utility index u and belief $\int_Q p \, d\mu(q)$, we have that for all $m \geq M$

$$\gamma^{xEz}_{\succsim^{h_{t_{n_m}}}} < \gamma \text{ and } \gamma^{xEz}_{\geq^{h_{t_{n_m}}}} = k.$$

By Corollary 2, this contradicts Asymptotic Leniency as then $\succsim^{h_{t_{n_m}}}$ and $\geq^{h_{t_{n_m}}}$ are not $(x, y, E, k - \gamma)$-similar for $m \geq M$. ∎

A.3 Computations supporting

A.3.1 Example 2

Observe that, compared Esponda and Pouzo (2016), we are adding (arbitrarily small) noise $l\varepsilon_1$ to the true tax schedule, fixing a problem in their original example. Indeed, without this modification, the relative entropy between the true and conjectured distribution is infinity for every model. We have

$$
\begin{aligned}
R\left(p_a^* \| q_a^\theta\right) &= \text{const.} + \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{\exp\left(\frac{(t - \tau(a + \omega_a))^2}{-2l^2}\right)}{-l\sqrt{2\pi}} \log\left(\frac{\exp\left(\frac{\left(\frac{t}{a+\omega_a} - \theta\right)^2}{-2}\right)}{\sqrt{2\pi}}\right) dt \, dp_a^*(\omega_a) \\
&= \text{const.} + \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{\frac{1}{l\sqrt{2\pi}} \exp\left(\frac{-\varepsilon_1^2}{2l^2}\right)\left(\frac{\tau(a+\omega_a)+\varepsilon_1}{a+\omega_a} - \theta\right)^2}{2} d\varepsilon_1 \, dp_a^*(\omega_a) \\
&= \text{const.} + \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{\left(\frac{1}{l\sqrt{2\pi}} \exp\left(\frac{-\varepsilon_1^2}{2l^2}\right)\left(\theta^2 - \frac{2\left(\tau(a+\omega_a)+(l\varepsilon_1)^2\right)\theta}{a+\omega_a}\right)\right)}{2} d\varepsilon_1 \, dp_a^*(\omega_a)
\end{aligned}
$$

taking the FOC, we get

$$\theta = \mathbb{E}_{p_a^*}\left[\frac{\tau(a + \omega_a) + l^2}{a + \omega_a}\right]$$

and so $Q(a) \sim \left\{q^{\mathbb{E}_{p_a^*}\left[\frac{\tau(a+\omega_a)}{a+\omega_a}\right]}\right\}$ for small l. The condition for not switching from an action a to an action a' with $a \geq a'$ in a Berk-Nash equilibrium in which the belief is concentrated

on θ is

$$
\begin{aligned}
\mathbb{E}_{\omega_a,\varepsilon_{2,a}}\left[(a-a')(1-\theta-\varepsilon_2)\right] &= \mathbb{E}_{\omega_a,\varepsilon_{2,a}}\left[(a+\omega_a)(1-\theta-\varepsilon_2)\right] - \mathbb{E}_{\omega_{a'},\varepsilon_{2,a}}\left[(a'+\omega_{a'})(1-\theta-\varepsilon_2)\right] \\
&\geq c(a) - c(a').
\end{aligned}
$$

By Proposition 1.4.2 in Dupuis and Ellis (2011), the condition for not switching from an action a to an action a' with $a \geq a'$ in a k-robust equilibrium in which the belief is concentrated on θ is

$$
\begin{aligned}
&\frac{-k}{R\left(p_a^*||q_a^\theta\right)} \log \mathbb{E}_{\omega_a,\omega_{a'},\varepsilon_{2,a}}\left[\exp\left(\frac{R\left(p_a^*||q_a^\theta\right)\left[(a'+\omega_{a'})(1-\theta-\varepsilon_{2,a})-(a+\omega_a)(1-\theta-\varepsilon_{2,a})\right]}{k}\right)\right] \\
&= \frac{-k}{R\left(p_a^*||q_a^\theta\right)} \log\left(\frac{\mathbb{E}_{\omega_a,\varepsilon_{2,a}}\left[\exp\left(-\frac{R\left(p_a^*||q_a^\theta\right)(a+\omega_a)(1-\theta-\varepsilon_{2,a})}{k}\right)\right]}{\mathbb{E}_{\omega_{a'},\varepsilon_{2,a}}\left[\exp\left(-\frac{R\left(p_a^*||q_a^\theta\right)(a'+\omega_{a'})(1-\theta-\varepsilon_{2,a})}{k}\right)\right]}\right) \\
&= \frac{-k}{R\left(p_a^*||q_a^\theta\right)} \log \mathbb{E}_{\omega_a,\varepsilon_{2,a}}\left[\exp\left(-\frac{R\left(p_a^*||q_a^\theta\right)(a+\omega_a)(1-\theta-\varepsilon_{2,a})}{k}\right)\right] \\
&\quad +\frac{k}{R\left(p_a^*||q_a^\theta\right)} \log \mathbb{E}_{\omega_{a'},\varepsilon_{2,a}}\left[\exp\left(-\frac{R\left(p_a^*||q_a^\theta\right)(a'+\omega_{a'})(1-\theta-\varepsilon_{2,a})}{k}\right)\right] \\
&\geq c(a) - c(a').
\end{aligned}
$$

Since $\mathbb{E}_{\omega_a,\varepsilon_{2,a}}\left[|(a+\omega_a)(1-\theta-\varepsilon_{2,a})|\right]$ and $\mathbb{E}_{\omega_{a'},\varepsilon_{2,a}}\left[|(a'+\omega_{a'})(1-\theta-\varepsilon_{2,a})|\right]$ are finite, by Jensen inequality (see 10.2.6 in Dudley, 2018 for the version that applies here) the LHS is lower in the second case, and we obtain the desired conclusion.

A.3.2   Example 3

The condition for not switching from action 0 to an action a with $a \geq 0$ in a Berk-Nash equilibrium is

$$
p_a^*(s \leq a)\left(\mathbb{E}_{p_a^*}(v) - a\right) \leq 0.
$$

By Proposition 1.4.2 in Dupuis and Ellis (2011), the condition for not switching from action 0 to an action a with $a \geq 0$ in a k-robust equilibrium is

$$
-\frac{k}{R\left(p_a^*||q_a\right)} \log \int_{\mathbb{R}} \int_{\mathbb{R}} \exp\left(-\frac{R\left(p_a^*||q_a\right)}{k}[v-a]\,\mathbb{I}_{[0,a]}(s)\right) \mathrm{d}p_a^*(s)\,\mathrm{d}p_a^*(v) \leq 0.
$$

Since $\mathbb{E}_{p_a^*}(|v|) < \infty$, by Jensen inequality (see 10.2.6 in Dudley, 2018 for the version that applies here) the LHS is lower in the second case, and we obtain the desired conclusion.

A.3.3 Example of Dynamic Inconsistency

Example 4. *To provide a simple illustration of dynamic inconsistency, we consider the two-period truncated problem. An urn contains black (b) or green (g) balls. At each time t, the DM is asked to bet 1 dollar on the color of the ball drawn from the urn or to opt out (o) and observe the drawn with a payoff of 0.6. That is, $u(a, y) = \mathbb{I}_{\{y\}}(a)$ if $a \in \{b, g\}$ and $u(o, y) = 0.6$. Suppose that at period 0, the level of concern for misspecification is $\Lambda(h_0) = 0$ and that the agent considers two models, q, q', that assign respectively probability 0.7 and 0.3 to drawing a black ball, independently of the agent action. The prior μ assigns equal probability to these two models.*

To illustrate the possibility of dynamic inconsistencies of a forward-looking agent, we introduce a discount factor equal to $\delta = 0.9$ and suppose that $\Lambda((0, b)) = 2$. In this case, at time 0, the decision maker would like to commit to the following plan: opt out in the first period and then, in the second period, bet on the color of the ball drawn in the first period. However, the increase in concern for misspecification created by the observation of the black drawn makes this plan not feasible: at history $(0, b)$, the agent will opt out again.

References

Aliprantis, Charambolos and Kim Border (2013). *Infinite Dimensional Analysis: A Hitchhiker's Guide.* Springer-Verlag.

Anderson, Evan W, Lars Peter Hansen, and Thomas J Sargent (2003). "A quartet of semigroups for model specification, robustness, prices of risk, and model detection". *Journal of the European Economic Association* 1, pp. 68–123.

Anderson, Robert M, Haosui Duanmu, Aniruddha Ghosh, and M Ali Khan (2022). "On Existence of Berk-Nash Equilibria in Misspecified Markov Decision Processes with Infinite Spaces". *arXiv preprint arXiv:2206.08437.*

Angeletos, George-Marios, Zhen Huo, and Karthik A Sastry (2021). "Imperfect macroeconomic expectations: Evidence and theory". *NBER Macroeconomics Annual* 35, pp. 1–86.

Arrow, KJ and JR Green (1973). "Notes on Expectations Equilibria in Bayesian Settings". Working Paper No. 33, Stanford University.

Aubin, J-P and Arrigo Cellina (2012). *Differential inclusions: set-valued maps and viability theory.* Vol. 264. Springer Science & Business Media.

Ba, Cuimin (2022). "Robust Model Misspecification and Paradigm Shifts". *arXiv preprint arXiv:2106.12727.*

Banerjee, Abhijit V, Sylvain Chassang, Sergio Montero, and Erik Snowberg (2020). "A theory of experimenters: Robustness, randomization, and balance". *American Economic Review* 110, pp. 1206–30.

Barillas, Francisco, Lars Peter Hansen, and Thomas J Sargent (2009). "Doubts or variability?" *journal of economic theory* 144, pp. 2388–2418.

Battigalli, Pierpaolo (1987). "Comportamento razionale ed equilibrio nei giochi e nelle situazioni sociali". *Unpublished undergraduate dissertation, Bocconi University, Milano.*

Battigalli, Pierpaolo, Simone Cerreia-Vioglio, Fabio Maccheroni, Massimo Marinacci, and Thomas Sargent (2022). "A framework for the analysis of self-confirming policies". *Theory and Decision*, pp. 1–58.

Battigalli, Pierpaolo, Alejandro Francetich, Giacomo Lanzani, and Massimo Marinacci (2019). "Learning and self-confirming long-run biases". *Journal of Economic Theory* 183, pp. 740–785.

Benaim, Michel, Josef Hofbauer, and Sylvain Sorin (2005). "Stochastic approximations and differential inclusions". *SIAM Journal on Control and Optimization* 44, pp. 328–348.

Berk, Robert H (1966). "Limiting behavior of posterior distributions when the model is incorrect". *The Annals of Mathematical Statistics*, pp. 51–58.

Bohren, J Aislinn and Daniel N Hauser (2021). "Learning with heterogeneous misspecified models: Characterization and robustness". *Econometrica* 89, pp. 3025–3077.

Cabrales, Antonio, Olivier Gossner, and Roberto Serrano (2013). "Entropy and the value of information for investors". *American Economic Review* 103, pp. 360–77.

Casella, George and Roger L Berger (2021). *Statistical inference.* Cengage Learning.

Cerreia-Vioglio, Simone, Lars Peter Hansen, Fabio Maccheroni, and Massimo Marinacci (2022). "Making decisions under model misspecification".

Cerreia-Vioglio, Simone, Fabio Maccheroni, Massimo Marinacci, and Luigi Montrucchio (2013a). "Ambiguity and robust statistics". *Journal of Economic Theory* 148, pp. 974–1049.

— (2013b). "Classical subjective expected utility". *Proceedings of the National Academy of Sciences* 110, pp. 6754–6759.

Chew, Hong Soo and Jacob S Sagi (2006). "Event exchangeability: Probabilistic sophistication without continuity or monotonicity". *Econometrica* 74, pp. 771–786.

Cho, In-Koo and Kenneth Kasa (2015). "Learning and model validation". *The Review of Economic Studies* 82, pp. 45–82.

Cho, In-Koo and Jonathan Libgober (2022). "Learning Underspecified Models". *arXiv preprint arXiv:2207.10140.*

Clarida, Richard, Jordi Gali, and Mark Gertler (2000). "Monetary policy rules and macroeconomic stability: evidence and some theory". *The Quarterly journal of economics* 115, pp. 147–180.

Cogley, Timothy and Thomas J Sargent (2005). "The conquest of US inflation: Learning and robustness to model uncertainty". *Review of Economic dynamics* 8, pp. 528–563.

De Filippis, Roberta, Antonio Guarino, Philippe Jehiel, and Toru Kitagawa (2022). "Non-Bayesian updating in a social learning experiment". *Journal of Economic Theory* 199, p. 105188.

Dean, Mark and Pietro Ortoleva (2017). "Allais, Ellsberg, and preferences for hedging". *Theoretical Economics* 12, pp. 377–424.

— (2019). "The empirical relationship between nonstandard economic behaviors". *Proceedings of the National Academy of Sciences* 116, pp. 16262–16267.

Dekel, Eddie, Jeffrey C Ely, and Okan Yilankaya (2007). "Evolution of preferences". *The Review of Economic Studies* 74, pp. 685–704.

Denti, Tommaso and Luciano Pomatto (2022). "Model and predictive uncertainty: A foundation for smooth ambiguity preferences". *Econometrica* 90, pp. 551–584.

Diamond, Peter A (1982). "Aggregate demand management in search equilibrium". *Journal of political Economy* 90, pp. 881–894.

Dillenberger, David, Juan Sebastián Lleras, Philipp Sadowski, and Norio Takeoka (2014). "A theory of subjective learning". *Journal of Economic Theory* 153, pp. 287–312.

Dudley, Richard M (2018). *Real Analysis and Probability.* Chapman and Hall/CRC.

Dunford, Nelson and Jacob T Schwartz (1988). *Linear operators, part 1: general theory.* Vol. 10. John Wiley & Sons.

Dupuis, Paul and Richard S Ellis (2011). *A weak convergence approach to the theory of large deviations.* Vol. 902. John Wiley & Sons.

Enke, Benjamin and Florian Zimmermann (2019). "Correlation neglect in belief formation". *The Review of Economic Studies* 86, pp. 313–332.

Epstein, Larry G and Shaolin Ji (2022). "Optimal learning under robustness and time-consistency". *Operations Research* 70, pp. 1317–1329.

Epstein, Larry G and Martin Schneider (2007). "Learning under ambiguity". *The Review of Economic Studies* 74, pp. 1275–1303.

Esponda, Ignacio (2008). "Behavioral equilibrium in economies with adverse selection". *American Economic Review* 98, pp. 1269–91.

Esponda, Ignacio and Demian Pouzo (2016). "Berk–Nash equilibrium: A Framework for Modeling Agents with Misspecified Models". *Econometrica* 84, pp. 1093–1130.

Esponda, Ignacio, Demian Pouzo, and Yuichi Yamamoto (2021a). "Asymptotic behavior of Bayesian learners with misspecified models". *Journal of Economic Theory* 195, p. 105260.

— (2021b). "Corrigendum: Asymptotic behavior of Bayesian learners with misspecified models". *Journal of Economic Theory* 195, p. 105260.

Esponda, Ignacio, Emanuel Vespa, and Sevgi Yuksel (2022). *Mental models and learning: The case of base-rate neglect.* Tech. rep.

Fishburn, Peter C (1970). *Utility theory for decision making.* Wiley.

Foutz, Robert V and RC Srivastava (1977). "The performance of the likelihood ratio test when the model is incorrect". *The annals of Statistics*, pp. 1183–1194.

Frick, Mira, Ryota Iijima, and Yuhta Ishii (2023). "Belief Convergence under Misspecified Learning: A Martingale Approach". *Review of the Economic Studies* 90, pp. 781–814.

Fudenberg, Drew and David M Kreps (1993). "Learning mixed equilibria". *Games and economic behavior* 5, pp. 320–367.

Fudenberg, Drew and Giacomo Lanzani (2023). "Which misperceptions persist?" *Theoretical Economics, Forthcoming.*

Fudenberg, Drew, Giacomo Lanzani, and Philipp Strack (2021). "Limit points of endogenous misspecified learning". *Econometrica* 89, pp. 1065–1098.

— (2022a). "Pathwise Concentration Bounds for Misspecified Bayesian Beliefs". *Available at SSRN 3805083.*

— (2022b). "Selective Memory Equilibrium". *Available at SSRN 4015313.*

Fudenberg, Drew and David K Levine (1993). "Self-confirming equilibrium". *Econometrica*, pp. 523–545.

— (1995). "Consistency and cautious fictitious play". *Journal of Economic Dynamics and Control* 19, pp. 1065–1089.

Gagnon-Bartsch, Tristan, Matthew Rabin, and Joshua Schwartzstein (2022). "Channeled Attention and Stable Errors".

Ghirardato, Paolo and Massimo Marinacci (2002). "Ambiguity made precise: A comparative foundation". *Journal of Economic Theory* 102, pp. 251–289.

Giacomini, Raffaella, Vasiliki Skreta, and Javier Turen (2020). "Heterogeneity, inattention, and bayesian updates". *American Economic Journal: Macroeconomics* 12, pp. 282–309.

Gilboa, Itzhak, Stefania Minardi, and Larry Samuelson (2020). "Theories and cases in decisions under uncertainty". *Games and Economic Behavior* 123, pp. 22–40.

Gilboa, Itzhak and David Schmeidler (1989). "Maxmin expected utility with non-unique prior". *Journal of Mathematical Economics* 18, pp. 141–153.

Gul, Faruk and Wolfgang Pesendorfer (2014). "Expected uncertain utility theory". *Econometrica* 82, pp. 1–39.

Hall, Peter and Christopher C Heyde (2014). *Martingale limit theory and its application.* Academic press.

Hansen, Lars Peter and Thomas J Sargent (2001). "Robust control and model uncertainty". *American Economic Review* 91, pp. 60–66.

— (2007). "Recursive robust estimation and control without commitment". *Journal of Economic Theory* 136, pp. 1–27.

— (2011). "Robustness". *Robustness.* Princeton university press.

— (2022). "Risk, Ambiguity, and Misspecification: Decision Theory, Robust Control, and Statistics".

Hansen, Lars Peter, Thomas J Sargent, Gauhar Turmuhambetova, and Noah Williams (2006). "Robust control and model misspecification". *Journal of Economic Theory* 128, pp. 45–90.

Hart, Sergiu and Andreu Mas-Colell (2000). "A simple adaptive procedure leading to correlated equilibrium". *Econometrica* 68, pp. 1127–1150.

Hausman, Jerry A (1978). "Specification tests in econometrics". *Econometrica: Journal of the econometric society*, pp. 1251–1271.

He, Kevin and Jonathan Libgober (2022). "Evolutionarily stable (mis) specifications: Theory and applications".

Kallenberg, Olav (1973). "Random measures". *Random Measures.* De Gruyter.

Karni, Edi and Marie-Louise Vierø (2013). ""Reverse Bayesianism": A choice-based theory of growing awareness". *American Economic Review* 103, pp. 2790–2810.

Klibanoff, Peter, Massimo Marinacci, and Sujoy Mukerji (2005). "A smooth model of decision making under ambiguity". *Econometrica* 73, pp. 1849–1892.

— (2009). "Recursive smooth ambiguity preferences". *Journal of Economic Theory* 144, pp. 930–976.

Kreps, David (1988). *Notes On The Theory Of Choice.* Westview Press.

Liebman, Jeffrey B and Richard J Zeckhauser (2004). "Schmeduling".

Liese, Friedrich and Igor Vajda (1987). *Convex statistical distances.* Vol. 95. Teubner.

Maccheroni, Fabio, Massimo Marinacci, and Aldo Rustichini (2006a). "Ambiguity aversion, robustness, and the variational representation of preferences". *Econometrica* 74, pp. 1447–1498.

— (2006b). "Dynamic variational preferences". *Journal of Economic Theory* 128, pp. 4–44.

Maenhout, Pascal J (2004). "Robust portfolio rules and asset pricing". *Review of financial studies* 17, pp. 951–983.

Mu, Xiaosheng, Luciano Pomatto, Philipp Strack, and Omer Tamuz (2021). "Monotone additive statistics". *arXiv preprint arXiv:2102.00618.*

Nyarko, Yaw (1991). "Learning in Mis-specified Models and the Possibility of Cycles". *Journal of Economic Theory* 55, pp. 416–427.

Ok, Efe A (2011). "Real analysis with economic applications". *Real Analysis with Economic Applications.* Princeton University Press.

Ortoleva, Pietro (2012). "Modeling the change of paradigm: Non-Bayesian reactions to unexpected news". *American Economic Review* 102, pp. 2410–36.

Parthasarathy, Kalyanapuram R (2005). *Probability Measures on Metric Spaces.* American Mathematical Soc.

Pathak, Parag (2002). "Notes on robust portfolio choice". *unpublished paper, Harvard University* 3, p. 14.

Primiceri, Giorgio E (2005). "Time varying structural vector autoregressions and monetary policy". *The Review of Economic Studies* 72, pp. 821–852.

Rabin, Matthew (2002). "Inference by believers in the law of small numbers". *The Quarterly Journal of Economics* 117, pp. 775–816.

Rees-Jones, Alex and Dmitry Taubinsky (2020). "Measuring "schmeduling"". *The Review of Economic Studies* 87, pp. 2399–2438.

Reny, Philip J (1999). "On the existence of pure and mixed strategy Nash equilibria in discontinuous games". *Econometrica* 67, pp. 1029–1056.

Robatto, Roberto and Balázs Szentes (2017). "On the biological foundation of risk preferences". *Journal of Economic Theory* 172, pp. 410–422.

Royden, Halsey Lawrence and Patrick Fitzpatrick (1988). *Real analysis.* Vol. 32. Macmillan New York.

Sargent, Thomas J (1999). "The conquest of American inflation". *The Conquest of American Inflation.* Princeton University Press.

— (2008). "Evolution and intelligent design". *American Economic Review* 98, pp. 5–37.

Savage, Leonard J (1954). *The foundations of statistics.* Courier Corporation.

Schwartzstein, Joshua and Adi Sunderam (2021). "Using models to persuade". *American Economic Review* 111, pp. 276–323.

Serfozo, Richard (1982). "Convergence of Lebesgue integrals with varying measures". *Sankhyā: The Indian Journal of Statistics, Series A*, pp. 380–402.

Shmaya, Eran and Leeat Yariv (2016). "Experiments on decisions under uncertainty: A theoretical framework". *American Economic Review* 106, pp. 1775–1801.

Simon, Jacques (2020). *Continuous Functions.* John Wiley & Sons.

Sims, Christopher A and Tao Zha (2006). "Were there regime switches in US monetary policy?" *American Economic Review* 96, pp. 54–81.

Siniscalchi, Marciano (2011). "Dynamic choice under ambiguity". *Theoretical Economics* 6, pp. 379–421.

Snowberg, Erik and Leeat Yariv (2021). "Testing the waters: Behavior across participant pools". *American Economic Review* 111, pp. 687–719.

Sobel, Joel (1984). "Non-linear prices and price-taking behavior". *Journal of Economic Behavior & Organization* 5, pp. 387–396.

Spiegler, Ran (2020). "Can agents with causal misperceptions be systematically fooled?" *Journal of the European Economic Association* 18, pp. 583–617.

Strzalecki, Tomasz (2011). "Axiomatic foundations of multiplier preferences". *Econometrica* 79, pp. 47–73.

Tversky, Amos and Daniel Kahneman (1971). "Belief in the law of small numbers." *Psychological bulletin* 76, p. 105.

Vuong, Quang H (1989). "Likelihood ratio tests for model selection and non-nested hypotheses". *Econometrica*, pp. 307–333.

Wakker, Peter P (2013). *Additive representations of preferences: A new foundation of decision analysis.* Vol. 4. Springer Science & Business Media.

Willard, Stephen (2012). *General topology.* Courier Corporation.