Robust Misspecified Models and Paradigm Shifts

Cuimin Ba*

December 13, 2023

– Click here to see the most recent version –

Abstract

This paper studies which misspecified models are likely to persist when individuals also entertain alternative models. The main result provides a characterization of such models based on two features that are straightforward to derive from the primitives: the model's asymptotic accuracy in predicting the equilibrium pattern of observed outcomes and the 'tightness' of the prior around such equilibria. Misspecified models can be robust in that they persist against a wide range of competing models—including the correct model—despite individuals observing an infinite amount of data. Moreover, simple misspecified models with entrenched priors can be more robust than correctly specified models.

*University of Pittsburgh. Email: bacuimin@gmail.com. I am deeply indebted to George Mailath, Aislinn Bohren, and Kevin He for their guidance and support at every stage of this paper. I thank Nageeb Ali, Ben Brooks, Hanming Fang, Joel Flynn, Mira Frick, Drew Fudenberg, Yuan Gao, Alice Gindin, Marina Halac, Takuma Habu, Daniel Hauser, Ju Hu, Yuhta Ishii, Nawaaz Khalfan, Botond Köszegi, Changhwa Lee, Jonathan Libgober, Xiao Lin, Ce Liu, Steven Matthews, Guillermo Ordoñez, Pietro Ortoleva, Wolfgang Pesendorfer, Andrew Postlewaite, Alvaro Sandroni, Pedro Solti, Juuso Toikka, Marcus Tomaino, Rakesh Vohra, Xi Weng, Ece Yegane and conference audiences for helpful comments and suggestions.

1 Introduction

People use models to guide decision making, but the subjective nature of models suggests that model misspecification can be pervasive. Misspecification can stem from the need to simplify the complex world as well as from behavioral biases such as overconfidence or correlation neglect. To explore how misspecified models impact beliefs and actions, the growing literature on misspecified learning focuses on the case of a dogmatic agent who uses a particular misspecified model and never considers changing this model.[1] While this simplifies the environment in a way that yields tractable characterizations of long-run beliefs, it leaves open the question of whether it is realistic to expect a decision-maker to never abandon a wrong model.

A plethora of evidence suggests that people often switch models when an alternative seems more compelling. For example, scientists adopt a new paradigm if it fits the observable data significantly better in terms of accuracy and simplicity (i.e., Kuhn's (1962) theory of paradigm shifts). One classic example is the paradigm shift from the Ptolemaic model to the Copernican model in astronomy. Likewise, economists adopt new models when evidence comes to light that important economic forces are missing from old models. People also alter their subjective assumptions about the world in daily life, such as changing thinking patterns in cognitive behavioral therapy or overcoming implicit biases through introspection (Di Stefano, Gino, Pisano, and Staats, 2015). People are influenced by and attracted to different political narratives as they receive more information (Fisher, 1985; Braungart and Braungart, 1986).

If individuals consider switching to competing models, which (if any) misspecified models should we expect to persist and when? Answering these questions is essential for understanding the enduring implications of model misspecification in the long term and for devising policies to tackle persistent model misspecification. This paper proposes a novel learning framework to address these questions. In this framework, an agent uses models to learn an unknown fixed data-generating process (henceforth

[1] Examples include: a monopolist trying to estimate the slope of the demand function when the true slope lies outside of the support of his prior (Nyarko, 1991; Fudenberg, Romanyuk, and Strack, 2017); agents learning from private signals and other individuals' actions while neglecting the correlation between the observed actions (Eyster and Rabin, 2010; Ortoleva and Snowberg, 2015; Bohren, 2016) or overestimating how similar others' preferences are to their own (Gagnon-Bartsch, 2017); overconfident agents falsely attributing low outcomes to an adverse environment (Heidhues, Kőszegi, and Strack, 2018, 2019; Ba and Gindin, 2022); a decision-maker imposing false causal interpretations on observed correlations (Spiegler, 2016, 2019, 2020); a gambler who flips a fair coin mistakenly believing that future tosses must exhibit systematic reversal (Rabin and Vayanos, 2010; He, 2022); individuals narrowly focusing their attention on only a few aspects rather than a complete state space (Mailath and Samuelson, 2020).

DGP) that governs the relationship between her action choices and random outcomes. Each model is a parametric theory of how actions may affect the outcome distribution. Formally, it consists of a collection of DGPs that are considered possible, each indexed by a distinct parameter value. For example, consider a monopolist who chooses production quantities based on a linear model of consumer demand. Here, for each pair of parameter values—fixing the slope and intercept of the demand curve—the model prescribes a mapping from production quantities to distributions of consumer demand. Such a model is *misspecified* if the true DGP is not included in the predicted mappings. While the *dogmatic modeler* typically considered in the misspecified learning literature uses the same model throughout, in my framework the agent is a *switcher* who subscribes to one model in any period but can switch between multiple models. For the main analysis, this agent starts with an initial model and entertains exactly one competing model. She has a prior over the parameters within each model, updates beliefs as she observes realized outcomes, and plays the optimal action based on the current model given the updated posterior. To decide whether to switch to the competing model, the agent keeps track of the *Bayes factor*—the likelihood ratio of the competing model relative to the initial model given the observed data—and switches if it exceeds a fixed switching threshold. She switches back to the initial model if the Bayes factor drops below the inverse of the threshold. As the switching threshold increases, model switching requires more evidence and becomes stickier.

One might question why, given the agent is already considering and comparing multiple models, she does not perform Bayesian updating over the models and aggregate their predictions. That is, instead of switching between models, the agent could potentially form a "hypermodel" that encompasses all DGPs from these models and then act like a standard Bayesian agent. First, it is important to note that the model-switching framework allows for the nesting of models. In fact, the initial model and the competing model may each consist of a group of smaller models. However, as pointed out in Savage's (1972) *Foundations of Statistics*, Bayesianism is a reasonable description of human behavior only when decision-makers focus on "modest little worlds."[2] It is unrealistic to expect them to formulate and act upon "a model of everything" because of the cognitive demand it imposes. Furthermore, models may be built upon fundamentally conflicting ideas, such as a geocentric model versus a heliocentric model, or a liberal

[2] Savage (1972, p. 16) describes it as "utterly ridiculous" to demand that "one envisage every conceivable policy for the government of his whole life (at least from now on) in its most minute details, in the light of the vast number of unknown states of the world, and decide here and now on one policy."

worldview versus a conservative worldview, which makes it hard for an individual to employ them in decision making simultaneously. Given these considerations, it seems plausible that people start with an initial model, which encompasses all sub-models they are comfortable employing simultaneously, and expand this model or switch to a different one when necessary.

Within this framework, a model *persists* against a given competing model if, with positive probability, the agent eventually stops switching and sticks to this model forever. A natural question is which models are *globally robust* in that, fixing a prior over model parameters, they persist against *every* possible competing model, regardless of its predictions and the associated prior over the competing model's parameters. It is not immediately clear whether this robustness notion permits *any* form of model misspecification; however, if a misspecified model turns out to be globally robust, this is a compelling argument for why we should expect it to sustain over time. Nevertheless, global robustness is undoubtedly a strong requirement since it places no restrictions on the competing model but not every model is equally likely to arise in competition. When the agent is conservative or has limited knowledge about the environment, she may only entertain competing models that are close to her current model rather than those representing completely different paradigms. This idea gives rise to another natural notion—*local robustness*—which requires a model to persist against every local perturbation with similar predictions and similar priors. These notions provide a language to compare the robustness properties of models across different environments, and their formalization is a central conceptual contribution of my framework.

The main results of the paper characterize both robustness notions based on two properties that are easily derived from the primitives of a model, *asymptotic accuracy* and *prior tightness*, as summarized in Table 1. A model has perfect asymptotic accuracy if it gives rise to a *self-confirming equilibrium* that satisfies a stability condition that I refer to as *p-absorbingness*. In a self-confirming equilibrium, the agent plays actions against a *consistent* belief over model parameters such that the model prediction perfectly coincides with the objective outcome distribution (Battigalli, 1987; Fudenberg and Levine, 1993). The stability condition requires that, with positive probability, a dogmatic modeler who only uses this model eventually only plays actions in the support of the equilibrium. With perfect asymptotic accuracy, the model has weakly higher explanatory power than any other competing model in the limit with positive probability. However, this alone does not imply persistence, because the learning dynamics may induce the agent to switch away before her belief moves sufficiently close to the equilibrium belief. If, in addition, the prior is tight in the sense of being concentrated

	Notions of robustness	
Properties	global	local
asymptotic accuracy	perfect	perfect
prior tightness	yes	no

Table 1: Summary of results.

around the set of p-absorbing self-confirming equilibria, the explanatory power of the model remains consistently high across all periods.

I first characterize *which* models can be locally or globally robust under at least one (full-support) prior. Theorem 1 establishes that a model is globally robust for at least one prior if and only if it is locally robust for at least one prior, and both amount to a requirement for perfect asymptotic accuracy. Hence, while we may conjecture local robustness to be much weaker than global robustness, these two notions characterize the exact same set of models. Importantly, Theorem 1 holds for all levels of switching stickiness as long as the switching threshold is strictly larger than 1. Even if the agent is extremely reluctant to switch, the set of robust models does not expand, because the accumulation of evidence over time eventually leads to the abandonment of a less accurate model.

Moving forward, I explore *when*, or under which priors, models exhibiting perfect asymptotic accuracy are locally or globally robust. Theorem 2 highlights the real distinction between global and local robustness: the former requires prior tightness but the latter does not. I provide a closed-form quantification of the required level of tightness in terms of the switching threshold. Specifically, the prior probability assigned to the parameters involved in the p-absorbing self-confirming equilibria must exceed the *inverse* of the switching threshold. As the threshold decreases to 1, the agent must start with a prior fully concentrated on the p-absorbing self-confirming equilibria to ensure that the initial model is globally robust. Therefore, higher switching stickiness facilitates the persistence of model misspecification not by broadening the set of robust misspecified models, but by enabling asymptotically accurate misspecified models to persist under a more extensive range of priors.

My characterization provides a formal learning foundation for the persistence of asymptotically accurate misspecified models.[3] Such misspecified models can be glob-

[3]Note that, however, perfect asymptotic accuracy does not equate to high efficiency because strictly suboptimal actions can be played in a self-confirming equilibrium if the model yields wrong predictions off-path.

ally robust and persist against any *arbitrary* competing model—including the true model—despite the agent having an infinite amount of data. Moreover, the results provide off-the-shelf tools to predict which underlying biases are more relevant in specific contexts. Applying these tools in different applications can be useful in generating testable predictions for empirical research and suggesting behavioral policies to mitigate the consequences of misspecification. To illustrate this, in Section 5.1 I apply my results to a workhorse model in the misspecified learning literature where the agent has a wrong perception of a payoff-relevant fundamental and learns about another fundamental (Heidhues et al., 2018; Ba and Gindin, 2022; Murooka and Yamamoto, 2023). I show that the asymptotic accuracy of misspecified models is closely linked to whether they induce positively or negatively reinforcing belief dynamics, the direction of which can be determined by examining how beliefs about the different fundamentals affect the optimal choice of action. In a leading example, I show that overconfidence in one's ability is globally robust, while underconfidence is not locally robust for a wide range of parameters. This suggests that underconfidence requires less intervention than overconfidence as it can be self-correcting.

The characterization also provides fresh insights into how qualitative features of the model and the learning environment contribute to persistence. First, an interesting contrast emerges between the robustness properties of misspecified and correctly specified models. On one hand, all correctly specified models have perfect asymptotic accuracy while a subset of misspecified models can achieve this. On the other hand, correct specification does not imply prior tightness, but the latter property can be easily achieved by a misspecified model with a small parameter space or one that yields a large number of self-confirming equilibria. In combination, these observations convey an intriguing negative message: certain misspecified models can prove more robust than correctly specified models, precisely because they are sufficiently extreme and misleading. Second, lower switching stickiness can be a double-edged sword, since it makes global robustness harder to attain for any model, whether correctly or incorrectly specified. As the agent adopts a lower switching threshold, it becomes easier to switch away from a misspecified model. However, it also becomes likely to abandon a correctly specified model due to small pieces of noisy information and get trapped with a misspecified alternative. In Section 5.2, I apply these insights to a media consumption problem, demonstrating that a simplistic binary model of the world can outperform a correct yet more flexible model in terms of robustness properties and even replace the correct model as the prevailing worldview when the switching threshold is sufficiently low. Such model misspecification can result in enduring polarization of political beliefs.

Related Literature. This paper contributes to the growing literature on learning with misspecified models. Much of the literature mentioned in Footnote 1 focuses on case-by-case analyses of misspecified models when the decision-maker holds on to a particular model. This paper provides a microfoundation for the persistence of certain types of misspecified models. Another strand of this literature studies equilibrium concepts to characterize the decision-maker's steady state behavior, among which the self-confirming equilibrium is the most relevant for this paper (Battigalli, 1987; Fudenberg and Levine, 1993).[4] The idea that a decision-maker can be trapped in a self-confirming equilibrium and fail to realize his misperception has been floating around in the literature, for which this paper provides a formal proof.[5] Crucially, my formal treatment of the model-switching dynamics yields new insights into the impact of environmental factors, such as the switching stickiness and the range of competing models, on model persistence—insights that go beyond what a simple equilibrium analysis can reveal.

Recent developments in the literature focus on characterizing asymptotic beliefs and actions in general environments (Bohren and Hauser, 2021; Frick, Iijima, and Ishii, 2023). This paper faces many of the same technical challenges as the works in this area since model persistence partly hinges on the asymptotic behavior of a dogmatic modeler. Esponda, Pouzo, and Yamamoto (2021) find conditions for a single agent's action frequency to converge to a Berk-Nash equilibrium using tools from stochastic approximation. Fudenberg, Lanzani, and Strack (2021) establish that a uniformly strict Berk-Nash equilibrium is uniformly stable in the sense that starting from any prior that is sufficiently close to the equilibrium belief, the dogmatic modeler's action converges to the equilibrium with arbitrarily high probability. In this paper, I show that robustness is related to p-absorbingness—a different stability notion that does not require the dogmatic modeler's action to converge, but her action to enter and eventually stay within the support of an equilibrium. The main technical contribution of this paper is to integrate model switching into active learning. Given that the agent considers multiple models, we need to keep track of multiple belief processes, all of which are generated by endogenous data. Since the Bayes factor that governs the model switching process interacts and correlates with all belief processes even when

[4] Esponda and Pouzo (2016) propose the concept of Berk-Nash equilibrium, generalizing the self-confirming equilibrium by relaxing the requirement that the subjective prediction fully coincides with the objective reality. Other related concepts include the analogy-based expectation equilibrium in Jehiel and Koessler (2008) and the cursed equilibrium in Eyster and Rabin (2005). As pointed out by Esponda and Pouzo (2016), these two solution concepts coincide with Berk-Nash or self-confirming equilibrium under appropriately specified feedback structures.

[5] See for example Sargent (1999).

there is no switch, the characterization of the agent's behavior requires new techniques.

This paper is part of a body of research that explores why misspecified models persist. Gagnon-Bartsch, Rabin, and Schwartzstein (2020) study the stability of models when the agent entertains a correctly specified alternative model. They focus on a setting where data is exogenous but the agent only pays attention to the data they deem decision-relevant given the current model. This contrasts with my framework where data is endogenously generated but the agent pays attention to all available data. Cho and Kasa (2015) study model switching with endogenously generated data in a continuous setting. They restrict attention to models that induce a unique globally stable self-confirming equilibrium and characterize dominant models based on the asymptotic rate of parameter drift leading to an escape from the equilibrium. In contrast, my results provide insights into the nuanced role of initial conditions in determining model persistence.[6] Montiel Olea, Ortoleva, Pai, and Prat (2022) characterize the "winner" model in a contest environment where agents use models to predict an exogenous data-generating process and make auction bids based on their subjective model prediction error. They identify a trade-off between model fit and model estimation uncertainty when the dataset is small. My paper complements their finding by showing that a similar trade-off between asymptotic accuracy and prior tightness exists in a model-switching framework even with an infinite amount of data.

A set of papers approach this problem from an evolutionary or payoff perspective. Fudenberg, Lanzani et al. (2022) study the evolutionary dynamics when a small share of a large population mutates to enlarge their models at a Berk-Nash equilibrium. They find that an equilibrium can resist mutations that yield a better statistical fit but induce worse-performing actions. Similar to this paper, they show that a self-confirming equilibrium resists every mutation.[7] He and Libgober (2020) also consider individuals who evaluate competing models based on their expected objective payoffs but examine multi-agent strategic games where misspecification can lead to beneficial misinferences. Frick, Iijima, and Ishii (2021) study welfare comparisons of learning biases and find that some biases can outperform Bayesian updating because they may lead to correct learning at a faster speed. In addition to comparing objective payoffs,

[6]The difference in results mainly stems from the different switching rules we consider. Cho and Kasa (2015) employ a Lagrange Multiplier (LM) test in model selection, which is calculated only based on the maximum likelihood estimation of the initial model, while the Bayes factor is sensitive to the prior beliefs and is used to formalize the notion that "it takes a model to beat a model".

[7]The underlying mechanisms of our results are different. Their result follows from the assumption that every mutation is an expansion of the original model. Since a self-confirming equilibrium must remain an equilibrium under the mutated model, it is possible for all individuals in the population to stick to the same behavior as before the mutation.

individuals might also choose to adopt models that promise the highest future payoffs (Eliaz and Spiegler, 2020; Levy, Razin, and Young, 2022).

An extensive literature in decision theory studies the behavior of a decision-maker who has access to multiple models or priors over states. A number of canonical decision criteria capture aversion to model uncertainty, which is absent from my framework since the agent maximizes the expected utility based on her current model (Gilboa and Schmeidler, 1989; Klibanoff, Marinacci, and Mukerji, 2005; Hansen and Sargent, 2001). Ortoleva (2012) proposes and axiomatically characterizes an amendment to Bayes' rule, called the Hypothesis Testing model, where the agent switches to a better alternative prior (if exists) upon observing an event to which she assigns a probability below a threshold. This contrasts with my framework where the agent switches if the ratio of the probability of the observed outcomes under the current model *relative to* the competing model is sufficiently low. Karni and Vierø (2013) provide a choice-based decision theory to model a self-correcting agent who can expand his universe of subjective states.

This paper is connected to a literature that considers the notion that individuals can shift models based on statistical fit and explores its ramifications across various contexts. As individuals tend to adopt better-fitting models based on observed information, this creates opportunities for a self-interested sender to manipulate the signal structure or introduce alternative models (Galperti, 2019; Schwartzstein and Sunderam, 2021; Aina, 2023). My results suggest that people could be misled into believing in a misspecified model and act suboptimally even when they have infinite data. Finally, this paper is also related to the statistics literature on model selection.[8] My paper focuses on the Bayes factor rule and differs with the statistical literature by studying an endogenous data-generating process.

The rest of the paper is organized as follows. Section 2 provides an illustrative example. Section 3 introduces the learning framework. Section 4 presents the main results and Section 5 develops two applications. Section 6 discusses several extensions and Section 7 concludes. Appendix A contains useful auxiliary results, Appendix B includes proofs of the main results, and Appendix C contains omitted examples.

[8]Statisticians have developed a number of criteria that differ in their cost of computation and penalty for overfitting, such as the Bayes factor, Akaike information criterion (AIC), Bayesian information criterion (BIC), and likelihood-ratio test (LR test), and the machine learning community favors cross-validation due to its flexibility and ease of use (Chernoff, 1954; Akaike, 1974; Stone, 1977; Schwarz et al., 1978; Kass and Raftery, 1995; Konishi and Kitagawa, 2008).

2 Illustrative Example

As a simple illustration of the learning framework and the main results, consider the following example. An artist chooses how much effort to exert in creating artwork in every period, $a_t \in \{0, 1, 2\}$, for $t = 0, 1, 2, \ldots$. Upon exerting effort, he incurs a cost $a_t(a_t+0.5)$ and obtains revenue from the sales of his work. The sales revenue is given by $y_t = (a_t + b)\omega + \epsilon_t$, where $b \in \mathbb{R}$ captures how talented the artist is, $\omega \in \mathbb{R}$ captures an unknown market demand for arts, and ϵ_t is a random noise with a known distribution. Suppose the artist's true talent is $b^* = 1$ and the true market demand is $\omega^* = 2$.

The artist holds a non-degenerate prior belief about the market demand and hopes to learn about it by repeatedly exerting effort and observing the realized sales. If the artist knows his true talent, he will be able to correctly infer the market demand from the sales data, allowing him to eventually choose the optimal effort $a^* = 1$. Suppose, however, the artist is potentially biased in his self-perception and assigns probability 1 to $\hat{b} \in \{0, 1, 2\}$. Specifically, $\hat{b} = 2$ corresponds to overconfidence and $\hat{b} = 0$ corresponds to underconfidence. This bias gives rise to a misspecified model of how sales are generated—the artist overestimates or underestimates the expected amount of sales for each possible level of effort and each possible value of the market demand.[9] Suppose the artist considers a competing model and may switch to this model if it fits the data significantly better. Are underconfidence and overconfidence equally likely to persist? My results reveal an interesting asymmetry—overconfidence tends to be more robust than underconfidence.[10]

Let's first consider the case of an underconfident artist who believes he has little talent, $\hat{b} = 0$, but entertains a correctly specified competing model that attaches probability 1 to the true talent $b^* = 1$. We can show that the underconfidence model does not persist against the competing model, as it consistently produces lower accuracy in fitting the data. To understand why, first note that underconfidence leads the artist to mistakenly attribute higher-than-expected sales to strong market demand, thereby encouraging a high level of effort. What comes next is critical: this high effort leads to a partial correction in the artist's overestimation of market demand. Due to complementarity, the marginal return to demand increases with effort, allowing the artist to

[9]This modeling approach captures the idea that individuals often commit fundamental attribution errors and are slower in changing self-perceptions than in updating beliefs about the outside environment (Miller and Ross, 1975). Heidhues et al. (2018) also adopt this approach and show that both over- and underconfidence lead to wrong inferences about market demand and inefficient choices of effort in the long run.

[10]In Section 5, I extend this example to allow for more general payoffs and outcome distributions.

explain sales data without an excessive overestimation of market demand. However, repeating this logic, a lower belief in market demand reduces effort and raises belief again, generating a negative feedback loop. Specifically, repeatedly choosing $\hat{a}^1 = 1$ shifts the artist's belief about market demand toward $\hat{\omega}^1 = 4$. This strong market demand then incentivizes a higher effort, $\hat{a}^2 = 2$. However, choosing $\hat{a}^2$ subsequently shifts belief toward a weaker market demand, $\hat{\omega}^2 = 3$, making the lower effort $\hat{a}^1$ optimal. Mathematically, this oscillation between efforts is illustrated by the equations:

$$(\hat{a}^1 + b^*) \cdot \omega^* = (1+1) \cdot 2 = (\hat{a}^1 + \hat{b}) \cdot \hat{\omega}^1 = (1+0) \cdot \mathbf{4}, \tag{1}$$
$$(\hat{a}^2 + b^*) \cdot \omega^* = (2+1) \cdot 2 = (\hat{a}^2 + \hat{b}) \cdot \hat{\omega}^2 = (2+0) \cdot \mathbf{3}. \tag{2}$$

As a result, the artist's effort perpetually cycles between 1 and 2, and no single market demand value can perfectly explain all the data—the model lacks a self-confirming equilibrium. In contrast, the correct competing model consistently achieves perfect prediction accuracy in the long run. Consequently, the artist, regardless of initial reluctance, would abandon the underconfidence model and correct his biased self-perception.

Now, let's consider an overconfident artist who believes his talent level is instead given by $\hat{b} = 2$ while also entertaining a correctly specified competing model. In contrast to the previous case, the overconfidence model exhibits perfect asymptotic accuracy. Overconfidence leads the artist to mistakenly attribute disappointing sales to low demand and respond by exerting a low effort. Crucially, a choice of lower effort induces an even lower belief—the marginal return to market demand decreases as effort drops, necessitating a larger inference-truth gap to rationalize the unsatisfactory sales. The positively reinforcing dynamics eventually drive the artist to believe that the market demand is $\hat{\omega} = 1$ and exert an inefficiently low amount of effort $\hat{a} = 0$. This steady state constitutes a self-confirming equilibrium—zero effort is optimal against the misguided belief about market demand, and this low belief perfectly aligns with the sales data given the misspecified model:

$$(\hat{a} + b^*) \cdot \omega^* = (0+1) \cdot 2 = (\hat{a} + \hat{b}) \cdot \hat{\omega} = (0+2) \cdot \mathbf{1}. \tag{3}$$

At this steady state, the initial model and the competing model generate equally accurate predictions, which suggests that the artist may stick with the overconfidence model forever. But this is not the end of the story. The equilibrium analysis indicates that overconfidence has the potential to persist, but it does not rule out switches in the course of converging to the steady state. The dynamic model switching framework introduced in Section 3 addresses this concern. My characterization implies that the overconfidence model is globally robust, persisting against the correctly specified competing model (and many other models), when the associated prior attaches sufficiently high probability to the low demand $\hat{\omega} = 1$. Moreover, as discussed in more detail in Section 4, the required prior tightness is inversely related to the switching stickiness.

3 Framework

3.1 Setup

Objective Environment. Consider an infinitely repeated decision problem with a myopic agent.[11] In each period $t = 0, 1, 2, ...$, the agent chooses an action a_t from a finite set $\mathcal{A}$ and subsequently observes the realization of a random outcome y_t from $\mathcal{Y}$. The set of possible outcomes $\mathcal{Y}$ is either an Euclidean space or a compact subset of an Euclidean space with at least two elements. The agent's choice of action may affect the distribution of the immediate outcome. Conditional on a_t, outcome y_t is independently drawn from the probability measure $Q^*(\cdot|a_t) \in \Delta\mathcal{Y}$. The true data-generating process $Q^* \in (\Delta\mathcal{Y})^{|\mathcal{A}|}$ remains fixed throughout. At the end of period t, the agent obtains a flow payoff $u_t := u(a_t, y_t) \in \mathbb{R}$. The payoff function u is known to the agent. Let $h_t := (a_\tau, y_\tau)_{\tau=0}^{t-1}$ denote the observable history in the beginning of period t and $H_t = (\mathcal{A} \times \mathcal{Y})^t$ denote the set of all such histories. The true DGP and the payoff function satisfy the following standard assumptions.

Assumption 1. *For all $a \in \mathcal{A}$: (i) $Q^*(\cdot|a)$ is absolutely continuous w.r.t. a common measure ν, and the Radon-Nikodym derivative $q^*(\cdot|a)$ is positive and continuous; (ii) $u(a, \cdot) \in L^1(\mathcal{Y}, \mathbb{R}, Q^*(\cdot|a))$.*[12]

Under Assumption 1 (i), the true DGP admits a positive and continuous density $q^*(\cdot|a)$ for each action $a \in \mathcal{A}$. When $\mathcal{Y}$ is discrete, $q^*(\cdot|a)$ is the probability mass function and ν is the counting measure; when $\mathcal{Y}$ is a continuum, $q^*(\cdot|a)$ is the probability density function and ν is the Lebesgue measure. Assumption 1 (ii) ensures that the agent's objective expected period-t payoff, $\overline{u}_t := \int_{\mathcal{Y}} u(a_t, y) q^*(y|a_t) \nu(dy)$, is well-defined and an objectively optimal action exists.

[11] I show that most of the results extend to a non-myopic agent who maximizes the expected discounted sum of payoffs based on her current model (see Section 6.2).

[12] $L^p(\mathcal{Y}, \mathbb{R}, \nu)$ denotes the space of all functions $g : \mathcal{Y} \to \mathbb{R}$ s.t. $\int |g(y)|^p \nu(dy) < \infty$.

Subjective Models. The decision problem becomes straightforward if the agent knows the true DGP—she can simply play an objectively optimal action in every period. However, the agent does not necessarily have access to this knowledge and instead relies on models to guide decisions. The universe of models, denoted by $\overline{\Theta}$, is the set of all possible finite collections of data-generating processes, i.e. $\overline{\Theta} := \{\theta \in \mathcal{P}\left((\Delta\mathcal{Y})^{|\mathcal{A}|}\right) : |\theta| < \infty\}$. Each model $\theta \in \overline{\Theta}$ consists a finite collection of predictions regarding the DGP. For ease of interpretation, each prediction is labeled by a parameter value within a model-specific *parameter space* Ω^θ. Given a parameter value $\omega \in \Omega^\theta$, model θ predicts a data-generating process $\{Q^\theta(\cdot|a,\omega)\}_{a\in\mathcal{A}}$. A model with a larger parameter space allows for a greater number of potential DGPs. I assume the agent can only entertain models satisfying Assumption 2 and denote the set of such models as $\Theta \subset \overline{\Theta}$.

Assumption 2. *For each $\theta \in \Theta$ and each $a \in \mathcal{A}$: (i) for all $\omega \in \Omega^\theta$, $Q^\theta\left(\cdot|a,\omega\right)$ is absolutely continuous w.r.t. measure ν, and the Radon-Nikodym derivative $q^\theta\left(\cdot|a,\omega\right)$ is positive and continuous; (ii) for all $\omega \in \Omega^\theta$, $u\left(a,\cdot\right) \in L^1\left(\mathcal{Y},\mathbb{R},Q^\theta\left(\cdot|a,\omega\right)\right)$; (iii) for all $\omega \in \Omega^\theta$, there exists $r_a \in L^2\left(\mathcal{Y},\mathbb{R},\nu\right)$ such that $\left|\ln\frac{q^*(\cdot|a)}{q^\theta(\cdot|a,\omega)}\right| \leq r_a\left(\cdot\right)$ a.s.-$Q^*\left(\cdot|a\right)$.*

Assumption 2 (i) and (ii) mirror Assumption 1, ensuring the existence of a density function and that the expected payoffs predicted by any model are well-defined. Assumption 2 (iii) requires that the log-likelihood ratio between the predictions of any model and the true DGP is bounded almost surely, which also implies that no models rule out events that occur with positive probability under the true DGP.

A model θ is said to be *correctly specified* if its predictions include the true DGP—there exists $\omega \in \Omega^\theta$ such that $q^*\left(\cdot|a\right) \equiv q^\theta\left(\cdot|a,\omega\right), \forall a \in \mathcal{A}$—and *misspecified* otherwise. I denote the smallest correctly specified model as θ^*. Namely, this model solely consists of the true DGP, and hence I refer to θ^* as the *true model*.

3.2 The Switcher's Problem

The agent considers a restricted set of models, $\Theta^\dagger \subseteq \Theta$. It is often assumed in the literature that the decision-maker is a *dogmatic modeler* who uses a single model throughout. For convenience, I refer to a dogmatic modeler with $\Theta^\dagger = \{\theta\}$ as a *θ-modeler*. My primary focus here diverges from this conventional assumption, as I explore the concept of a *switcher*. A switcher employs only one model at any given moment but can switch between different models across periods. For the main analysis, I restrict attention to the two-model case where $\Theta^\dagger = \{\theta, \theta'\}$.[13] The agent's model choice in period t is denoted by $m_t \in \Theta^\dagger$, where the initial model choice is $m_0 = \theta$. A switcher's learning environment can be summarized by a quadruple, $E = (\theta, \theta', \pi_0^\theta, \pi_0^{\theta'})$, where the first two elements represent the *initial model* and the *competing model*, respectively, and the last two correspond to the agent's prior beliefs regarding the parameters of these models, denoted by $\pi_0^\theta \in \Delta\Omega^\theta$ and $\pi_0^{\theta'} \in \Delta\Omega^{\theta'}$.[14] Without loss of generality, all priors are assumed to have full support. I now describe the events in period t in chronological order.

Model switching. In period 0, the agent adopts θ and proceeds to the phase of choosing an action. At the beginning of period $t \geq 1$, the agent employs a *Bayes factor rule* to determine whether she should switch. The agent calculates the Bayes factor λ_t, which is the ratio of the likelihood of the data under model θ' to the likelihood of the data under model θ. Here, the likelihood of the data under a model is a weighted sum of the likelihoods of the data under all DGPs included in the model, with the weights given by the prior. That is,

$$\lambda_t := \frac{\ell_t(\theta')}{\ell_t(\theta)} := \frac{\sum_{\omega \in \Omega^\theta} \pi_0^\theta(\omega)\ell_t(\theta, \omega)}{\sum_{\omega' \in \Omega^{\theta'}} \pi_0^{\theta'}(\omega')\ell_t(\theta', \omega')}, \tag{4}$$

where

$$\ell_t(\theta, \omega) := \prod_{\tau=0}^{t-1} q^\theta \left(y_\tau | a_\tau, \omega\right), \tag{5}$$

and $\ell_t(\theta', \omega')$ is defined analogously. Fix a constant *switching threshold* $\alpha \geq 1$. If $m_{t-1} = \theta$, then the agent switches to θ' if and only if the Bayes factor exceeds the switching threshold, $\lambda_t > \alpha$. Conversely, if $m_{t-1} = \theta'$, the agent switches back to θ if and only if λ_t drops below the inverse of the switching threshold, $\lambda_t < 1/\alpha$. In cases where $1/\alpha \leq \lambda_t \leq \alpha$, the agent does not consider the existing evidence sufficient to warrant a model switch. The parameter α thus serves as a measure of the *stickiness* of the switching process, with larger values of α requiring stronger evidence for a switch.[15]

[13] This model switching framework can be extended to allow for three or more models in $\Theta^\dagger$. This extension is analyzed in Section 6.1.

[14] I treat these priors as part of the learning environment rather than as primitives of the subjective models. This choice is made for the sake of expositional convenience, allowing me to separately characterize which models persist or are robust under at least one prior and to identify which specific priors confer properties (see Section 4).

[15] The symmetry in the switching threshold is not important for the main results.

Assumptions on the switching rule are discussed in Section 3.3.

Learning. The agent then updates her beliefs over the parameters of both models based on the last realized outcome.[16] This is captured by two recursive belief processes, $\pi_t^\theta = B^\theta(a_{t-1}, y_{t-1}, \pi_{t-1}^\theta)$ and $\pi_t^{\theta'} = B^{\theta'}(a_{t-1}, y_{t-1}, \pi_{t-1}^{\theta'})$, where $B^\theta : \mathcal{A} \times \mathcal{Y} \times \Delta\Omega^\theta \to \Delta\Omega^\theta$ is the Bayesian operator for model θ and $B^{\theta'}$ is the Bayesian operator for θ'.[17] Note that now we can alternatively express the Bayes factor recursively in terms of the posteriors,

$$\lambda_t = \lambda_{t-1} \times \frac{\sum_{\omega' \in \Omega^{\theta'}} \pi_{t-1}^{\theta'}(\omega')\, q^{\theta'}(y_{t-1}|a_{t-1}, \omega')}{\sum_{\omega \in \Omega^\theta} \pi_{t-1}^{\theta}(\omega)\, q^{\theta}(y_{t-1}|a_{t-1}, \omega)} \tag{6}$$

$$= \prod_{\tau=0}^{t-1} \frac{\sum_{\omega' \in \Omega^{\theta'}} \pi_{\tau}^{\theta'}(\omega')\, q^{\theta'}(y_{\tau}|a_{\tau}, \omega')}{\sum_{\omega \in \Omega^\theta} \pi_{\tau}^{\theta}(\omega)\, q^{\theta}(y_{\tau}|a_{\tau}, \omega)} \tag{7}$$

This expression carries an intuitive interpretation: the first term on the right-hand side of (6) compares how well the models explain the observable history up until the last period, and the second term compares how well they explain the most recent observation, taking into account parameter estimates derived from past data. Hence, the agent assesses a model's performance not only based on its current explanatory power but also on its full historical track record.

Actions. The agent selects an action based on a pure policy that is optimal under the current model m_t. The policy under θ, denoted by $f^\theta : \Delta\Omega^\theta \to \mathcal{A}$, is a selection from the correspondence $A_m^\theta : \Delta\Omega^\theta \rightrightarrows \mathcal{A}$ that returns all myopically optimal actions at a given belief. For any belief $\pi_t^\theta \in \Delta\Omega^\theta$,

$$A_m^\theta(\pi_t^\theta) \coloneqq \arg\max_{a \in \mathcal{A}} \sum_{\omega \in \Omega^\theta} \pi_t^\theta(\omega) \int_{y \in \mathcal{Y}} u(a, y) q^\theta(y|a, \omega)\nu(dy). \tag{8}$$

Analogously, the policy under θ', denoted by $f^{\theta'}$, is a selection from $A_m^{\theta'}$.

Within any single period t, a θ-modeler behaves identically as a switcher with $m_t = \theta$ who shares the same belief over the parameters of θ. That is, both types of agents update their belief over the parameters of model θ and then choose a myopically optimal

[16]Note that the agent revises her beliefs by incorporating all observed outcomes, including those generated under an alternative model. This is because all data provide information about the true DGP, and there is no justification for the agent to ignore any observations.

[17]Nevertheless, the agent only needs to keep track of her posterior belief for model m_t to make choices. Note that neither the switching decision captured by (4) nor the action rule captured by (8) requires the agent to use her posterior beliefs for the model that is not currently in use.

action. However, while a θ-modeler always uses a fixed decision rule f^θ, the switcher's decision rule depends on her current model choice. Consequently, a θ-modeler and a switcher may exhibit significant differences in behavior and beliefs across periods.

3.3 Discussion on the Switching Rule

Bayes factor. The Bayes factor rule is a common model selection criterion in Bayesian statistics. Several recent studies in the literature on model-based learning and persuasion have also used the Bayes factor rule (Schwartzstein and Sunderam, 2021; Aina, 2023). It is a natural choice in this environment for the following reasons. First, the Bayes factor has a strong Bayesian flavor, so the agent maintains some form of conceptual consistency in belief updating and model switching. To see this, observe that if our agent were to calculate the posterior odds of two models based on a uniform prior, she would find that the posterior odds ratio is precisely given by the Bayes factor. Second, the Bayes factor rule is flexible in that it could be easily formulated for any model and any data-generating processes, without imposing assumptions about the underlying parametric structure.[18] Finally, it is well known from the statistics literature that the Bayes factor automatically includes a penalty for including too much structure into the model and helps prevent overfitting (Kass and Raftery, 1995).[19] This is consistent with the empirical observation that people in general favor simple models.

Sticky switching. I allow the agent to exhibit switching stickiness, as indicated by the assumption that $\alpha \geq 1$. Switching stickiness is well observed in reality and can stem from a variety of causes, such as conservatism, concerns about overreaction to noises, or mental and physical costs associated with model switching. For example, universities base their promotion decisions on models of faculty performance that heavily rely on key statistics such as the rankings of the journals. While such an evaluation system can be deeply flawed, implementing a new system is highly costly, and thus a model switch

[18] Common alternative rules in statistics such as the BIC and the AIC are shown to approximate the Bayes factor under certain assumptions about the parametric family and the prior.

[19] The use of the prior in Bayes factor calculations plays a pivotal role in this observation. Suppose the agent instead calculates the likelihood ratio of the models using maximum likelihood estimates. That is, she calculates $\tilde{\lambda}_t := \tilde{l}_t(\theta')/\tilde{l}_t(\theta)$, where $\tilde{l}_t(\theta) := \max_{\omega \in \Omega^\theta} \prod_{\tau=0}^{t-1} l_t(\theta, \omega)$ and $\tilde{l}_t(\theta') := \max_{\omega' \in \Omega^{\theta'}} \prod_{\tau=0}^{t-1} l_t(\theta', \omega')$. In this case, the likelihood associated with a model always weakly improves as the model expands to include more DGPs, indicating a preference for larger models. However, a nuanced trade-off between accuracy and simplicity arises when using the Bayes factor. In particular, the prior becomes more diffuse with the inclusion of additional DGPs, and the weighted average likelihood of a model may decrease if the newly added DGPs exhibit poorer fit compared to the original simpler model considered as a whole.

only occurs when a meta-analysis of other potential evaluation systems points to a clear winner. In the statistics literature on Bayesian model selection, Kass and Raftery (1995) recommend using a threshold of 20 as the requirement of "strong evidence" in favor of the competing model. One important goal of this paper is to delineate the implications of stickiness on model persistence, but a full-fledged micro-foundation for sticky switching is outside the scope of this paper.

4 Main Results

4.1 Definitions

Model θ is said to *persist* against θ' if the agent eventually stops switching and settles on model θ with positive probability.[20] This implies that if θ persists against θ', then with positive probability, the agent resembles a θ-modeler in the long term. Conversely, if θ fails to persist against θ', then the agent either adopts the competing model permanently or oscillates forever between the two models. In both scenarios, the learning outcomes of the agent can significantly diverge from that of a θ-modeler.

Definition 1. Model θ *persist against* θ' at priors π_0^θ and $\pi_0^{\theta'}$ if, given $E = (\theta, \theta', \pi_0^{\theta'}, \pi_0^{\theta'})$, there exists $T \geq 0$ such that with positive probability, $m_t = \theta$ for all $t \geq T$.

Two observations about Definition 1 are in order. First, note that persistence is *prior-sensitive*—a model could persist against a given competing model under certain priors but not under other priors. In fact, priors potentially affect persistence in two ways. First, priors play a direct role in the calculation of the Bayes factor. Second, priors affect the agent's behavior and, as a result, endogenously influence the distribution of the random outcome. We are interested in not only identifying which models persist but also understanding how their persistence depends on the agent's prior belief about the data-generating process. Second, persistence is a *relative* concept—a model could persist against some competing models but not against others. However, the specific competing models that may arise and the priors assigned to them are context-dependent and unpredictable. This observation motivates a robustness approach, i.e., characterizing models that persist against a wide range of competing models.

Critically, the scope of robustness can vary considerably with the set of admissible competing models and priors, particularly in terms of their distance from the initial model and its prior (i.e. allowable maximum step size of switching). I introduce two

[20] I construct the underlying probability space in Appendix A.

concepts of robustness that establish upper and lower boundaries: Global robustness allows for an unlimited step size of switching, while local robustness restricts to the most minimal allowable step size of switching. Formally, model θ is said to be globally robust at a given prior if it persists irrespective of the competing model it is compared against and the prior assigned to that model. Note that the property of being prior-sensitive is inherited by global robustness from the concept of persistence. When θ is not globally robust at any prior, one can always identify a competing model associated with some prior that almost surely replaces θ infinitely often.

Definition 2 (Global robustness)**.** Model $\theta \in \Theta$ is *globally robust at prior* π_0^θ if θ persists against every competing model $\theta' \in \Theta$ at π_0^θ and $\pi_0^{\theta'}$ for every $\pi_0^{\theta'} \in \Delta\Omega^{\theta'}$.

To define local robustness, I first provide a measure to quantify the distance between two arbitrary models θ and θ' and their priors. Since both models are a finite collection of DGPs, a natural candidate for measuring their distance is the Hausdorff distance between the two sets of DGPs, where the distance between any two DGPs is measured based on the Prokhorov metric.[21] For convenience, denote the DGP to which model θ and parameter ω correspond by $Q^{\theta,\omega}$. I define the distance between $Q^{\theta,\omega}$ and $Q^{\theta',\omega'}$ as the maximum Prokhorov distance between the outcome distributions across all actions, $d(Q^{\theta,\omega}, Q^{\theta',\omega'}) := \max_{a\in\mathcal{A}} d_P(Q_a^{\theta,\omega}, Q_a^{\theta',\omega'})$. The distance between model θ and θ' is then given by the Hausdorff metric, $d(\theta,\theta') := d_H\left(\{Q^{\theta,\omega}\}_{\omega\in\Omega^\theta}, \{Q^{\theta',\omega'}\}_{\omega'\in\Omega^{\theta'}}\right)$. This allows us to define an ϵ-*neighborhood* of model θ, denoted by $N_\epsilon(\theta) := \{\theta' \in \Theta : d(\theta,\theta') < \epsilon\}$. Finally, note that while prior π_0^θ and prior $\pi_0^{\theta'}$ are defined on potentially different parameter spaces, each of them corresponds to a distribution over the set of all DGPs. With an abuse of notation, we use $d_P(\pi_0^\theta, \pi_0^{\theta'})$ to denote the Prokhorov distance between the implied distributions over DGPs, and define a ϵ-neighborhood of prior π_0^θ within the set of possible priors for θ' as $N_\epsilon(\pi_0^\theta) := \left\{\pi_0^{\theta'} \in \Delta\Omega^{\theta'} : d_P(\pi_0^\theta, \pi_0^{\theta'}) < \epsilon\right\}$.

Definition 3 (Local robustness)**.** Model $\theta \in \Theta$ is *locally robust at prior* π_0^θ if there exists $\epsilon > 0$ such that θ persists against every competing model $\theta' \in N_\epsilon(\theta)$ at priors π_0^θ and $\pi_0^{\theta'}$ for every $\pi_0^{\theta'} \in N_\epsilon(\pi_0^\theta)$.

[21] The Prokhorov metric measures the distance between any two probability distributions on the same metric space. For any two probability measures μ and μ' over $\mathcal{Y}$, the Prokhorov distance is given by $d_P(\mu,\mu') := \inf\{\epsilon > 0 | \mu(Y) \le \mu'(B_\epsilon(Y)) + \epsilon \text{ and } \mu'(Y) \le \mu(B_\epsilon(Y)) + \epsilon \text{ for all } Y \subseteq \mathcal{Y}\}$. The results in this paper also hold for alternative metrics such as Kullback-Leibler divergence or total variation. The Hausdorff metric measures how far two subsets of the same metric space are from each other. For any two sets X and Z, their Hausdorff distance is $d_H(X,Z) = \max\{\sup_{x\in Z} d(x,Z), \sup_{z\in Z} d(X,z)\}$.

Local robustness requires that there exists some positive ϵ such that the model persists against nearby models with nearby priors within the relevant ϵ-neighborhoods. Hence, a locally robust model persists as long as the agent takes sufficiently small steps by considering sufficiently close perturbations. By contrast, if θ is not locally robust, there must exist a local perturbation of θ that replaces model θ infinitely often.

4.2 Which Models Can Be Robust?

I begin by characterizing models that are locally or globally robust for at least one full-support prior. This characterization is useful because it directly speaks to the question of which models *can* be robust—failing to be robust at any full-support prior implies non-robustness across all initial conditions. Since all priors are assumed to be full-support without loss, I sometimes omit this adjective for convenience.

It is instructive to start our analysis with a special case: which models can persist against a correctly specified model? It is a well-known result that with a correctly specified model, a learner assigns probability 1 to a data-generating process that predicts the true outcome distribution in the limit (Easley and Kiefer, 1988). It follows that such a model perfectly matches the data in the long term, and thus any model that persists in its presence must also have perfect prediction accuracy in the limit. Since outcomes are endogenously generated by actions, this observation suggests that the agent, possibly after a lot of back-and-forth switching, eventually converges to a *self-confirming equilibrium* as defined below.

Definition 4. A strategy $\sigma \in \Delta\mathcal{A}$ is a *self-confirming equilibrium* (SCE) under model θ if there exists a supporting belief $\pi^\theta \in \Delta\Omega^\theta$ such that the following conditions hold.

(1) Optimality: σ is myopically optimal against π^θ, i.e. $\sigma \in \Delta A^\theta_m(\pi^\theta)$.

(2) Consistency: π^θ is consistent at σ in that $q^\theta(\cdot|a,\omega) \equiv q^*(\cdot|a)$ for all $a \in \mathrm{supp}(\sigma)$ and all $\omega \in \mathrm{supp}(\pi)$.

In an SCE, the agent plays actions that are myopically optimal based on a consistent belief which ensures that the model prediction fully aligns with the objective outcome distribution. It's worth noting that an SCE may not necessarily be efficient. While consistency implies correct predictions regarding the payoffs achieved in equilibrium, the model could have entirely incorrect predictions for payoffs off the equilibrium path.

But persistence against a correct model implies more than the existence of a SCE—the SCE must also be stable and reachable. In particular, the agent should, with positive probability, end up playing *only* the equilibrium actions under model θ. If non-SCE actions are played infinitely often, the Bayes factor would still diverge to infinity and result in a permanent abandonment of model θ. Note that on paths where θ is adopted forever, a switcher eventually behaves no differently than a θ-modeler. Thus, a necessary condition for model θ to persist is that a θ-modeler eventually only plays the equilibrium actions with positive probability. I refer to this stability notion as p-absorbingness, where "p" means that the equilibrium is absorbing *with positive probability.*

Definition 5. Strategy $\sigma \in \Delta\mathcal{A}$ is *p-absorbing* under θ if there exists a full-support prior π_0^θ and some integer $T \geq 0$ such that, with positive probability, a θ-modeler only plays actions in $\text{supp}(\sigma)$ for all $t \geq T$.

P-absorbingness differs from absorbingness or other existing stability notions of self-confirming equilibria in the literature. In particular, it does not imply that the θ-modeler's or the switcher's action sequence converges to a single action in the support of σ, or their action frequency converges to σ, or convergence of any kind occurs with probability 1.[22] Rather, it allows for non-convergent behavior within the support of σ, but rules out the scenario where the modeler almost surely plays actions outside the support of σ infinitely often.[23] Although p-absorbingness is a relatively weak requirement, not all self-confirming equilibria are p-absorbing. Since a θ-modeler's learning outcome is independent from the model switching process, we can further characterize p-absorbingness with conditions based on the primitives of model θ only. Lemma 1 shows that quasi-strictness is a simple sufficient condition for p-absorbingness.[24]

Lemma 1. *A self-confirming equilibrium is p-absorbing under θ if it is quasi-strict.*

Quasi-strictness ensures that at any belief sufficiently close to the equilibrium belief, it is strictly optimal to play the actions prescribed by the equilibrium. Since the equilibrium is self-confirming, the equilibrium belief is consistent with the actual outcome distribution, and thus with positive probability, a θ-modeler's belief stays within

[22]For example, p-absorbingness is weaker than the stability notion proposed by Fudenberg et al. (2021). By their definition, a pure equilibrium a^* under θ is stable if for every $\kappa \in (0,1)$, there exists a belief $\pi \in \Delta\Omega^\theta$ such that for any prior π_0^θ sufficiently close to π, the dogmatic modeler's action sequence a_t converges to a^* with probability larger than κ. They do not define a stability notion for a mixed equilibrium.

[23]Indeed, a dogmatic modeler's action may never converge even when she eventually only plays the actions in the support of a p-absorbing SCE (see Example 4 in Appendix C).

[24]As conventional in the literature, a self-confirming equilibrium σ is said to be *quasi-strict* if there exists a supporting belief π^θ such that $\text{supp}(\sigma) = A_m^\theta(\pi^\theta)$, that is, any action outside the support of σ is strictly suboptimal given π^θ.

a small neighborhood of the equilibrium belief. Taken together, this implies that starting from a prior sufficiently close to the equilibrium supporting belief, the θ-modeler plays the SCE forever with positive probability. Therefore, such an SCE is p-absorbing.

I conclude our analysis of a correctly specified competing model with Lemma 2.

Lemma 2. *If model θ persists against a correctly specified model θ' at some priors π_0^θ and $\pi_0^{\theta'}$, then there exists a p-absorbing SCE under θ.*

I say that a model has *perfect asymptotic accuracy* or is *asymptotically accurate* if it induces a p-absorbing SCE. At first glance, Lemma 2 offers merely a necessary condition for global robustness. On one hand, perfect asymptotic accuracy appears unnecessarily strong for local robustness, given that local robustness only requires persistence against local perturbations. This is because any local perturbation of any misspecified model is necessarily misspecified and Lemma 2 does not apply. On the other hand, it is unclear whether the existence of a p-absorbing SCE alone would be sufficient for global robustness, even if the agent can start from any arbitrary full-support prior. Lemma 2 does not tell us whether perfect asymptotic accuracy is also sufficient for a model to persist against a correctly specified model, and even if this is true, it does not necessarily imply persistence against every other competing model in Θ. Surprisingly, as I show in Theorem 1, when switching exhibits stickiness, perfect asymptotic accuracy is both necessary for local robustness and sufficient for global robustness, which equates the two robustness notions provided flexibility in the prior.

Theorem 1. *Suppose $\alpha > 1$, then the following are equivalent:*

(i) Model θ is globally robust for at least one full-support prior.

(ii) Model θ is locally robust for at least one full-support prior.

(iii) There exists a p-absorbing SCE under model θ.

First, Theorem 1 provides a formal microfoundation for the persistence of certain types of misspecified models. A misspecified model can persist against any arbitrary competing model as long as it has perfect asymptotic accuracy as captured by the existence of a p-absorbing SCE. Notably, Theorem 1 does not depend on switching threshold α as long as $\alpha > 1$, meaning that the sets of models that can be locally or globally robust do not expand as switching becomes stickier.

Second, Theorem 1 offers a new perspective for understanding local and global robustness by showing the equivalence between (i) and (ii). If a model fails to be globally robust, the switcher need not go far to find an attractive alternative—models that do not persist against major paradigm shifts are also vulnerable to local changes. This observation highlights the necessity of perfect asymptotic accuracy for robust misspecified models.

Third, Theorem 1 reveals that the demanding notion of global robustness amounts to the requirement that θ persists against *one* correctly specified model at some prior. In other words, if θ can persist against a competing model that assigns a tiny probability to the true DGP, then it is also capable of persisting against the true model, or any other model with an arbitrarily complex parameter space. Conversely, a model that fails to be globally robust does not persist as long as the agent starts considering any correctly specified competing model. As an immediate corollary, any correctly specified model is globally robust since a model must persist against itself.

The above observation, combined with Lemma 1, leads to the following corollary. This result simplifies the application of Theorem 1 in specific contexts, allowing us to quickly determine whether a given model can be locally or globally robust by confirming its correctness or its ability to induce a quasi-strict SCE.

Corollary 1. *Suppose $\alpha > 1$, then model θ is locally or globally robust for at least one prior if θ is correctly specified or there exists a quasi-strict SCE under θ.*

When switching exhibits no stickiness, meaning $\alpha = 1$, the set of models that can be locally or globally robust for at least one full-support prior shrinks discontinuously. In this case, robustness requires not only asymptotic accuracy but also the prior to be fully concentrated on the set of p-absorbing self-confirming equilibria. A detailed exploration of this case is deferred to Section 4.4, where we will investigate the role of the prior.

4.3 Proof idea of Theorem 1

In this section, I provide an explanation and intuition for the proof of Theorem 1, following the order of (i)⇒(ii)⇒(iii)⇒(i).

From (i) to (ii). It directly follows from the definitions of global and local robustness that the former necessarily implies the latter.

From (ii) to (iii). Local robustness requires perfect asymptotic accuracy. Consider an initial model θ that is not asymptotically accurate. Although the agent is restricted to contemplate only local perturbations, such perturbations are unconstrained and can be made towards the direction of higher asymptotic accuracy. Specifically, we could construct a competing model θ' within the ϵ-neighborhood of θ such that θ' fits data better than θ. To do this, we simply take the predictions of θ' to be a convex combination between the predictions of θ and the true DGP for every action in $\mathcal{A}$. Since the Kullback-Leibler (KL henceforth) divergence is convex, the mixture model θ' yields a strictly lower KL divergence than model θ in the limit. While the discrepancy in the KL divergence between the models can be arbitrarily small, the Bayes factor diverges to infinity almost surely as the agent draws a sufficient number of outcome realizations. Therefore, the Bayes factor must eventually surpass the switching threshold and the agent will switch to θ' permanently.[25]

From (iii) to (i). Asymptotically accurate models can be globally robust. This result may appear straightforward at first glance, as p-absorbingness ensures that the SCE is reachable from a full-support prior and consistency ensures that model θ fits the data weakly better than any competing model in the equilibrium. However, the fact that the SCE is reachable for a θ-modeler from a certain prior does not imply that it is also reachable for our switcher agent from the same prior or any other full-support prior. To illustrate this, first note that in the active learning framework that is considered here, a dogmatic modeler's behavior and beliefs are endogenous and can mutually influence each other. For a switcher, in addition to her behavior and beliefs within each model, her model choice is also endogenous, and all three of these endogenous objects can influence one another. This complex interaction can lead to challenges that prevent the agent from converging to the SCE, even when the SCE is p-absorbing. In particular, the outcome realizations that drive a dogmatic modeler to the SCE may, in fact, trigger a switch away from model θ, making its adoption *self-defeating*.

Such challenges are inherent to the multiple-model learning framework and thus may be of independent interest to future research pursuits on problems other than persistence and robustness. Therefore, I will first take some time to elucidate this issue further with Example 1. For simplicity, in this example I take the competing model to be the true model, but a similar phenomenon can occur with a competing model that is arbitrarily close to the initial model.

Example 1 (Self-defeating Models)**.** In each period, an agent chooses from two tasks

[25]The Kullback-Leibler divergence of a density q from another density q' is given by $D_{KL}(q \parallel q') = \int_{\mathcal{Y}} q \ln(q/q')\, \nu(dy)$. The KL divergence is an asymmetric non-negative distance measure between q and q', which is minimized to zero if and only if q and q' coincide almost everywhere. It is convex in the following sense: for any two pairs of densities (q_1, q_2) and (q_1', q_2') and any $\gamma \in [0,1]$, we have $D_{KL}(\lambda q_1 + (1-\lambda)q_2 \parallel \lambda q_1' + (1-\lambda)q_2') \leq \lambda D_{KL}(q_1 \parallel q_1') + (1-\lambda)D_{KL}(q_2 \parallel q_2')$.

$q^{\theta}(1\|a,\omega)$	ω^1	ω^2
a^1	0.5	0.3
a^2	0.4	0.4

$q^{\theta^*}(1\|a,\omega)$	ω^*
a^1	0.5
a^2	0.5

Table 2: Initial model θ and competing model θ' in Example 1.

$a_t \in \{a^1, a^2\}$ and observes the output of the chosen task $y_t \in \{0, 1\}$, where 0 represents failure and 1 represents success. The true DGP prescribes that successes and failures happen with equal probability 0.5 for either task. The agent is an expected output maximizer, so he would be indifferent between the tasks if the true DGP was known.

The agent holds a subjective model θ that presumes the success rate may depend on both the task type and his luck $\omega \in \Omega^\theta = \{\omega^1, \omega^2\}$, where ω^1 represents good luck and ω_2 represents bad luck (see Table 2). The agent believes that his luck is fixed and has a uniform prior over his luck, i.e. $\pi_0^\theta(\omega_1) = 0.5$. Under model θ, the agent believes Task 1 is risky and success occurs more often if he has good luck, while Task 2 is safe and its outcome is independent of his luck. Besides, the agent is overall *pessimistic* under θ because the assumed success rate is always (weakly) lower than its true level. His policy under θ prescribes Task 1 iff good luck is more likely than bad luck, i.e. $\pi_t^\theta(\omega_1) \geq 0.5$.[26] In addition, the agent entertains the competing model θ^* that correctly predicts the true success rate. Under model θ^*, the agent is indifferent and always chooses Task 1 (see Table 2). We consider the case where his switching threshold is given by $\alpha = 1.1$.

Choosing Task 1 is a strict self-confirming equilibrium under θ, supported by a degenerate belief at ω^1. To see why, note that the risky task is strictly optimal when the agent believes he has good luck; meanwhile, the superstitious belief of good luck offsets the overall pessimism and the model correctly predicts the success rate.

Starting from a uniform prior, a θ-modeler converges to playing Task 1 forever with positive probability. However, it turns out that model θ does not persist against model θ^* at the given priors, because any sequence of outcome realizations that leads to choosing a^1 must eventually trigger a model switch. As illustrated in Fig. 1, if the first realized outcome is a failure, the agent believes that he is more likely to have bad luck and thus switches his task choice to the safe task a^2 (Scenario 1); if the first realized outcome is a success, the agent switches his model choice to the more optimistic model θ^* in the next period (Scenario 2). In Scenario 1, the safe task choice causes the agent to stop updating on his luck. As a result, the agent never changes back to the risky

[26] The uniform prior is assumed for simple exposition. The mechanism in this example does not depend on the fact that the agent starts off being exactly indifferent between the tasks.

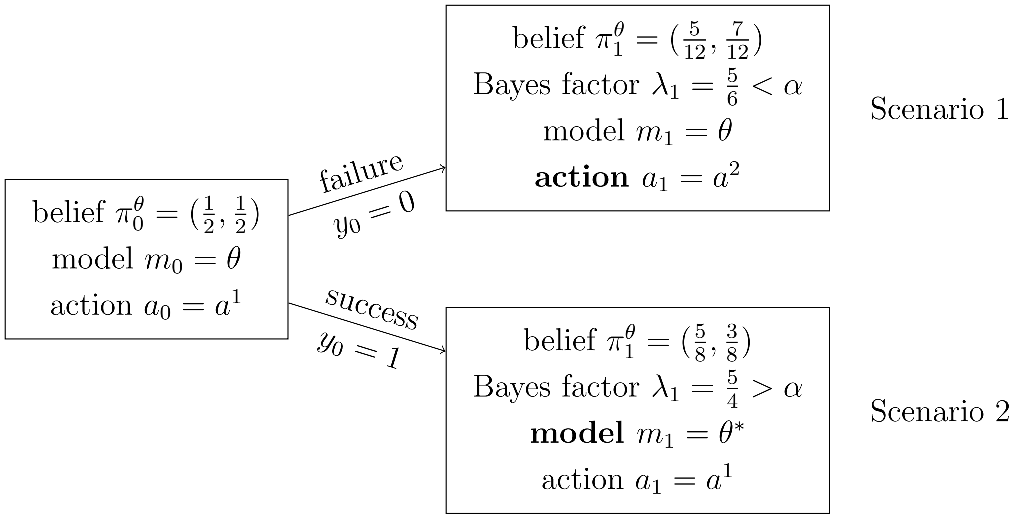

Figure 1: Scenario analysis in Example 1.

task a^1 as long as he remains under model θ. Since θ is incorrectly pessimistic when a^2 is chosen, the agent eventually switches to the correct model θ^* and enters Scenario 2. Once Scenario 2 occurs, the agent switches back to the overall pessimistic model θ only under the circumstance that he observes more failures than successes. But if so, the resulted posterior π_t^θ assigns higher probability to bad luck than good luck, which again induces the agent to choose the safe task a^2, bringing the agent back to Scenario 1. Eventually, the agent must abandon model θ and adopt the competing model θ^* forever. Therefore, θ does not persist against θ^* at the given priors.

Several factors contribute to the self-defeating result in Example 1. First, the agent's model choice and his belief on luck are tightly correlated given the particular structure of the models. In order for the agent to choose the risky task, he must believe in good luck more than bad luck, but the successes needed to induce this belief inevitably lead to a switch to the more optimistic competing model. Second, the agent's the task choice and model choice are too sensitive to early outcome realizations. Since the agent's prior π_0^θ is relatively far away from the SCE supporting belief and the switching threshold α is relatively low, a single observation is powerful enough to sway the agent's task choice or the model choice. Last but not least, the safe choice constitutes an absorbing *trap* in model θ because it causes the agent to stop updating his belief within θ. Indeed, the agent never finds it optimal to choose the risky option again under model θ once the agent enters the trap.

However, the agent can avoid falling into the trap if their initial beliefs are sufficiently close to the SCE supporting belief. Specifically, if the agent starts with a prior that assigns a sufficiently high probability to good luck, his task choice becomes less sensitive to early failures; meanwhile, his initial model becomes overall less pessimistic, and thus his model choice becomes less sensitive to early successes. The proof of Theorem 1 uses precisely this idea to show that there exists a full-support prior π_0^θ sufficiently close to the SCE supporting belief, such that self-defeating behavior does not arise with probability 1 and the initial model θ persists. First, I establish the existence of a full-support prior such that a θ-modeler consistently plays a p-absorbing SCE, and her belief remains within a small neighborhood of the equilibrium belief with arbitrarily high probability. Second, I employ Ville's maximal inequality to show that the probability of the likelihood ratio between the competing model and the true model ever exceeding the switching threshold α is bounded away from 1 (this step is trivial in Example 1). Finally, I show that this aforementioned likelihood ratio approximates the Bayes factor when the agent continues playing the SCE and the prior is highly concentrated around the SCE. Taken together, it follows that the probability of the agent playing the SCE indefinitely without switching to the competing model is strictly positive.

4.4 When are Models Robust?

Theorem 1 characterizes which models can be locally and globally robust based on their asymptotic accuracy but it remains silent about under which priors these models are locally or globally robust. The analysis of self-defeating models in Section 4.3 reveals that determining such priors becomes challenging in the presence of traps, though we know such priors do exist. To explore the potential role of priors, I propose two assumptions that eliminate traps from the model while keeping the prior fixed. Consider Example 1, where if model θ predicts even a slight dependence of success rate on the agent's luck for the safe task, the agent continues updating on luck after choosing the safe task. This eliminates the trap that would otherwise lock in the agent's task choice, rendering model θ non-self-defeating—we can construct a finite sequence of outcomes, though non-trivial and requiring proof, to make the agent confident in model θ while attaching high probability to having good luck. To rule out traps of the sort described in Example 1, we assume the model is *identifiable*, as defined below.

Definition 6. Model θ is *identifiable* if the predicted DGPs in θ prescribe different outcome distributions for each action, i.e. $q^\theta(\cdot|a,\omega) \neq q^\theta(\cdot|a,\omega')$ for all distinct $\omega, \omega' \in \Omega^\theta$ and $a \in \mathcal{A}$.

Non-identifiability is not the only cause of traps; other types of traps are more technical and arise when the p-absorbing SCE under model θ is not quasi-strict. In this case, there exists an action that is optimal given the equilibrium supporting belief but fails to be self-confirming. Under specific policies f^θ, these actions can also function as traps, meaning that once played, the agent is precluded from reverting to playing the SCE actions under the same model.Definition 7 collects the two no-trap conditions, both of which are relatively mild and can be easily verified from the primitives.[27]

Definition 7. Model θ has *no traps* if θ is identifiable and all p-absorbing SCEs (if exists) under θ are quasi-strict.

In the absence of traps, a reasonable conjecture is that any asymptotically accurate model must be locally and globally robust at all priors. I will now demonstrate that while this conjecture holds for local robustness, global robustness still requires the prior to be concentrated around the p-absorbing SCEs, referred to as a property of *prior tightness*. Theorem 2 establishes that in the absence of traps, prior tightness is a sufficient and necessary condition for global robustness given a specific prior. To state the result, let C^θ represent the set of parameters in θ that support at least one p-absorbing SCE, referred to as the set of *consistent* parameters in θ. In other words, for each $\omega \in C^\theta$, there exists a p-absorbing SCE under θ with supporting belief δ_ω.[28] Notice that C^θ is non-empty if and only if model θ admits a p-absorbing SCE .

Theorem 2. *Suppose $\alpha > 1$ and model θ has no traps, then the following are true:*

(i) Model θ is locally robust at all full-support priors if and only if $C^\theta \neq \emptyset$.

(ii) Model θ is globally robust at prior π_0^θ if and only if $C^\theta \neq \emptyset$ and $\pi_0^\theta(C^\theta) \geq 1/\alpha$.

(iii) Model θ is globally robust at all full-support priors if and only if $C^\theta = \Omega^\theta$.

Theorem 2 clarifies the fundamental distinction between the two notions of robustness: local robustness is prior-free, but global robustness is prior-dependent. Hence, limiting the maximal allowable step size of switching does not expand the set of robust models, but allows robust models to persist under more diverse priors. Specifically, local robustness at any single full-support prior automatically implies local robustness at all full-support priors. This, together with the insights from Theorem 1, underscores

[27]For example, it is straightforward to show that the artist's overconfidence model in Section 6 has no traps since each level of market demand corresponds to a different outcome distribution for all effort choices, and the model either induces a unique strict SCE.

[28]Note that when θ is identifiable, no parameters predict the same outcome distribution, and thus the supporting belief of any SCE must be pure.

that achieving perfect asymptotic accuracy, as indicated by $C^\theta \neq \emptyset$, is both a sufficient and necessary condition for a model to exhibit robustness when the agent engages in local exploration for an alternative model.

Theorem 2 provides a closed-form quantification of how concentrated the prior must be on C^θ in order to support global robustness. In particular, the tightness of the prior, quantified by $\pi_0^\theta(C^\theta)$, multiplied by the switching stickiness, quantified by the switching threshold α, must be weakly larger than 1. This relationship implies perfect substitutability between the roles of prior tightness and switching stickiness in facilitating model robustness. When switching is highly sticky and the agent demands substantial evidence for a switch, the prior tightness requirement has less bite—in fact, any asymptotically accurate model can be globally robust at any given full-support prior, provided that switching is sufficiently sticky. Conversely, when switching is relatively smooth and the agent requires minimal evidence for a switch, global robustness requires priors that are tightly centered around the set of SCEs.

Theorem 2 also suggests a trade-off when considering the impact of a model's size on its robustness. While including more predictions in a model increases the likelihood of achieving perfect asymptotic accuracy, it may also dilute the prior and result in a decrease in the prior probability assigned to consistent parameters. In fact, for any fixed $\alpha > 1$, global robustness holds at *all* priors if and only if every parameter in model θ induces a p-absorbing self-confirming equilibrium, $C^\theta = \Omega^\theta$. Models that exhibit the strongest form of robustness are those asymptotically accurate models that are simple enough for the number of p-absorbing SCEs they induce to be equal to the number of their total predictions. Conversely, when $C^\theta \neq \Omega^\theta$, global robustness fails at any given full-support prior, provided that the switching threshold is sufficiently close to 1.

Theorems 1 and 2 together draw an interesting comparison between misspecified models and correctly specified models in terms of their robustness properties. On one hand, all correctly specified models are locally robust at all priors and globally robust for at least one full-support prior since they have perfect asymptotic accuracy, which is achieved only by a subset of misspecified models. On the other hand, some misspecified models can be globally robust at more priors if they have a simple structure or induce a large set of SCEs. I further illustrate this comparison in Application 5.2.

Finally, I examine the scenario of perfectly non-sticky switching, i.e. $\alpha = 1$, and provide a characterization in Corollary 2. In this case, the existence of a p-absorbing SCE is no longer sufficient for either local robustness or global robustness for at least one full-support prior. Instead, we need every parameter in the model to induce a p-absorbing SCE.

Corollary 2. *Suppose model θ has no traps and $\alpha = 1$, then model θ is locally or globally robust for any full-support prior if and only if $C^\theta = \Omega^\theta$.*

Corollary 2 generates three immediate takeaways. First, the set of models that can be locally robust or globally robust shrinks discontinuously at $\alpha = 1$, highlighting the important role of switching stickiness in the persistence of misspecified models. Second, this result reveals an equivalence between two strong notions of robustness—global robustness when switching is non-sticky and global robustness at all priors—both of which are characterized by a simple condition, $C^\theta = \Omega^\theta$. Finally, the gap between local and global robustness has closed when switching is perfectly non-sticky, since both require perfect asymptotic accuracy and full prior concentration on p-absorbing SCEs.

Additionally, Corollary 2 and Theorem 2 (ii) suggest that a lower level of switching stickiness may not always benefit the agent. This observation follows from our previous finding that lowering α makes prior tightness (hence global robustness) harder to attain for *any* model, whether correctly or incorrectly specified. In fact, assuming no traps, the *only* correctly specified model that is globally robust when $\alpha = 1$ is the true model θ^*. As α approaches 1, it becomes likely for the agent to switch from a correctly specified model to a misspecified alternative model due to heightened sensitivity to initial noisy information, potentially resulting in the agent getting stuck with the misspecified model.[29]

4.5 Proof idea of Theorem 2

To understand why Theorem 2 is true, we first apply the theorem to the simplest case where outcomes are exogenously generated, specifically when $\mathcal{A}$ consists of a single action $\bar{a}$. Since any DGP now reduces to a single outcome distribution, Theorem 2 leads to a simple prediction: a model is locally robust across all priors if and only if its predictions contain the true outcome distribution; it is globally robust at a given prior if and only if the prior assigns probability weakly higher than $1/\alpha$ to the true outcome distribution. Example 2 illustrates the logic behind this result.

Example 2 (Exogenous data)**.** Suppose the agent works on a single task $\mathcal{A} = \{\bar{a}\}$ and observes the failure/success of the task, $\mathcal{Y} = \{0, 1\}$. As in Example 1, the true DGP prescribes that the success rate is 0.5. The agent holds a subjective model θ, under

[29] Determining an *optimal* switching threshold requires an analysis of Type I and Type II errors in the presence of endogenously generated data, a task beyond the scope of this paper.

$q^{\theta}(1\|a,\omega)$	ω^1	ω^2
$\bar{a}$	0.5	0.3

$q^{\theta'}(1\|a,\omega)$	ω^1	ω^2
$\bar{a}$	0.5	0.3+ϵ

$q^{\theta^*}(1\|a,\omega)$	ω^*
$\bar{a}$	0.5

Table 3: Initial model θ and competing model θ' in Example 2.

which he presumes the success rate depends on his luck $\omega \in \Omega^{\theta} = \{\omega^1, \omega^2\}$, where ω^1 represents good luck and ω^2 represents bad luck (see Table 3). Note that model θ is correctly specified because it predicts the true success rate of 0.5 for good luck. Corollary 1 implies that θ is both locally and globally robust for at least one prior.

As predicted by Theorem 2, model θ is locally robust at all full-support priors. To see why, assume the agent entertains a nearby competing model θ' that predicts a success rate of $0.3 + \epsilon$ for bad luck. When $\epsilon > 0$ is small, model θ' is slightly more optimistic than model θ. It is straightforward to calculate the Bayes factor in the beginning of period t,

$$\frac{\ell_t(\theta')}{\ell_t(\theta)} = \frac{\pi_0^{\theta'}(\omega^1)0.5^t + \pi_0^{\theta'}(\omega^2)(0.3+\epsilon)^{S_t}(0.7-\epsilon)^{F_t}}{\pi_0^{\theta}(\omega^1)0.5^t + \pi_0^{\theta}(\omega^2)0.3^{S_t}0.7^{F_t}}, \tag{9}$$

where S_t and F_t are the number of successes and failures observed before period t and $S_t + F_t = t$. As the agent accumulates more evidence, the likelihood associated with bad luck vanishes as compared to that associated with good luck in both models. In the limit, the Bayes factor converges to the ratio of the prior odds of good luck, $\pi_0^{\theta'}(\omega^1)/\pi_0^{\theta}(\omega^1)$, which is bounded above by $\alpha > 1$ when the priors are sufficiently close. This argument can be generalized to other nearby competing models to prove that model θ is locally robust at all full-support priors.

However, model θ is globally robust at a given prior if and only if the prior assigns sufficiently high probability to good luck, $\pi_0^{\theta}(\omega^1) \geq 1/\alpha$. To illustrate, suppose the competing model is the true model θ^*. The Bayes factor is given by

$$\lambda_t = \frac{\ell_t(\theta^*)}{\ell_t(\theta)} = \frac{0.5^t}{\pi_0^{\theta}(\omega^1)0.5^t + \pi_0^{\theta}(\omega^2)0.3^{S_t}0.7^{F_t}}. \tag{10}$$

If $\pi_0^{\theta}(\omega^1) < 1/\alpha$, the agent must eventually abandon model θ. As the likelihood of bad luck diminishes in comparison to that of good luck, λ_t converges to $1/\pi_0^{\theta}(\omega^1)$, which is strictly larger than α. Consequently, the competing model almost surely replaces the initial model, rendering it not globally robust. In contrast, when $\pi_0^{\theta}(\omega^1) \geq 1/\alpha$, λ_t is bounded above by α for any history, preventing the agent from switching to the competing model. Intuitively, this is because the explanatory power of model θ is at least $1/\alpha$ times as large as the true model, yet a switch is triggered only when the Bayes factor is strictly larger than α. The argument for other competing models is analogous, indicating that model θ is globally robust if and only if $\pi_0^\theta(\omega^1) \geq 1/\alpha$.

The key force driving the necessity of prior tightness for global robustness lies in the Bayes factor rule acting akin to *Occam's Razor*—it favors parsimonious models with tight priors while penalizing complex models with diffuse priors. In the case of exogenously generated data, conditional on having the same asymptotic accuracy, model θ with a diffuse prior can fit much worse than a "simple and accurate" competing model for which the prior is concentrated around the true DGP. When the agent exhibits mild switching stickiness, the competing model is sufficiently more attractive than the initial model due to its simplicity, leading to a permanent switch to the competing model. As switching becomes stickier, the agent's tolerance for diffuse priors also increases. In contrast, when the agent compares a model with its local perturbation, their priors are similarly tight around similar DGPs. Sticky switching then implies that the Bayes factor remains under the switching threshold with positive probability.

When outcomes are endogenously generated with multiple available actions, perfect asymptotic accuracy is achieved not only by the true model but also by models that induce at least one p-absorbing self-confirming equilibrium. Hence, an analogous prior tightness condition is that $\pi_0^\theta(C^\theta) \geq 1/\alpha$. The proof of Theorem 2 establishes the validity of this generalization, overcoming two complications that arise due to endogenous data. First, the parameters in C^θ represent data-generating processes, mapping actions to potentially distinct outcome distributions. Hence, these parameters may not accurately predict the true outcome distribution when non-self-confirming equilibrium actions are played. As a result, we cannot directly establish bounds for the Bayes factor using the prior ratio $\pi_0^{\theta'}(C^\theta)/\pi_0^\theta(C^\theta)$ for arbitrary action histories as in Example 2. Second, different parameters in C^θ may support different self-confirming equilibria to which the agent may converge. It remains unclear whether the prior should concentrate around one of the parameters in C^θ or the entire set.

I prove the necessity of prior tightness for global robustness by construction. Suppose model θ is globally robust at a full-support prior π_0^θ. Instead of constructing a "simple and accurate" competing model in the case of exogenous data, here we construct a potentially "extreme and misleading" competing model that replaces the initial model almost surely if prior tightness fails. Consider a competing model θ' characterized by a parameter space $\Omega^{\theta'} = C^\theta \cup \{\omega^*\}$, where any $\omega \in C^\theta$ represents the same

DGP that it corresponds to in model θ, and ω^* represents the true DGP. Let the prior $\pi_0^{\theta'}$ assign probabilities to parameters within C^θ in proportion to those assigned by π_0^θ, while assigning probability $\epsilon > 0$ to ω^*. Note that for each parameter $\omega \in C^\theta$, the ratio between $\pi_0^{\theta'}(\omega)$ and $\pi_0^\theta(\omega)$ is close to $1/\pi_0^\theta(C^\theta)$ when ϵ is sufficiently small. Since the competing model is correctly specified, on the paths where the model choice eventually converges to the initial model, the agent must end up playing a self-confirming equilibrium, and thus her belief π_t^θ must end up concentrating on one parameter within C^θ, say $\hat{\omega}$. Therefore, intuitively, the ratio of the explanatory power of model θ' and θ—i.e. the Bayes factor—is asymptotically bounded below by the ratio of the prior probabilities assigned to $\hat{\omega}$, approximately given by $1/\pi_0^\theta(C^\theta)$. However, when the prior tightness condition fails, this ratio exceeds α, implying that the Bayes factor must surpass α, leading to a contradiction.

5 Applications

In this section, I present two applications to demonstrate how the main results uncover new insights about robust misspecified models. The first application revisits the comparison between over- and underconfidence in more general environments. The second application illustrates that simple misspecified worldviews may outperform more complex correct worldviews in a political context.

5.1 Overconfidence and Underconfidence

A wealth of evidence in psychology and economics suggests that overconfidence is more prevalent than underconfidence. A leading explanation for this phenomenon is that individuals derive ego utility from holding overconfident beliefs about their positive traits (Brunnermeier and Parker, 2005; Kőszegi, 2006; Oster, Shoulson, and Dorsey, 2013). In this application, I compare the robustness properties of over- and underconfidence in broadly defined environments with competing models. I restrict attention to the prior-free local robustness notion since the interesting difference between over- and underconfidence only concerns the induced equilibria rather than the prior. The results demonstrate that, under natural assumptions, any degree of overconfidence is locally robust, whereas underconfidence is not locally robust except on a union of unconnected intervals. This result breaks the symmetry between over- and underconfidence and provides a novel mechanism rooted in the learning environment itself as to why we might expect one bias to be more common than the other.

As in Section 2, consider an agent selecting an action a_t from a finite set $\mathcal{A} \subset [\underline{a}, \overline{a}]$ in each period. The agent receives and observes a flow payoff $u(a_t, y_t) = g(a_t, b^*, \omega^*) + \eta_t$, where function g is twice continuously differentiable, strictly increasing in b and ω, $b^* \in [\underline{b}, \bar{b}]$ represents the agent's ability, and $\omega^* \in [\underline{\omega}, \overline{\omega}]$ captures an environment fundamental, such as market demand or organizational quality. The noise term η_t follows a known zero-mean distribution. I assume that g is strictly concave in a over $[\underline{a}, \overline{a}]$ ($g_{aa} < 0$), and that the action and the fundamental are either always strict complements or always strict substitutes. Formally, either $g_{a\omega} < 0$ or $g_{a\omega} > 0$ for all $a \in [\underline{a}, \overline{a}]$, $b \in [\underline{b}, \bar{b}]$, $\omega \in [\underline{\omega}, \overline{\omega}]$. Building on Heidhues et al. (2018) and Ba and Gindin (2022), I assume that the impact of one's ability on optimal effort differs in direction from the impact of the fundamental. Specifically, $\text{sgn}(g_{ab}) \neq \text{sgn}(g_{a\omega})$. This assumption plays a critical role in determining the robustness properties of over- and underconfidence.

I examine misspecified models that assign probability 1 to some $\hat{b} \in [\underline{b}, \bar{b}]$, deviating from its correct value b^*. The agent is dogmatically overconfident about his ability when $\hat{b} > b^*$ and underconfident when $\hat{b} < b^*$. To avoid trivial cases of non-robustness, I focus on models whose parameter spaces are *complete*: if the model assigns probability 1 to $\hat{b}$, then for *any* $a \in \mathcal{A}$, there exists $\Omega^\theta(a) \in \Omega^\theta$ such that $g(a, \hat{b}, \Omega^\theta(a)) = g(a, b^*, \omega^*)$.In other words, the agent can always identify a fundamental value that perfectly explains the observed data for any fixed a. I use $\Theta^M \subset \Theta$ to denote the set of all models satisfying the above conditions. Proposition 1 shows that while any level of overconfidence is locally robust, underconfidence can only be locally robust on unconnected intervals.

Proposition 1. *The following statements are true:*

(i) Model $\theta \in \Theta^M$ with any level of overconfidence $\hat{b} > b^*$ is locally robust.

(ii) There exists a strictly decreasing sequence $\beta_N < ... < \beta_1 < \beta_0 = b^*$ such that, model $\theta \in \Theta^M$ with underconfidence $\hat{b} < b^*$ is locally robust if $\hat{b} \in (\beta_{2k+1}, \beta_{2k})$ for any $k \in \mathbb{N}$ and not locally robust if $\hat{b} \in (\beta_{2k}, \beta_{2k-1})$ for some $k \in \mathbb{N}_+$.

Proposition 1 holds because any model in Θ^M induces at least one p-absorbing SCE if $\hat{b} > b^*$ while this is not guaranteed if $\hat{b} < b^*$. The key determinant in the existence of a p-absorbing SCE is the direction of belief reinforcement. To illustrate, suppose the action and the fundamental are strict complements, $g_{a\omega} > 0$. In this case, higher degenerate beliefs about the fundamental motivate higher actions, i.e. $\max A^\theta(\delta_{\omega'}) \leq \min A^\theta(\delta_{\omega''})$ for all $\omega'' > \omega'$. When the agent is overconfident with $\hat{b} > b^*$, higher actions lead to even

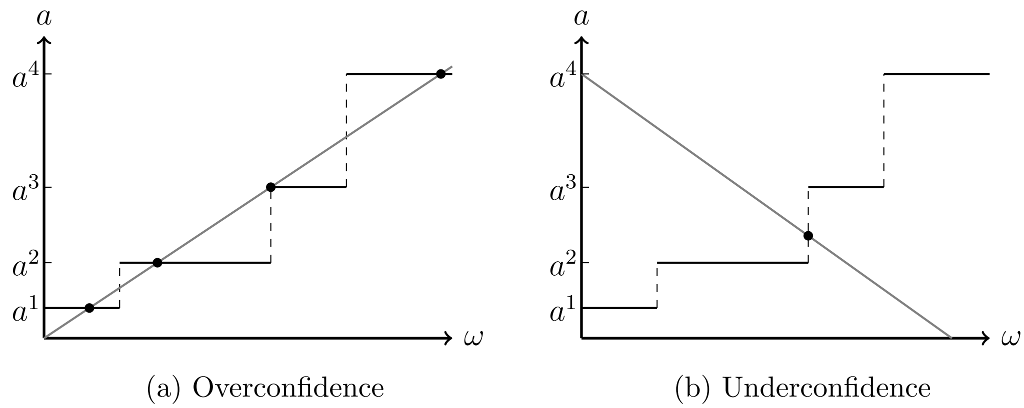

Figure 2: An example illustration of the model-induced equilibria when a and ω are strict complements. In both sub-figures, the black step function represents the agent's myopically optimal action as a function of the fundamental, the gray curve represents the agent's inferred fundamental $\Omega^\theta(a)$ as a function of the action, and the black dots represent potential long-run steady states. In the left figure, there are four self-confirming equilibria and three of them are quasi-strict (except a^3); in the right figure, there is no self-confirming equilibrium.

higher beliefs about the fundamental, i.e. $\Omega^\theta(a') < \Omega^\theta(a'')$ for all $a' < a''$, positively reinforcing the distortion. This relationship can be observed from the equation below,

$$g(a, \hat{b}, \Omega^\theta(a)) = g(a, b^*, \omega^*). \tag{11}$$

At a higher action, the return to fundamental ω is higher because $g_{a\omega} > 0$ and the return to ability b is weakly lower because, by assumption, $\text{sgn}(g_{ab}) \neq \text{sgn}(g_{a\omega})$. Therefore, the positive gap between the true fundamental ω^* and the inferred fundamental $\Omega^\theta(a)$ should be smaller such that expectations meet the reality, implying that $\Omega^\theta(a)$ is larger. In contrast, when the agent is underconfident with $\hat{b} < b^*$, higher actions in turn lead to lower beliefs, i.e. $\Omega^\theta(\delta_{a'}) > \Omega^\theta(\delta_{a''})$, negatively reinforcing the distortion.

As shown in Fig. 2, the optimal action is an increasing step function of the belief about the fundamental. Meanwhile, the inferred fundamental is a strictly increasing function of the action in cases of overconfidence and a strictly decreasing function in cases of underconfidence. With overconfidence, the optimal action curve and the inference curve must intersect at least once at a flat segment of the optimal action curve. This together with the assumption of a complete parameter space ensures the existence of at least one p-absorbing SCE. In contrast, with underconfidence, the optimal action curve and the inference curve may intersect at the vertical segment of the optimal action curve. This point of intersection corresponds to a steady state where the agent mixes two actions with a fixed frequency, and his belief eventually converges to the inferred value of the fundamental that best explains the data at the mixed action. Notably, however, this steady state is not a self-confirming equilibrium because the agent's belief about the fundamental cannot perfectly explain the data since it is generated by two distinct actions.[30] When the action and the fundamental are strict substitutes ($g_{a\omega} < 0$), the orientation of both the optimal action curve and the inferred fundamental curve is inverted, so Proposition 1 still applies.

Remark. While the condition $\text{sgn}(g_{ab}) \neq \text{sgn}(g_{a\omega})$ is sufficient for Proposition 1, it is not necessary. The result may still hold in cases where $\text{sgn}(g_{ab}) = \text{sgn}(g_{a\omega})$, but verifying that overconfidence is more robust than underconfidence in this scenario requires a case-by-case analysis of the direction of belief reinforcement, i.e. whether the inferred fundamental function is co-monotone with the optimal action function.[31]

5.2 Media Bias, Extremism, and Polarization

In this application, I consider a stylized model of media consumption and demonstrate how misconceptions about media bias (Groseclose and Milyo, 2005) can lead to stable polarization in political views despite no individual partisan bias. The misspecified model with these misconceptions and an extremism bias is globally robust regardless of the initial conditions. Even more surprisingly, people may abandon a correctly specified model, switch to such a misspecified model and then get stuck forever.

The agent has access to three media outlets and in each period she chooses one to consume, $\mathcal{A} = \{a^L, a^M, a^R\}$. The media outlets are indexed by their political leanings, left-wing, neutral, or right-wing. Each media outlet delivers two types of news, $\mathcal{Y} = \{l, r\}$, where l represents good stories for the leftists and r represents good stories for the rightists. The unknown state of the world $\omega \in \Omega = \{\omega^L, \omega^M, \omega^R\}$ governs the fraction of l and r stories happened in the real world and it remains fixed throughout the life of the agent. In particular, 60% of the stories are l stories (r stories) in state

[30]This steady state corresponds to a Berk-Nash equilibrium, where the equilibrium belief about the fundamental minimizes the weighted average Kullback-Leibler divergence of the true outcome distribution from the model prediction.

[31]For example, suppose the output g takes the form of $h(a)b + a\omega - c(a)$, where h and c are strictly increasing. If $h(a) = ka$ for some $k > 0$, then $\Omega^\theta(a)$ is independent of a for any $\hat{b} \neq b$, implying that both over- and underconfidence are locally robust. If $h(a) = ka^n$ for some $k > 0$ and $n < 1$, then it can be verified from Eq. (11) that $\Omega^\theta(a)$ is co-monotone with the optimal action function when $\hat{b} > b$, and not co-monotone if $\hat{b} < b$. In this case, Proposition 1 still applies. In contrast, if $h(a) = ka^n$ for some $k > 0$ and $n > 1$, then the opposite of Proposition 1 holds, i.e. underconfidence is more robust than overconfidence.

$q^{\theta}(l\|a,\omega)$	ω^L	ω^M	ω^R
a^L	0.7	0.6	0.5
a^M	0.6	0.5	0.4
a^R	0.5	0.4	0.3

$q^{\hat{\theta}}(l\|a,\omega)$	ω^L	ω^R
a^L	0.6	0.5
a^M	0.5	0.5
a^R	0.5	0.4

Table 4: The left panel summarizes the true fraction of l stories reported by each media outlet in each state of the world. It is also a description of the correctly specified model θ. The right panel describes the predictions of a misspecified model $\hat{\theta}$.

ω^L (ω^R), while an equal share of l and r stories happen in state ω^M. The three media outlets differ in their ways of news reporting: in each state of the world, media a^M truthfully reports the stories without bias, media a^L selectively reports l more than media a^M, and media a^R selectively reports r more than media a^M. The left panel of Section 5.2 summarizes the true fraction of l stories reported by the media in different states. We restrict attention on the world in state M, where the true fractions of l stories reported by the three media are given by $(0.6, 0.5, 0.4)$.

In this exercise, we focus attention on the comparison between two different models θ and $\hat{\theta}$ that I describe in Section 5.2. Model θ is correctly specified: a θ-modeler realizes that ω^M is a possible state of the world and are fully aware of the bias of both the left-wing and the right-wing media outlets. By contrast, model $\hat{\theta}$ is misspecified in two aspects. First, $\hat{\theta}$ features *extremism* because it only recognizes the possibility of the extreme states ω^L and ω^R. Second, $\hat{\theta}$ features *naivety* about media bias: a $\hat{\theta}$-modeler underestimates the selective reporting bias of the left-wing a^L and right-wing media a^R, and also underestimates the informativeness of the neutral media. As a result, when a $\hat{\theta}$-modeler subscribes to the left-wing media and finds that 60% of the stories are good stories for leftists, she does not interpret it as evidence for the middle state ω^M (which does not exist in her extreme worldview), but treats it as evidence for the left state ω^L; a similar logic applies to the right-wing media. She also mistakenly thinks that the reporting of the neutral media is totally uninformative about the state.

To highlight the core mechanism, I abstract away from specifying the payoff structure and outline the minimal assumptions that allow us to apply the characterization theorems in Section 4. It is straightforward to verify using Section 5.2 that the SCE supporting beliefs mentioned in Assumption 3 are indeed consistent.

Assumption 3. *When the true state is ω^M, the following are true:*

(i) Model θ admits a unique SCE a^M and it is strict, supported by belief δ_{ω^M}.

(ii) Model $\hat{\theta}$ admits only two strict SCEs, a^L and a^R, supported by δ_{ω^L} and δ_{ω^R}, respectively.

Assumption 3 is natural and intuitive. With the correctly specified model θ, the agent infers the true state and subscribes to the neutral media. With the misspecified model $\hat{\theta}$, however, the agent develops partisan bias and only subscribes to the media biased towards her political belief. The choices of the agent can be justified as the result of maximizing the sum of emotional and informational value from news consumption.[32]

Model $\hat{\theta}$ has an advantage over model θ due to its extremeness. Since both models θ and $\hat{\theta}$ admit at least one SCE, Theorem 1 tells us that both models are globally robust at some prior. Interestingly, Theorem 2 implies a counter-intuitive result (see Proposition 2 below): model θ is globally robust only when the associated prior assigns high enough probability to the true state ω^M, while model $\hat{\theta}$ is globally robust at *all* priors. In other words, model $\hat{\theta}$ is globally robust in a *robust* way.

Proposition 2. *Fix any $\alpha > 1$. Model θ is globally robust at prior π_0^θ if and only if $\pi_0^\theta(\omega^M) \geq 1/\alpha$, while model $\hat{\theta}$ is globally robust at all priors.*

Despite being misspecified, model $\hat{\theta}$ has a stronger global robustness property, because its narrative is simple, coherent, and balanced. To see this, notice that all parameter values in model $\hat{\theta}$ are consistent, i.e. $C^{\hat{\theta}} = \{\omega^L, \omega^R\} = \Omega^{\hat{\theta}}$. This makes it possible for model $\hat{\theta}$ to persist against any competing model. For example, model $\hat{\theta}$ can outperform a left-biased competing model in explaining the data when the agent happens to read a series of r stories, and similarly it can outperform a right-biased competing model when the agent happens to read a series of l stories. If the competing model is unbiased and correctly specified such as model θ, model $\hat{\theta}$ still persists because of its simplicity.

Proposition 2 characterizes the robustness properties of θ and $\hat{\theta}$ separately with the implicit assumption that they are the initial model choice of a switcher. What if a switcher originally adopts θ and entertains $\hat{\theta}$ as the competing model? Whether she will abandon θ in favor of $\hat{\theta}$ is *a priori* unclear. While θ may not be globally robust at a given prior, this only tells us that θ does not persist against some competing model, but this competing model may not be $\hat{\theta}$. Surprisingly, as I show in Proposition 3, $\hat{\theta}$ indeed replaces θ with positive probability if the switching threshold is low.

[32] A micro-foundation is provided in Appendix C.

Proposition 3. *Fix any full-support priors $\pi_0^\theta, \pi_0^{\hat\theta}$ and any $\alpha < 1/\pi_0^\theta(\omega^M)$. In the switcher's problem $(\theta, \hat\theta, \pi_0^\theta, \pi_0^{\hat\theta})$, the model choice m_t eventually equals $\hat\theta$ with positive probability.*

In summary, this application generates three novel insights about news consumption and political beliefs. First, extremism and naivety about media bias go hand in hand and their persistence is robust against arbitrary competing narratives. Second, individuals may abandon their correct models and switch to incorrect alternatives because of their extremeness/simplicity. Third, even though the extreme and naive model has no built-in political bias, individuals who hold such a model gradually develop a strong partisan bias over time. The direction of the partisan bias is random and path-dependent, leading to long-term political polarization.

6 Extensions

6.1 Multiple Competing Models

I now relax the assumption of the agent entertaining exactly one competing model and characterize robust models allowing for multiple competing models. The framework described in Section 3 can be easily extended to accommodate this modification. Let $\Theta' \subseteq \Theta$ denote the set of competing models that the agent considers, and $\Theta^\dagger := \Theta' \cup \theta$ denote the set of all models considered, including the initial model. Throughout, I maintain the assumption that Θ' is finite and contains at most $K \geq 1$ distinct models. At the beginning of period t, the agent compares her current model against all alternative models and switches to the most plausible one if the corresponding Bayes factor exceeds the switching threshold α. Specifically, the agent calculates the Bayes factors between models in $\Theta^\dagger$ and the model she used in the last period, $\boldsymbol{\lambda}_t := (\lambda_t^{\theta'})_{\theta' \in \Theta^\dagger}$, where

$$\lambda_t^{\theta'} = \ell_t(\theta')/\ell_t(m_{t-1}). \tag{12}$$

The agent makes a switch if $\max_{\theta' \in \Theta^\dagger} \lambda_t^{\theta'} > \alpha$ and switches to the model θ' with the highest Bayes factor. Let $\pi_0^{\Theta'}$ denote a vector of priors for the competing models in Θ'. The definitions of persistence and robustness can be modified by simply replacing θ' with Θ' and $\pi_0^{\theta'}$ with $\pi_0^{\Theta'}$. That is, model θ is globally robust at prior π_0^θ if it persists against every $\Theta' \subseteq \Theta$ of size no larger than K at π_0^θ and every vector of priors $\pi_0^{\Theta'}$; it is locally robust at π_0^θ if it persists whenever each competing model in Θ' is sufficiently close to θ, along with their corresponding priors.[33]

Introducing multiple competing models leads to overfitting. Specifically, when the number of competing models K exceeds one plus the switching threshold α, even the true model fails to be globally robust, as demonstrated in Example 3.[34]

Example 3 (Overfitting)**.** Consider an agent who repeated chooses between two actions, $\mathcal{A} = \{a^1, a^2\}$. The true DGP is a uniform distribution over K outcomes, $\mathcal{Y} = \{1, ..., K\}$, regardless of the chosen action. When the outcome is $y = 1$, the agent incurs a loss of $-K$, while receiving a payoff of 0 for any other outcome. The agent pays an additional cost $c > 0$ for playing a^1 and no cost if she plays a^2. Assuming that the agent's initial model θ is the true model θ^*, she optimally plays a^2 in the first period to avoid the cost. Suppose the agent evaluates K competing models that I describe below. Each model $\theta^k \in \{\theta^1, ..., \theta^K\}$ has a single parameter ω^k. When a^1 is played, model θ^k agrees with θ, correctly predicting a uniform outcome distribution. When a^2 is played, model θ^k diverges from θ. Specifically, for any $k > 1$, θ^k predicts

$$q^{\theta^k}(y|a^2, \omega^k) = \begin{cases} 1 - \frac{1}{K} - (K-1)\eta & \text{if } y = n, \\ \frac{1}{K} + \eta & \text{if } y = 1, \\ \eta & \text{if } y \in \mathcal{Y} \setminus \{1, n\}, \end{cases}$$

where η is a small positive constant. When $k = 1$, $q^{\theta^k}(\cdot|a^2, \omega^k)$ is given by

$$q^{\theta^1}(y|a^2, \omega^1) = \begin{cases} 1 - (K-1)\eta & \text{if } y = 1, \\ \eta & \text{if } y \in \mathcal{Y} \setminus \{1\}. \end{cases}$$

Note that model θ^k predicts that when a^2 is played, the outcome k is drawn with probability near 1. Given there is one such model for every possible outcome, the agent must switch to one of these competing models upon the first outcome realization when η is sufficiently small. If the realized outcome is k, the agent immediately switches to model θ^k when

$$\frac{\ell_1(\theta^k)}{\ell_1(\theta)} \geq \frac{1 - \frac{1}{K} - (K-1)\eta}{\frac{1}{K}} > \alpha.$$

[33] Note that if Θ' is not a singleton, then persistence against Θ' is not equivalent to persistence against each model in Θ', and neither implies the other. For examples, refer to Appendix C.

[34] Schwartzstein and Sunderam (2021) find that in a static setting, a decision-maker shifts from the true model when a persuader is permitted to propose an alternative model after the data is realized. In contrast, my results suggest that the persuader can achieve the same objective even if they are required to propose a model before the data is realized, and the data is infinite, as long as they can present multiple competing models.

Note that such η exists as $K > \alpha + 1$. Furthermore, since every competing model assigns a probability larger than $1/K$ to the outcome $y = 1$, once the switch occurs, the agent finds it optimal to play a^1 to avoid the loss associated with outcome 1 when c is sufficiently small. However, since all models yield the same correct predictions under a^1, the Bayes factors $\boldsymbol{\lambda_t}$ remain constant thereafter. Hence, despite starting with the true model, the agent becomes permanently trapped with a wrong model and chooses a suboptimal action.

The trap described in Example 3 differs from the trap constructed in Example 1. Note that θ^* satisfies the no-trap conditions in Definition 7 since it is identifiable and induces a quasi-strict SCE. The agent in Example 3 becomes trapped because a^1 is strictly dominant under the competing models, and all models have the exact same predictions once a^1 is being played, eliminating any possibility of future learning and switching. However, the driving force that leads the agent into the trap indeed stems from overfitting in both examples. In Example 1, the agent holds a diffuse prior, rendering his model choice to be sensitive to early outcome realizations. Similarly, in Example 3, overfitting arises in the short term and triggers an early switch to other models. The likelihood of such a switch increases as the agent evaluates more competing models. Therefore, a natural remedy is to make switching stickier so that the agent is less responsive to early outcome realizations. Indeed, Theorem 3 shows that if $\alpha > K$, the scope of local and global robustness does not change and both are still characterized by perfect asymptotic accuracy.

Theorem 3. *Suppose the agent evaluates at most K competing models and $\alpha > K$. Model θ is locally and globally robust for at least one prior if and only if there exists a p-absorbing SCE under θ.*

6.2 Non-myopic Agent

Our baseline framework focuses on a myopic agent and rules out any experimentation motives. This assumption can be less substantial than one might think. In this subsection, I discuss two potential ways of relaxing this assumption.

First, we may assume that the agent is non-myopic within each model but maintain that she is myopic across models. That is, when choosing an optimal action, the agent maximizes her expected discounted sum of payoffs assuming that she keeps her current model m_t in the future. An optimal policy f^θ solves the following dynamic programming problem,

$$U^{\theta}\left(\pi_t^{\theta}\right)=\max_{a\in\mathcal{A}}\sum_{\omega\in\Omega^{\theta}}\pi^{\theta}\left(\omega\right)\int_{y\in\mathcal{Y}}\left[u\left(a,y\right)+\delta U^{\theta}\left(B^{\theta}\left(a,y,\pi_t^{\theta}\right)\right)\right]q^{\theta}\left(y|a,\omega\right)v\left(dy\right).$$

How should we interpret the asymmetry between experimentation within models and no experimentation across models? This asymmetry again highlights the stickiness of switching models as opposed to the smoothness of Bayesian updating, and it is plausible when resources are constrained. For instance, consider an applied data scientist who uses one single model to guide data collection and make policy recommendations. While he is aware of potential misspecification, he may choose not to spare valuable resources in additional experiments to find the best model. However, he may indeed switch to a different model if the data at hand happens to suggest its superiority.[35]

If we relax the myopicity assumption this way, Theorem 1 goes through without changes. This claim may appear surprising at first, because experimentation motives should make it harder to sustain a self-confirming equilibrium or a Berk-Nash equilibrium, and thus the set of robust misspecified models might be smaller if the agent is non-myopic. This intuition is correct—as the agent becomes more patient, p-absorbingness is harder to achieve. However, note that the theorems only establish the equivalence relationship between the existence of p-absorbing equilibria and the models' robustness properties, so whether p-absorbingness can be achieved is irrelevant. In the Appendix, I provide stronger sufficient conditions for p-absorbingness such that variants of Corollary 1 continue to hold.[36]

Alternatively, we may assume the agent is forward-looking both within and across models. If the agent anticipates future model switches, she may intentionally take actions that allow her to distinguish different models, even if her current model predicts a different optimal action. Characterizing robust models in this environment is significantly more challenging and beyond the scope of this paper. I conjecture that the set of robust misspecified models will shrink as the agent becomes increasingly patient.

[35] This assumption is also natural in organizations where decision making and model estimation are handled by separate teams. For example, a manager (e.g. the chairman of a central bank) chooses policies based on the predictions made by the research team (e.g. a group of macroeconomists), while the research team focuses on estimating the models given the available data.

[36] In particular, any uniformly quasi-strict SCE is p-absorbing, and so is any uniformly strict BN-E.

6.3 Alternative Definitions of Persistence

This subsection discusses alternative definitions of model persistence. Our definition of persistence in Section 3 requires that if a switcher initially adopts this model, she eventually settles down with it with positive probability. This concept has a natural interpretation and can be used to predict whether a particular bias is likely to exist in a stable form. However, by relaxing or strengthening different parts of this definition, we can obtain a couple of variants that are also worth exploring. These alternative definitions are useful if one looks for models with stronger or weaker persistence properties. This investigation also helps us better understand the original definition since it sheds light on the importance of the different parts of the concept.

Almost sure eventual adoption. The first natural extension is to strengthen persistence by requiring that the model is eventually adopted with probability 1. That is, any such model is guaranteed to win out in the competition. But almost-sure persistence turns out to make global and local robustness impossible. In fact, given any model θ (including the true model θ^*), we can easily construct a nearby competing model θ' such that the competing model is eventually adopted with positive probability. The idea is that the agent can draw a sequence of outcome realizations that can be better explained by the competing model, and once a switch happens, the agent does not feel compelled to switch back since the predictions of the two models are identical in the limit. Therefore, almost sure eventual adoption can be too restrictive to be any useful.[37]

No switch. The current definition of persistence allows for back-and-forth switching before the eventual adoption of the model. A more conservative definition could have required the agent to adopt the same model throughout. It turns out that all theorems continue to hold even with this conservative definition. Indeed, model switching does not play any role in ensuring local and global robustness when there are no traps—in the proof of all main theorems, I show that there exists a sequence of outcome realizations that induce the agent to play a SCE while remaining under the same model. When there

[37]To do this, let us construct θ' such that it contains all DGPs in θ and one additional DGP that differs from any other DGPs for all actions in θ. That is, we have $\Omega^{\theta'} = \Omega^{\theta} \cup \{\hat{\omega}\}$, where $q^{\theta'}(\cdot|a,\omega) = q^{\theta}(a|\cdot,\omega)$ and $q^{\theta'}(\cdot|a,\hat{\omega}) \neq q^{\theta}(\cdot|a,\omega)$ for all $\omega \in \Omega^{\theta}$ and all actions $a \in \mathcal{A}$. In addition, let the prior π_0^{θ} be proportional to $\pi_0^{\theta'}$ for the all shared parameters. With this structure, the Bayes factor λ_t is bounded below by $\pi_0^{\theta'}(\Omega^{\theta})$. Note that since $\hat{\omega}$ predicts differently from model θ, it is always a positive-probability event that the agent finds model θ' particularly compelling and make a switch because of the existence of $\hat{\omega}$. But then the agent never switches back if the lower bound of the Bayes factor, $\pi_0^{\theta'}(\Omega^{\theta})$, is higher than $1/\alpha$, which can be achieved if we make $\pi_0^{\theta'}(\Omega^{\theta})$ sufficiently close to 1.

are traps in the model, however, a temporary switch to the competing model can be instrumental to the persistence of the initial model because switching to the competing model may happen to keep the agent away from the traps. A full characterization of this case is left for future research.

7 Concluding Remarks

In this paper, I propose a new theoretical framework to study the persistence of misspecified models when decision-makers are aware of potential model misspecification. I introduce sticky switching to the standard model of individual active learning and study the limit of the model choices. I explore local and global robustness and use them to derive novel insights about which models persist and when they persist. I show that both robustness notions can be characterized in terms of two properties, asymptotic accuracy and prior tightness.

The idea that the existence of self-confirming equilibria can explain why incorrect models persist can be found in the existing literature. Instead of assuming that the agent starts outright from an equilibrium, my framework incorporates full-fledged model switching dynamics into active learning processes. The characterization highlights the importance of this consideration. Robustness not only requires the existence of a self-confirming equilibrium but also needs it to be p-absorbing, which connects the notion of model persistence with the stability of equilibria. Furthermore, global robustness requires high prior tightness around the set of p-absorbsing self-confirming equilibria. This finding provides a theoretical justification for the empirical observation that simple narratives and entrenched worldviews tend to be more persistent.

The model-switching framework has great application value. My characterization of robust models, stated in the form of simple criteria that can be easily verified from the primitives, provides a learning foundation for certain misspecified models, some of which are already studied in misspecified learning literature. It can also be used to predict the persistence of given behavioral biases in specific contexts, which can be useful for guiding empirical work on behavioral economics and relevant policy making.

Within this general framework of model switching, there are many other interesting questions to pursue. For example, persistence requires a positive chance of eventual adoption, but this concept is silent about the size of this probability. New insights may emerge from studying how this probability is determined by key primitives of the model, such as whether it is correctly specified or misspecified, and features of the learning environment, such as the switching stickiness. Another potentially fruitful direction could be to restrict attention to a given small set of models and fully characterize the dynamics of model choices, i.e. how a decision-maker oscillates between two or more competing models persistently.

References

Aina, C. (2023): "Tailored Stories," *Mimeo.*

Akaike, H. (1974): "A New Look at the Statistical Model Identification," *IEEE Transactions on Automatic Control*, 19, 716–723.

Ba, C. and A. Gindin (2022): "A Multi-Agent Model of Misspecified Learning with Overconfidence," *Available at SSRN.*

Battigalli, P. (1987): "Comportamento Razionale ed Equilibrio nei Giochi e nelle Situazioni Sociali," *Unpublished thesis.*

Blackwell, D. (1965): "Discounted Dynamic Programming," *The Annals of Mathematical Statistics*, 36, 226–235.

Bohren, J. A. (2016): "Informational Herding with Model Misspecification," *Journal of Economic Theory*, 163, 222–247.

Bohren, J. A. and D. N. Hauser (2021): "Learning with Heterogeneous Misspecified Models: Characterization and Robustness," *Econometrica*, 89, 3025–3077.

Braungart, R. G. and M. M. Braungart (1986): "Life-Course and Generational Politics," *Annual Review of Sociology*, 12, 205–231.

Brunnermeier, M. K. and J. A. Parker (2005): "Optimal Expectations," *American Economic Review*, 95, 1092–1118.

Chernoff, H. (1954): "On the Distribution of the Likelihood Ratio," *The Annals of Mathematical Statistics*, 573–578.

Cho, I.-K. and K. Kasa (2015): "Learning and Model Validation," *The Review of Economic Studies*, 82, 45–82.

Di Stefano, G., F. Gino, G. P. Pisano, and B. Staats (2015): *Learning by Thinking: Overcoming the Bias for Action through Reflection*, Harvard Business School Cambridge, MA, USA.

Easley, D. and N. M. Kiefer (1988): "Controlling a Stochastic Process with Unknown Parameters," *Econometrica: Journal of the Econometric Society*, 1045–1064.

Eliaz, K. and R. Spiegler (2020): "A model of competing narratives," *American Economic Review*, 110, 3786–3816.

Esponda, I. and D. Pouzo (2016): "Berk–Nash Equilibrium: A Framework for Modeling Agents with Misspecified Models," *Econometrica*, 84, 1093–1130.

Esponda, I., D. Pouzo, and Y. Yamamoto (2021): "Asymptotic Behavior of Bayesian Learners with Misspecified Models," *Journal of Economic Theory*, 195, 105260.

Eyster, E. and M. Rabin (2005): "Cursed Equilibrium," *Econometrica*, 73, 1623–1672.

——— (2010): "Naive Herding in Rich-Information Settings," *American Economic Journal: Microeconomics*, 2, 221–43.

Fisher, W. R. (1985): "The Narrative Paradigm: An Elaboration," *Communications Monographs*, 52, 347–367.

Frick, M., R. Iijima, and Y. Ishii (2021): "Welfare Comparisons for Biased Learning," *Cowles Foundation Discussion Papers. 2605.*

——— (2023): "Belief convergence under misspecified learning: A martingale approach," *The Review of Economic Studies*, 90, 781–814.

Fudenberg, D., G. Lanzani, and P. Strack (2021): "Limit Points of Endogenous Misspecified Learning," *Econometrica*, 89, 1065–1098.

Fudenberg, D., G. Lanzani, et al. (2022): "Which Misspecifications Persist?" *Theoretical Economics.*

Fudenberg, D. and D. K. Levine (1993): "Self-Confirming Equilibrium," *Econometrica*, 523–545.

Fudenberg, D., G. Romanyuk, and P. Strack (2017): “Active Learning with a Misspecified Prior,” *Theoretical Economics*, 12, 1155–1189.

Gagnon-Bartsch, T. (2017): “Taste Projection in Models of Social Learning,” Tech. rep., Mimeo.

Gagnon-Bartsch, T., M. Rabin, and J. Schwartzstein (2020): *Channeled Attention and Stable Errors*, Harvard Business School.

Galperti, S. (2019): “Persuasion: The Art of Changing Worldviews,” *American Economic Review*, 109, 996–1031.

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

Groseclose, T. and J. Milyo (2005): “A Measure of Media Bias,” *The Quarterly Journal of Economics*, 120, 1191–1237.

Hansen, L. and T. J. Sargent (2001): “Robust Control and Model Uncertainty,” *American Economic Review*, 91, 60–66.

He, K. (2022): “Mislearning from Censored Data: The Gambler’s Fallacy and Other Correlational Mistakes in Optimal-Stopping Problems,” *Theoretical Economics*, 17, 1269–1312.

He, K. and J. Libgober (2020): “Evolutionarily Stable (Mis) specifications: Theory and Applications,” *arXiv preprint arXiv:2012.15007*.

Heidhues, P., B. Kőszegi, and P. Strack (2018): “Unrealistic Expectations and Misguided Learning,” *Econometrica*, 86, 1159–1214.

——— (2019): “Overconfidence and Prejudice,” *arXiv preprint arXiv:1909.08497*.

Jehiel, P. and F. Koessler (2008): “Revisiting Games of Incomplete Information with Analogy-Based Expectations,” *Games and Economic Behavior*, 62, 533–557.

Karni, E. and M.-L. Vierø (2013): “”Reverse Bayesianism”: A Choice-Based Theory Of Growing Awareness,” *American Economic Review*, 103, 2790–2810.

Kass, R. E. and A. E. Raftery (1995): “Bayes Factors,” *Journal of the American Statistical Association*, 90, 773–795.

Klibanoff, P., M. Marinacci, and S. Mukerji (2005): "A Smooth Model of Decision Making under Ambiguity," *Econometrica*, 73, 1849–1892.

Konishi, S. and G. Kitagawa (2008): *Information Criteria and Statistical Modeling*, Springer Science & Business Media.

Köszegi, B. (2006): "Ego Utility, Overconfidence, and Task Choice," *Journal of the European Economic Association*, 4, 673–707.

Kuhn, T. S. (1962): *The Structure of Scientific Revolutions*, University of Chicago Press.

Levy, G., R. Razin, and A. Young (2022): "Misspecified politics and the recurrence of populism," *American Economic Review*, 112, 928–962.

Mailath, G. J. and L. Samuelson (2020): "Learning under Diverse World Views: Model-Based Inference," *American Economic Review*, 110, 1464–1501.

Maitra, A. (1968): "Discounted Dynamic Programming on Compact Metric Spaces," *Sankhyā: The Indian Journal of Statistics, Series A*, 211–216.

Miller, D. T. and M. Ross (1975): "Self-Serving Biases in the Attribution of Causality: Fact or fiction?" *Psychological bulletin*, 82, 213.

Montiel Olea, J. L., P. Ortoleva, M. M. Pai, and A. Prat (2022): "Competing Models," *The Quarterly Journal of Economics*, 137, 2419–2457.

Murooka, T. and Y. Yamamoto (2023): "Higher-Order Misspecification and Equilibrium Stability," .

Nyarko, Y. (1991): "Learning in Mis-specified Models and the Possibility of Cycles," *Journal of Economic Theory*, 55, 416–427.

Ortoleva, P. (2012): "Modeling the Change of Paradigm: Non-Bayesian Reactions to Unexpected News," *American Economic Review*, 102, 2410–36.

Ortoleva, P. and E. Snowberg (2015): "Overconfidence in Political Behavior," *American Economic Review*, 105, 504–35.

Oster, E., I. Shoulson, and E. Dorsey (2013): "Optimal Expectations and Limited Medical Testing: evidence from Huntington disease," *American Economic Review*, 103, 804–30.

Rabin, M. and D. Vayanos (2010): "The Gambler's and Hot-Hand Fallacies: Theory and Applications," *Review of Economic Studies*, 77, 730–778.

Sargent, T. J. (1999): *The conquest of American inflation*, Princeton University Press.

Savage, L. J. (1972): *The Foundations of Statistics*, Courier Corporation.

Schwartzstein, J. and A. Sunderam (2021): "Using Models to Persuade," *American Economic Review*, 111, 276–323.

Schwarz, G. et al. (1978): "Estimating the Dimension of a Model," *The Annals of Statistics*, 6, 461–464.

Spiegler, R. (2016): "Bayesian Networks and Boundedly Rational Expectations," *The Quarterly Journal of Economics*, 131, 1243–1290.

——— (2019): "Behavioral Implications of Causal Misperceptions," *Annual Review of Economics*, 12.

——— (2020): "Can Agents with Causal Misperceptions be Systematically Fooled?" *Journal of the European Economic Association*, 18, 583–617.

Stone, M. (1977): "An Asymptotic Equivalence of Choice of Model by Cross-validation and Akaike's Criterion," *Journal of the Royal Statistical Society: Series B (Methodological)*, 39, 44–47.

A Auxiliary Lemmas

The underlying probability space $(Y, \mathcal{F}, \mathbb{P})$ is constructed as follows. The sample space is $\mathscr{Y} := (\mathcal{Y}^\infty)^{\mathcal{A}}$, each element of which consists of infinite sequences of outcome realizations $(y_{a,0}, y_{a,1}, ...)$ for all actions $a \in \mathcal{A}$, where $y_{a,t}$ denotes the outcome when the agent takes a in period t. Let us denote by $\mathbb{P}$ the probability measure over $\mathscr{Y}$ induced by independent draws from q^* and denote by $\mathcal{F}$ the product sigma algebra. Let $h := (a_t, y_t)_{t=0}^{\infty}$ denote an infinite history and $H := (\mathcal{A} \times \mathcal{Y})^\infty$ be the set of infinite histories. Combined with the switching threshold α, the switcher's problem $(\theta, \theta', \pi_0^\theta, \pi_0^{\theta'})$, and policies $(f^\theta, f^{\theta'})$, $\mathbb{P}$ induces a probability measure over H when the agent is a switcher, denoted by $\mathbb{P}_S$. Meanwhile, the measure $\mathbb{P}$, prior π_0^θ, and policy f^θ induce a different probability measure over H for a θ-modeler who uses the same prior and policy, denoted by $\mathbb{P}_D$. All probabilistic statements about a switcher are made with respect to $\mathbb{P}_S$ and all those about a θ-modeler are with respect to $\mathbb{P}_D$, unless indicated otherwise.

Lemma 3. *Consider any switcher's problem $(\theta, \theta', \pi_0^\theta, \pi_0^{\theta'})$ in which $\theta, \theta' \in \Theta$ and θ' is correctly specified. The ratio $\ell_t(\theta)/\ell_t(\theta')$ a.s. converges to a non-negative random variable with finite expectation.*

Proof. Let $\kappa_t = \ell_t(\theta)/\ell_t(\theta')$, then $\kappa_0 = 1$ and $\kappa_t \geq 0, \forall t$. I now construct the probability space in which κ_t is a martingale. Given prior $\pi_0^{\theta'}$, denote by $\mathbb{P}_S^{\theta'}$ the probability measure over the set of histories H as implied by model θ'. Formally, for any $\hat{H} \subseteq H$, we have $\mathbb{P}_S^{\theta'}\left(\hat{H}\right) = \sum_{\omega \in \Omega^{\theta'}} \pi_0^{\theta'}(\omega)\, \mathbb{P}_S^{\theta',\omega}\left(\hat{H}\right)$, where $\mathbb{P}_S^{\theta',\omega}$ is the probability measure over H induced by the switcher if the true DGP is as described by θ' and ω. Take the conditional expectation of κ_t with respect to $\mathbb{P}_S^{\theta'}$, then we have

$$\begin{aligned}
&\mathbb{E}^{\mathbb{P}_S^{\theta'}}\left(\kappa_t|h_t\right)\\
&=\mathbb{E}^{\mathbb{P}_S^{\theta'}}\left[\frac{\sum_{\omega\in\Omega^\theta}q^\theta\left(y_{t-1}|a_{t-1},\omega\right)\pi_{t-1}^\theta\left(\omega\right)}{\sum_{\omega'\in\Omega^{\theta'}}q^{\theta'}\left(y_{t-1}|a_{t-1},\omega'\right)\pi_{t-1}^{\theta'}\left(\omega'\right)}\kappa_{t-1}|h_t\right]\\
&=\kappa_{t-1}\sum_{\tilde{\omega}\in\Omega^{\theta'}}\pi_{t-1}^{\theta'}\left(\tilde{\omega}\right)\left[\int_{\mathcal{Y}}\frac{\sum_{\omega\in\Omega^\theta}q^\theta\left(y_{t-1}|a_{t-1},\omega\right)\pi_{t-1}^\theta\left(\omega\right)}{\sum_{\omega'\in\Omega^{\theta'}}q^{\theta'}\left(y_{t-1}|a_{t-1},\omega'\right)\pi_{t-1}^{\theta'}\left(\omega'\right)}q^{\theta'}\left(y_{t-1}|a_{t-1},\tilde{\omega}\right)\nu\left(dy_{t-1}\right)\right]\\
&=\kappa_{t-1}\int_{\mathcal{Y}}\left[\frac{\sum_{\omega\in\Omega^\theta}q^\theta\left(y_{t-1}|a_{t-1},\omega\right)\pi_{t-1}^\theta\left(\omega\right)}{\sum_{\omega'\in\Omega^{\theta'}}q^{\theta'}\left(y_{t-1}|a_{t-1},\omega'\right)\pi_{t-1}^{\theta'}\left(\omega'\right)}\left(\sum_{\tilde{\omega}\in\Omega^{\theta'}}q^{\theta'}\left(y_{t-1}|a_{t-1},\tilde{\omega}\right)\pi_{t-1}^{\theta'}\left(\tilde{\omega}\right)\right)\right]\nu\left(dy_{t-1}\right)\\
&=\kappa_{t-1}\int_{\mathcal{Y}}\left[\sum_{\omega\in\Omega^\theta}q^\theta\left(y_{t-1}|a_{t-1},\omega\right)\pi_{t-1}^\theta\left(\omega\right)\right]\nu\left(dy_{t-1}\right)\\
&=\kappa_{t-1}\sum_{\omega\in\Omega^\theta}\left[\int_{\mathcal{Y}}q^\theta\left(y_{t-1}|a_{t-1},\omega\right)\nu\left(dy_{t-1}\right)\right]\pi_{t-1}^\theta\left(\omega\right)=\kappa_{t-1}.
\end{aligned}$$

Hence, κ_t is a martingale w.r.t. $\mathbb{P}_S^{\theta'}$. Since $\kappa_t\geq 0,\forall t$, the Martingale Convergence Theorem implies that κ_t converges to κ almost surely w.r.t. $\mathbb{P}_S^{\theta'}$, and $\mathbb{E}^{\mathbb{P}_S^{\theta'}}\kappa\leq\mathbb{E}^{\mathbb{P}_S^{\theta'}}\kappa_0=1$. Since θ' is correctly specified, there exists a parameter $\omega^*\in\Omega^{\theta'}$ such that $q^*\left(\cdot|a\right)\equiv q^{\theta'}\left(\cdot|a,\omega^*\right),\forall a\in\mathcal{A}$. It then follows from $\pi_0^{\theta'}\left(\omega^*\right)>0$ that κ_t also converges to κ almost surely w.r.t. $\mathbb{P}_S^{\theta',\omega^*}$, which is the same measure as $\mathbb{P}_S$. Moreover, $\mathbb{E}\kappa<\infty$ because otherwise it contradicts $\mathbb{E}^{\mathbb{P}_S^{\theta'}}\kappa\leq 1$. $\square$

Lemma 4. *Suppose $\theta\in\Theta$ persists against a correctly specified model $\theta'\in\Theta$ at some full-support priors $\pi_0^\theta,\pi_0^{\theta'}$. Then on paths where m_t eventually equals θ, we have $\lambda_t\xrightarrow{a.s.}\lambda_\infty\leq\alpha$, $\pi_t^{\theta'}\xrightarrow{a.s.}\pi_\infty^{\theta'}$, and $\pi_t^\theta\xrightarrow{a.s.}\pi_\infty^\theta$.*

Proof. It immediately follows from Lemma 3 that $\ell_t(\theta')/\ell_t(\theta)\xrightarrow{\text{a.s.}}\iota\leq\alpha$ on paths where m_t converges to θ. I now show that π_t^θ and $\pi_t^{\theta'}$ also converge almost surely. Given any $\omega\in\Omega^\theta$, we can write

$$\begin{aligned}
\frac{\pi_t^\theta\left(\omega\right)}{\pi_0^\theta\left(\omega\right)}&=\frac{\prod_{\tau=0}^{t-1}q^\theta\left(y_\tau|a_\tau,\omega\right)}{\sum_{\omega'\in\Omega^\theta}\prod_{\tau=0}^{t-1}q^\theta\left(y_\tau|a_\tau,\omega'\right)\pi_0^\theta\left(\omega'\right)}\\
&=\frac{\ell_t(\theta')}{\ell_t(\theta)}\cdot\frac{\prod_{\tau=0}^{t-1}q^\theta\left(y_\tau|a_\tau,\omega\right)}{\sum_{\omega''\in\Omega^{\theta'}}\prod_{\tau=0}^{t-1}q^{\theta'}\left(y_\tau|a_\tau,\omega''\right)\pi_0^{\theta'}\left(\omega''\right)}\\
&:=\frac{\ell_t(\theta')}{\ell_t(\theta)}\cdot\frac{\ell_t(\theta,\omega)}{\ell_t(\theta')},
\end{aligned}$$

where the second term $\ell_t(\theta,\omega)/\ell_t(\theta')$ can be seen as the likelihood ratio of a model that consists of a single parameter ω and the competing model θ'. By Lemma 3, $\ell_t(\theta,\omega)/\ell_t(\theta')$ a.s. converges to a random variable with finite expectation. Consider the paths on which m_t converges to θ. On these paths, both $\ell_t(\theta')/\ell_t(\theta)$ and $\ell_t(\theta,\omega)/\ell_t(\theta')$ converges a.s., which implies that $\pi_t^{\theta}(\omega)$ a.s. converges to a random variable with finite expectation as well. Since this is true for all $\omega \in \Omega^{\theta}$, π_t^{θ} a.s. converges to some limit π_{∞}^{θ} on those paths. Analogously, for any $\omega' \in \Omega^{\theta'}$, we can write

$$\frac{\pi_t^{\theta'}(\omega')}{\pi_0^{\theta'}(\omega')} = \frac{\ell_t(\theta',\omega')}{\ell_t(\theta')},$$

which, again by Lemma 3, converges almost surely. □

Lemma 5. *Fix any $\theta,\theta' \in \Theta$, $\omega \in \Omega^{\theta}, \omega' \in \Omega^{\theta'}$ and any sequence of actions $(a_1, a_2, ...)$. For each infinite history $h \in (\mathcal{A} \times \mathcal{Y})^{\infty}$ that is generated according to $(a_1, a_2, ...)$ by the true DGP, let*

$$\xi_t(h) = \ln \frac{q^{\theta}(y_t|a_t,\omega)}{q^{\theta'}(y_t|a_t,\omega')} - \mathbb{E}\left(\ln \frac{q^{\theta}(y_t|a_t,\omega)}{q^{\theta'}(y_t|a_t,\omega')}|h_t\right).$$

Then for any fixed $t_0 \geq 1$,

$$\lim_{t\to\infty} (t - t_0 + 1)^{-1} \sum_{\tau=t_0}^{t} \xi_{\tau}(h) = 0, \;\; a.s..$$

Proof. Note that $\xi_t(h)$ is a martingale difference process since $E(\xi_t(h)|h_t) = 0$. So for any t_0, $\xi_{t_0}^t(h) := \sum_{\tau=t_0}^{t} (t-\tau+1)^{-1} \xi_{\tau}(h)$ is also a martingale difference process. To use the Martingale Convergence Theorem, I now show that $\sup_t \mathbb{E}\left(\left(\xi_{t_0}^t\right)^2\right) < \infty$.

Notice that

$$\begin{aligned}
\mathbb{E}\left(\left(\xi_{t_0}^{t}\right)^{2}\right) =&\mathbb{E}\left[\left(\sum_{\tau=t_0}^{t}(t-\tau+1)^{-1}\xi_{\tau}(h)\right)^{2}\right] \\
\leq&\sum_{\tau=t_0}^{t}(t-\tau+1)^{-2}\,\mathbb{E}\left[(\xi_{\tau}(h))^{2}\right] \\
\leq&\sum_{\tau=t_0}^{t}(t-\tau+1)^{-2}\,\mathbb{E}\left[\left(\ln\frac{q^{\theta}(y_t|a_t,\omega)}{q^{\theta'}(y_t|a_t,\omega')}\right)^{2}\right] \\
\leq&\sum_{\tau=t_0}^{t}(t-\tau+1)^{-2}\,\mathbb{E}\left[\left(\ln\frac{q^{*}(y_t|a_t)}{q^{\theta}(y_t|a_t,\omega)}\right)^{2}+\left(\ln\frac{q^{*}(y_t|a_t)}{q^{\theta'}(y_t|a_t,\omega')}\right)^{2}\right] \\
\leq&2\sum_{\tau=t_0}^{t}(t-\tau+1)^{-2}\max_{a}\mathbb{E}\left[(r_a(y))^{2}\right]<\infty,
\end{aligned}$$

where the first inequality follows from the fact that, for any $\tau' > \tau \geq t_0$, $\mathbb{E}(\xi_{\tau}(h)\,\xi_{\tau'}(h)) = \mathbb{E}(\mathbb{E}(\xi_{\tau'}(h)\,|h_{\tau})\,\xi_{\tau}(h)) = 0$ and the last inequality follows from Assumption 2. Now we can invoke the Martingale Convergence Theorem which implies that $\xi_{t_0}^{t}$ converges to a random variable $\xi_{t_0}^{\infty}$ almost surely with $\mathbb{E}\left(\left(\xi_{t_0}^{\infty}\right)^{2}\right) < \infty$. Since $\xi_{t_0}^{\infty} = \lim_{t\to\infty}\sum_{\tau=t_0}^{t}(t-\tau+1)^{-1}\xi_{\tau}(h)$ is finite a.s., it follows from the Kronecker Lemma that

$$\lim_{t\to\infty}(t-t_0+1)^{-1}\sum_{\tau=t_0}^{t}\xi_{\tau}(h)=0,\text{ a.s..}$$

□

Let us define action frequency $\sigma_t : \mathcal{A}^{t+1} \to \Delta\mathcal{A}$ to measure how frequent each action has been played up to period t. In particular, given an action sequence $(a_0, a_1, \ldots)$, let

$$\sigma_t(a)=\frac{\sum_{\tau=0}^{t}\mathbf{1}(a_t=a)}{t}.$$

Lemma 6. *Fix any $\theta \in \Theta$. Suppose the action frequency of a θ-modeler converges to σ, then her belief π_t^{θ} satisfies $\pi_t^{\theta}\left(\Omega^{\theta}(\sigma)\right) \xrightarrow{a.s.} 1$. Similarly, if the action frequency of a switcher with $\Theta^{\dagger} \ni \theta$ converges to σ, then her belief π_t^{θ} also satisfies $\pi_t^{\theta}\left(\Omega^{\theta}(\sigma)\right) \xrightarrow{a.s.} 1$.*

Proof. The proof is completely identical for either a θ-modeler or a switcher. Since Ω^{θ}

is finite, for any given σ, there exists $\epsilon > 0$ such that

$$\sum_{\mathcal{A}} \sigma(a) \left[D_{KL}\left(q^*(\cdot|a) \parallel q^\theta(\cdot|a,\omega)\right) - D_{KL}\left(q^*(\cdot|a) \parallel q^\theta(\cdot|a,\omega')\right)\right] < -\epsilon \tag{13}$$

for all $\omega \in \Omega^\theta(\sigma)$ and $\omega' \in \Omega^\theta / \Omega^\theta(\sigma)$. For any $\omega \in \Omega^\theta(\sigma)$ and $\omega' \in \Omega^\theta / \Omega^\theta(\sigma)$ at time t,

$$\begin{aligned} \frac{\pi_t^\theta(\omega')}{\pi_t^\theta(\omega)} &= \frac{\prod_{\tau=0}^{t-1} q^\theta(y_\tau|a_\tau,\omega')\,\pi_0^\theta(\omega')}{\prod_{\tau=0}^{t-1} q^\theta(y_\tau|a_\tau,\omega)\,\pi_0^\theta(\omega)} \\ &= \exp\left(\sum_{\tau=0}^{t-1} \ln \frac{q^\theta(y_\tau|a_\tau,\omega')}{q^\theta(y_\tau|a_\tau,\omega)} + \ln \frac{\pi_0^\theta(\omega')}{\pi_0^\theta(\omega)}\right). \end{aligned}$$

We are done if this ratio converges to 0. Notice that

$$\begin{aligned} &\frac{1}{t}\sum_{\tau=0}^{t-1} \mathbb{E}\left(\ln \frac{q^\theta(y_\tau|a_\tau,\omega')}{q^\theta(y_\tau|a_\tau,\omega)}|h_t\right) \\ = &-\sum_{\mathcal{A}} \sigma_t(a)\left[D_{KL}\left(q^*(\cdot|a) \parallel q^\theta(\cdot|a,\omega')\right) - D_{KL}\left(q^*(\cdot|a) \parallel q^\theta(\cdot|a,\omega)\right)\right], \end{aligned}$$

which converges to the left-hand side of Eq. (13) as σ_t converges to σ. Hence, there exists T_1 such that

$$\frac{1}{t}\sum_{\tau=0}^{t-1} \mathbb{E}\left(\ln \frac{q^\theta(y_\tau|a_\tau,\omega')}{q^\theta(y_\tau|a_\tau,\omega)}|h_t\right) < -\frac{\epsilon}{2}, \forall t > T_1.$$

By Lemma 5, there exists T_2 such that when $t > T_2$,

$$\frac{1}{t}\sum_{\tau=0}^{t-1} \ln \frac{q^\theta(y_\tau|a_\tau,\omega')}{q^\theta(y_\tau|a_\tau,\omega)} < \frac{1}{t}\sum_{\tau=0}^{t-1} \mathbb{E}\left(\ln \frac{q^\theta(y_\tau|a_\tau,\omega')}{q^\theta(y_\tau|a_\tau,\omega)}|h_t\right) + \frac{\epsilon}{3}$$

It follows that when $t > \max\{T_1, T_2\}$,

$$\sum_{\tau=0}^{t-1} \ln \frac{q^\theta(y_\tau|a_\tau,\omega')}{q^\theta(y_\tau|a_\tau,\omega)} < t\cdot\left(-\frac{\epsilon}{6}\right).$$

Hence, $\frac{\pi_t^\theta(\omega')}{\pi_t^\theta(\omega)}$ converges to 0 for all $\omega \in \Omega^\theta(\sigma)$ and $\omega' \in \Omega^\theta / \Omega^\theta(\sigma)$. $\square$

Lemma 7. *For any $\theta \in \Theta$, the optimal action correspondence $A^\theta : \Delta\Omega^\theta \rightrightarrows \mathcal{A}$ is upper hemicontinuous in both the belief π and the agent's discount factor δ.*

Proof. This is a standard result directly following from Blackwell (1965) and Maitra (1968). $\square$

Lemma 8. *For any $\theta \in \Theta$, the set of all Berk-Nash equilibria under θ is compact.*

Proof. Denote the set of all Berk-Nash equilibria under model θ as $BN^{\theta} \subseteq \Delta\mathcal{A}$. Since $\Delta\mathcal{A}$ is bounded, we only need to show that BN^{θ} is closed. Suppose σ is the limit of some sequence $(\sigma_n)_n$ of Berk-Nash equilibria, but σ is not a Berk-Nash equilibrium, i.e. $\sigma \notin BN^{\theta}$. Then for every belief $\pi \in \Delta\Omega^{\theta}(\sigma)$, we have that $\sigma \notin \Delta A_m^{\theta}(\pi)$. Since $\Omega^{\theta}(\cdot)$ is upper hemicontinuous, it must be that $\Omega^{\theta}(\sigma_n) \subseteq \Omega^{\theta}(\sigma)$ for large enough n. Hence, we have $\sigma \notin \Delta A_m^{\theta}(\pi)$ for every belief $\pi \in \Delta\Omega^{\theta}(\sigma_n)$ when n is large enough. However, we know that $\text{supp}(\sigma) \subseteq \text{supp}(\sigma_n)$ for large enough n, which implies that $\sigma_n \notin \Delta A_m^{\theta}(\pi)$ for large n. This is a contradiction. $\square$

B Proofs of Main Results

I prove all theorems under the assumption that the agent may be non-myopic within each model but is myopic across models (see Section 6.2). This includes the special case where the agent is myopic everywhere.

B.1 Proof of Theorem 1: (i)$\Leftrightarrow$(iii)

I first prove Lemma 2, which implies the necessity of a p-absorbing SCE for global robustness. I then prove Lemma 9, which is then used to show sufficiency.

Proof of Lemma 2. By Lemma 4, on paths where θ is eventually forever adopted, beliefs π_t^{θ} and $\pi_t^{\theta'}$ both converge almost surely. Consider any $\hat{\omega}$ such that with positive probability, m_t eventually equals θ and $\hat{\omega} \in \text{supp}\big(\pi_{\infty}^{\theta}\big)$. Let $A^{-}(\hat{\omega}) \equiv \big\{a \in \mathcal{A} : q^{\theta}(\cdot|a,\hat{\omega}) \neq q^{*}(\cdot|a)\big\}$. I now show that every action in $A^{-}(\hat{\omega})$ is played at most finite times a.s. on the paths where m_t converges to θ and $\hat{\omega} \in \text{supp}\big(\pi_{\infty}^{\theta}\big)$. Suppose instead that actions in $A^{-}(\hat{\omega})$ are played infinitely often. Then there must exist some $\gamma > 0$ such that $\mathbb{E}\ln\frac{q^{*}(y|a_t)}{q^{\theta}(y|a_t,\hat{\omega})} > \gamma$ for infinitely many t. Since θ' is correctly specified, there exists a parameter $\omega^{*} \in \Omega^{\theta'}$ such that $q^{*}(\cdot|a) \equiv q^{\theta'}(\cdot|a,\omega^{*}), \forall a \in \mathcal{A}$. Hence, $\mathbb{E}\ln\frac{q^{\theta'}(y|a_t,\omega^{*})}{q^{\theta}(y|a_t,\hat{\omega})} > \gamma$ for infinitely many

t. Notice that

$$\begin{aligned}
\frac{\ell_t(\theta')}{\ell_t(\theta)} &= \frac{\sum_{\omega'\in\Omega^{\theta'}}\prod_{\tau=0}^{t-1} q^{\theta'}\left(y_\tau|a_\tau,\omega'\right)\pi_0^{\theta'}\left(\omega'\right)}{\sum_{\omega\in\Omega^{\theta}}\prod_{\tau=0}^{t-1} q^{\theta}\left(y_\tau|a_\tau,\omega\right)\pi_0^{\theta}\left(\omega\right)} \\
&> \pi_t^\theta\left(\hat\omega\right)\frac{\pi_0^{\theta'}\left(\omega^*\right)}{\pi_0^\theta\left(\hat\omega\right)}\frac{\prod_{\tau=0}^{t-1} q^{\theta'}\left(y_\tau|a_t,\omega^*\right)}{\prod_{\tau=0}^{t-1} q^{\theta}\left(y_\tau|a_\tau,\hat\omega\right)} \\
&= \pi_t^\theta\left(\hat\omega\right)\frac{\pi_0^{\theta'}\left(\omega^*\right)}{\pi_0^\theta\left(\hat\omega\right)}\exp\left[\sum_{\tau=0}^{t-1} 1_{\{a_\tau\in A^-(\hat\omega)\}}\ln\frac{q^{\theta'}\left(y_\tau|a_t,\omega^*\right)}{q^{\theta}\left(y_\tau|a_\tau,\hat\omega\right)}\right],
\end{aligned}$$

which, by Lemma 5, a.s. increases to infinity as $t\to\infty$, contradicting the assumption that m_t converges to θ. Therefore, on the paths where m_t eventually equals θ, almost surely, there exists T such that $a_t \in \mathcal{A}\backslash\cup_{\omega'\in\text{supp}(\pi_\infty^\theta)} A^-\left(\hat\omega\right), \forall t > T$.

Since $q^\theta\left(\cdot|a,\omega'\right)\equiv q^*\left(\cdot|a\right)$ for all $\omega'\in\text{supp}\left(\pi_\infty^\theta\right)$ and all $a\in\mathcal{A}\backslash\cup_{\omega'\in\text{supp}(\pi_\infty^\theta)} A^-\left(\omega'\right)$, the actions that are played in the limit have no experimentation value and are myopically optimal. Therefore, any strategy that takes support on the limit actions is a self-confirming equilibrium. Fixing a particular value of π_∞^θ that is a limit belief for a positive measure of histories where m_t eventually equals θ, there exists a set of actions $\hat A\subseteq A_m^\theta\left(\pi_\infty^\theta\right)$ such that on those histories, the agent only plays actions from this set in the limit. Since m_t eventually converges to θ, it must be true that with positive probability, a θ-modeler who inherits the switcher's prior and policy from the period when the last switch happens also only plays actions from $\hat A$ in the limit with positive probability. Therefore, take any strategy σ with $\text{supp}\left(\sigma\right)=\hat A$, it is a p-absorbing self-confirming equilibrium under θ. □

Lemma 9. *If σ is a p-absorbing SCE, then for any $\gamma\in(0,1)$ and $\epsilon>0$, there exists a full-support prior π_0^θ under which, with probability higher than γ, a θ-modeler only plays actions in* $\text{supp}(\sigma)$ *and her belief stays within $B_\epsilon(\Delta\Omega^\theta(\sigma))$ for all periods.*[38]

Proof of Lemma 9. Suppose there exists a p-absorbing SCE σ under θ. Consider the learning process of a θ-modeler. By definition, there exists a full-support prior $\pi_0^\theta\in\Delta\Omega^\theta$ such that with positive probability, she eventually only plays actions in $\text{supp}\left(\sigma\right)$ and each element of $\text{supp}\left(\sigma\right)$ is played infinitely often (this is without loss of generality). Denote those paths by $\tilde H$. Then by a similar argument as in the proof of Lemma 2, π_t^θ a.s. converges to a limit π_∞^θ on $\tilde H$, with $\text{supp}\left(\pi_\infty^\theta\right)\subseteq\Omega^\theta(\sigma)=\{\omega\in\Omega^\theta : q^*\left(\cdot|a\right)=q^\theta\left(\cdot|a,\omega\right),\forall a\in\text{supp}\left(\sigma\right)\}$.

[38] For any set of finite probability distributions Z over sample space S, I use $B_\epsilon(Z)$ to denote the set of probability distributions whose minimum distance from any element in Z is smaller than ϵ, i.e. $B_\epsilon(Z)=\{z\in\Delta S:\min_{z'\in Z} d_P(z,z')<\epsilon\}$, where d_P represents the usual Prokhorov metric over ΔS.

This implies the existence of an integer $T > 0$ such that, with positive probability, we have (1) $a_t \in \text{supp}(\sigma), \forall t \geq T$, (2) π_t^θ converges to a limit π_∞^θ with $\text{supp}(\pi_\infty^\theta) \subseteq \Omega^\theta(\sigma)$. Pick any $\epsilon > 0$. Since the learning processes are Markov, we can find a new prior $\tilde{\pi}_0^\theta \in B_\epsilon(\Delta\Omega^\theta(\sigma))$ under which, on a positive measure of histories, a θ-modeler behaves such that (1′) $a_t \in \text{supp}(\sigma), \forall t \geq 0$, and (2′) the posterior $\tilde{\pi}_t^\theta$ almost surely converges to π_∞^θ and never leaves $B_\epsilon(\Delta\Omega^\theta(\sigma))$ for all $t \geq 0$.

Denote the event described by (1′) and (2′) by E . I now show for any constant $\gamma \in (0, 1)$, there exists a full-support prior $\hat{\pi}_0^\theta$ under which $\mathbb{P}_D(E) > \gamma$. Suppose for contradiction that this is not true. Denote the probability of E under any full-support prior by $\gamma(\pi^\theta)$ and let $\overline{\gamma} := \sup_{\pi_0^\theta \in \text{int}(\Delta\Omega^\theta)} \gamma(\pi^\theta)$, where $\text{int}(\Delta\Omega^\theta)$ denotes all full-support beliefs over Ω^θ, then it follows that $\overline{\gamma} < 1$. By definition, for any $\psi > 0$, there exists some prior $\pi_0^{\theta,\psi}$ such that $\gamma(\pi_0^{\theta,\psi}) > \overline{\gamma} - \psi$. But under this prior, with probability $1 - \gamma(\pi_0^{\theta,\psi})$, the dogmatic modeler eventually either arrives at some posterior $\pi_t^{\theta,\psi}$ that either leads her to play an action outside $\text{supp}(\sigma)$ or leaves the neighborhood $B_\epsilon(\Delta\Omega^\theta(\sigma))$. Hence, there exists an integer $T > 0$ such that

$$\mathbb{P}_D\left(\gamma(\pi_T^{\theta,\psi}) = 0\right) > \gamma(\pi_0^{\theta,\psi}) - \psi > \overline{\gamma} - 2\psi.$$

Now, consider the supremum probability that E is achieved if the agent starts with a prior that is equal to one of the possible posteriors $\pi_T^{\theta,\psi}$. Since

$$\gamma(\pi_0^{\theta,\psi}) = \mathbb{E}_{h_T \in H_T}^{\mathbb{P}_D} \gamma(\pi_T^{\theta,\psi}),$$

we have

$$\begin{aligned}\sup_{h_T \in H_T} \gamma(\pi_T^{\theta,\psi}) &> \frac{\gamma(\pi_0^{\theta,\psi})}{1 - \mathbb{P}_D\left(\gamma(\pi_T^{\theta,\psi}) = 0\right)} \\ &> \frac{\overline{\gamma} - \psi}{1 - \overline{\gamma} + 2\psi}.\end{aligned}$$

But notice that when ψ is sufficiently small, the term $\frac{\overline{\gamma} - \psi}{1 - \overline{\gamma} + 2\psi}$ is strictly larger than $\overline{\gamma}$, contradicting the assumption that $\overline{\gamma}$ is the supremum of $\gamma(\pi^\theta)$ over all full-support beliefs.

□

Proof of Theorem 1 (iii)⇒(i). Pick any competing model $\theta' \in \Theta$ and any full-support prior $\pi_0^{\theta'} \in \Delta\Omega^{\theta'}$. Let $S_t := \ell_t(\theta')/\ell_t(\theta^*)$, then S_t is a martingale with respect to both

$\mathbb{P}_D$ and $\mathbb{P}_S$ by Lemma 3. By the Ville's maximal inequality for supermartingales, the probability that S_n is bounded above by a positive constant larger than 1 is bounded away from 0. In particular, for any $\eta \in (1, \alpha)$,

$$\mathbb{P}_D(S_t \leq \eta, \forall t \geq 0) \geq 1 - \frac{\mathbb{E}^{\mathbb{P}_D} S_0}{\eta} = 1 - \frac{1}{\eta}.$$

Note that this inequality holds for any model θ'.

Denote by σ a p-absorbing SCE under θ. By Lemma 9, we know that for any $\eta \in (1, \alpha)$ and $\epsilon > 0$, there exist a prior $\pi_0^\theta \in B_\epsilon(\Delta\Omega^\theta(\sigma))$ such that $\mathbb{P}_D(E) > 1/\eta$ (the event E is defined in the proof of Lemma 9). Therefore,

$$\begin{aligned} & \mathbb{P}_D(E \text{ occurs and } S_t \leq \eta, \forall t \geq 0) \\ \geq \quad & \mathbb{P}_D(E) + \mathbb{P}_D(S_t \leq \eta, \forall t \geq 0) - 1 > 0. \end{aligned}$$

Denote the histories where E occurs and $S_t \leq \eta, \forall t \geq 0$ by $\hat{H}$. When ϵ is small enough, we have that on $\hat{H}$,

$$\begin{aligned} \lambda_t = \frac{\ell_t(\theta')}{\ell_t(\theta)} &= \frac{\sum_{\omega' \in \Omega^{\theta'}} \pi_0^{\theta'}(\omega') \prod_{\tau=0}^{t-1} q^{\theta'}(y_\tau | a_\tau, \omega')}{\sum_{\omega \in \Omega^\theta} \pi_0^\theta(\omega) \prod_{\tau=0}^{t-1} q^\theta(y_\tau | a_\tau, \omega)} \\ &< \frac{\sum_{\omega' \in \Omega^{\theta'}} \pi_0^{\theta'}(\omega') \prod_{\tau=0}^{t-1} q^{\theta'}(y_\tau | a_\tau, \omega')}{\pi_0^\theta(\Omega^\theta(\sigma)) \prod_{\tau=0}^{t-1} q^*(y_\tau | a_\tau)} \\ &\leq \frac{\eta}{1-\epsilon} < \alpha \end{aligned}$$

where the first inequality follows from the fact that π_0^θ is full-support and the second inequality follows from the definition of $\hat{H}$. Thus, on $\hat{H}$, the switcher never makes any switch to the competing model θ', i.e. $m_t = \theta, \forall t \geq 0$, and her action choices would be identical to the θ-modeler. Therefore, if we endow the switcher with the same prior π_0^θ, event $\hat{H}$ also occurs with positive probability under $\mathbb{P}_S$. $\square$

B.2 Proof of Theorem 1: (ii)⇒(iii)

I show that if θ is locally robust at some prior, then it must admit a p-absorbing SCE. Construct a competing model θ' as follows. Let θ' have the identical parameter space as θ, i.e. $\Omega^{\theta'} = \Omega^\theta$, and let its predictions be given by $q^{\theta'}(\cdot|a, \omega) = \mu q^\theta(\cdot|a, \omega) + (1 - \mu) q^*(\cdot|a)$, for all $a \in \mathcal{A}$ and all $\omega \in \Omega^\theta$, where $\mu \in (0, 1)$. For any $\epsilon > 0$, when μ is close enough to 1, we have $\theta' \in N_\epsilon(\theta)$. By the definition of local robustness, there exists

$\epsilon > 0$ such that θ persists against θ' under some full-support priors π_0^θ and $\pi_0^{\theta'} = \pi_0^\theta$. Consider any $\hat{\omega} \in \Omega^\theta$ such that

$$\mathbb{P}_S \left(m_t \text{ eventually equals } \theta \text{ and } \liminf_{t\to\infty} \pi_t^\theta(\hat{\omega}) > 0 \right) > 0.$$

Let $A^-(\hat{\omega}) := \{a \in \mathcal{A} : q^\theta(\cdot|a, \hat{\omega}) \neq q^*(\cdot|a)\}$. Then every action in $A^-(\hat{\omega})$ is played at most finite times a.s. on the path where m_t eventually equals θ and $\liminf_{t\to\infty} \pi_t^\theta(\hat{\omega}) > 0$. Suppose instead that actions in $A^-(\hat{\omega})$ are played infinitely often. Then there must exist some $\gamma > 0$ such that $\mathbb{E} \ln \frac{q^*(y|a_t)}{q^\theta(y|a_t,\hat{\omega})} > \gamma$ for infinitely many t. So we have

$$\mathbb{E} \ln \frac{q^{\theta'}(y|a_t, \hat{\omega})}{q^\theta(y|a_t, \hat{\omega})} = \mathbb{E} \ln \left(\mu + (1-\mu) \frac{q^*(y|a_t)}{q^\theta(y|a_t, \hat{\omega})} \right) > (1-\mu)\gamma$$

where the inequality follows from the concavity of the logarithm function. Therefore,

$$\begin{aligned} \lambda_t =& \frac{\sum_{\omega\in\Omega^\theta} \prod_{\tau=0}^{t-1} q^{\theta'}(y_\tau|a_\tau, \omega)\, \pi_0^\theta(\omega)}{\sum_{\omega\in\Omega^\theta} \prod_{\tau=0}^{t-1} q^{\theta}(y_\tau|a_\tau, \omega)\, \pi_0^\theta(\omega)} \\ >& \pi_t^\theta(\hat{\omega}) \frac{\pi_0^\theta(\hat{\omega})}{\pi_0^\theta(\hat{\omega})} \frac{\prod_{\tau=0}^{t-1} q^{\theta'}(y_\tau|a_t, \hat{\omega})}{\prod_{\tau=0}^{t-1} q^{\theta}(y_\tau|a_\tau, \hat{\omega})} \\ =& \pi_t^\theta(\hat{\omega}) \exp\left[\sum_{\tau=0}^{t-1} 1_{\{a_\tau \in A^-(\hat{\omega})\}} \ln \frac{q^{\theta'}(y_\tau|a_t, \hat{\omega})}{q^\theta(y_\tau|a_\tau, \hat{\omega})} \right], \end{aligned}$$

which, by Lemma 5, a.s. increases to infinity when m_t converges to θ and $\liminf_{t\to\infty} \pi_t^\theta(\hat{\omega}) > 0$. This implies that, letting $\hat{\Omega}^\theta := \{\omega \in \Omega^\theta : \liminf_{t\to\infty} \pi_t^\theta(\hat{\omega}) > 0\}$, on the paths where m_t eventually equals θ, there almost surely exists T such that $a_t \in \mathcal{A} \backslash \cup_{\hat{\omega}\in\hat{\Omega}^\theta} A^-(\hat{\omega}), \forall t > T$. Since $q^\theta(\cdot|a, \hat{\omega})$ is equal to $q^*(\cdot|a)$ for all $\hat{\omega} \in \hat{\Omega}^\theta$ and all $a \in \mathcal{A} \backslash \cup_{\hat{\omega}\in\hat{\Omega}^\theta} A^-(\hat{\omega})$, the posterior π_t^θ must converge to a limit π_∞^θ. The rest of the arguments are identical to those in the proof of Lemma 2; it follows that θ must admit a p-absorbing SCE. □

B.3 Proof of Corollary 1

I prove Corollary 1 assuming the agent is myopic. I show below that any quasi-strict SCE satisfies a stability property stronger than p-absorbingness, which implies Corollary 1. Next, I show that if we strengthen quasi-strictness with uniform quasi-strictness, then Corollary 1 holds for a non-myopic agent as well.

Lemma 10. *Suppose σ is a quasi-strict SCE with supporting belief $\hat{\pi}$, then for any $\gamma \in (0,1)$, there exists $\epsilon > 0$ such that starting from any prior $\pi_0^\theta \in B_\epsilon(\hat{\pi})$, the probability that the θ-modeler always plays actions in* $\operatorname{supp}(\sigma)$ *for all periods is strictly larger than γ.*

Proof. If σ is quasi-strict, then $\operatorname{supp}(\sigma) = A_m^\theta(\hat{\pi})$. Since A^θ is upper hemicontinuous (Lemma 7), there exists $\tilde{\epsilon} > 0$ small enough such that $\operatorname{supp}(\sigma) \supset A^\theta(\pi)$ for all $\pi \in B_{\tilde{\epsilon}}(\hat{\pi})$.

Suppose $a_t \in \operatorname{supp}(\sigma), \forall t \geq 0$, then for every $\omega \in \Omega^\theta \backslash \Omega^\theta(\sigma)$,

$$\begin{aligned}
\mathbb{E}\left[\frac{\pi_t^\theta(\omega)}{\pi_t^\theta(\Omega^\theta(\sigma))}|h_t\right] &= \mathbb{E}\left[\frac{\pi_0^\theta(\omega)\prod_{\tau=0}^{t-1} q^\theta(y_\tau|a_\tau,\omega)}{\sum_{\omega'\in\Omega^\theta(\sigma)}\pi_0^\theta(\omega')\prod_{\tau=0}^{t-1} q^\theta(y_\tau|a_\tau,\omega')}|h_t\right] \\
&= \mathbb{E}\left[\frac{\pi_0^\theta(\omega)}{\pi_0^\theta(\Omega^\theta(\sigma))}\frac{\prod_{\tau=0}^{t-1} q^\theta(y_\tau|a_\tau,\omega)}{\prod_{\tau=0}^{t-1} q^*(y_\tau|a_\tau)}|h_t\right] \\
&= \frac{\pi_0^\theta(\omega)\prod_{\tau=0}^{t-2} q^\theta(y_\tau|a_\tau,\omega)}{\pi_0^\theta(\Omega^\theta(\sigma))\prod_{\tau=0}^{t-2} q^*(y_\tau|a_\tau)} = \frac{\pi_{t-1}^\theta(\omega)}{\pi_{t-1}^\theta(\Omega^\theta(\sigma))}
\end{aligned}$$

Therefore, $\frac{\pi_t^\theta(\omega)}{\pi_t^\theta(\Omega^\theta(\sigma))}$ is a non-negative supermartingale for every $\omega \in \Omega^\theta \backslash \Omega^\theta(\sigma)$. It follows that $\frac{\pi_t^\theta(\Omega^\theta \backslash \Omega^\theta(\sigma))}{\pi_t^\theta(\Omega^\theta(\sigma))}$ is also non-negative supermartingale. By the Ville's maximal inequality for supermartingales, for any $\eta > 0$,

$$\mathbb{P}_D\left(\frac{\pi_t^\theta(\Omega^\theta \backslash \Omega^\theta(\sigma))}{\pi_t^\theta(\Omega^\theta(\sigma))} \geq \eta \text{ for some } t\right) < \frac{1}{\eta}\frac{\pi_0^\theta(\Omega^\theta \backslash \Omega^\theta(\sigma))}{\pi_0^\theta(\Omega^\theta(\sigma))}.$$

Since $\pi_t^\theta(\Omega^\theta(\sigma)) = 1 - \pi_t^\theta(\Omega^\theta \backslash \Omega^\theta(\sigma))$, the above inequality implies that

$$\mathbb{P}_D\left(\pi_t^\theta(\Omega^\theta \backslash \Omega^\theta(\sigma)) \geq \frac{\eta}{1+\eta} \text{ for some } t\right) < \frac{1}{\eta}\frac{\pi_0^\theta(\Omega^\theta \backslash \Omega^\theta(\sigma))}{\pi_0^\theta(\Omega^\theta(\sigma))}.$$

Pick some $\epsilon \in (0, \tilde{\epsilon})$ and $\pi_0^\theta \in B_\epsilon(\hat{\pi})$, then $\pi_0^\theta(\Omega^\theta \backslash \Omega^\theta(\sigma)) < \epsilon$. Notice that the ratio $\frac{\pi_t^\theta(\omega)}{\pi_t^\theta(\omega')}$ remain unchanged throughout all periods for any $\omega, \omega' \in \Omega^\theta(\sigma)$. Hence, if $\pi_t^\theta \notin B_{\tilde{\epsilon}}(\hat{\pi})$ for some $t \geq 0$, then there exists t such that $\pi_t^\theta(\Omega^\theta \backslash \Omega^\theta(\sigma)) \geq \pi_0^\theta(\Omega^\theta \backslash \Omega^\theta(\sigma)) +$

$\tilde{\epsilon} - \epsilon$. Using the previous inequality,

$$\begin{aligned}
& \mathbb{P}_D\left(\pi_t^\theta \notin B_{\tilde{\epsilon}}(\hat{\pi}) \text{ for some } t \geq 0\right) \\
\leq\ & \mathbb{P}_D\left(\pi_t^\theta\left(\Omega^\theta \backslash \Omega^\theta(\sigma)\right) \geq \pi_0^\theta\left(\Omega^\theta \backslash \Omega^\theta(\sigma)\right) + \tilde{\epsilon} - \epsilon \text{ for some } t\right) \\
<\ & \left(\frac{1}{\pi_0^\theta\left(\Omega^\theta \backslash \Omega^\theta(\sigma)\right) + \tilde{\epsilon} - \epsilon} - 1\right) \frac{\pi_0^\theta\left(\Omega^\theta \backslash \Omega^\theta(\sigma)\right)}{\pi_0^\theta\left(\Omega^\theta(\sigma)\right)} \\
<\ & \left(\frac{1}{\tilde{\epsilon} - \epsilon} - 1\right) \frac{\epsilon}{1 - \epsilon}
\end{aligned}$$

which converges to 0 as ϵ approaches 0. This implies that for any $\gamma \in (0,1)$ we have $\mathbb{P}_D\left(\pi_t^\theta \in B_{\tilde{\epsilon}}(\hat{\pi}), \forall t \geq 0\right) > \gamma$ when ϵ is sufficiently small. Notice that $\pi_t^\theta \in B_{\tilde{\epsilon}}(\hat{\pi}), \forall t \geq 0$ in turn implies that $a_t \in \operatorname{supp}(\sigma), \forall t \geq 0$, validating our assumption.

□

Say that a SCE or a BN-E σ with supporting belief π is *uniformly quasi-strict* if $\operatorname{supp}(\sigma) = A_m^\theta(\pi)$ for every belief $\pi \in \Delta\Omega^\theta(\sigma)$. The following lemma implies that given any discount factor, a uniformly quasi-strict SCE is p-absorbing.

Lemma 11. *Suppose the θ-modeler has discount factor $\delta \in (0,1)$. Suppose σ is a uniformly quasi-strict SCE with supporting belief $\hat{\pi}$, then for any $\gamma \in (0,1)$, there exists $\epsilon > 0$ such that starting from any prior $\pi_0^\theta \in B_\epsilon(\hat{\pi})$, the probability that the θ-modeler always plays actions in* $\operatorname{supp}(\sigma)$ *for all periods is strictly larger than γ.*

Proof. Since σ is uniformly quasi-strict with supporting belief π, $\operatorname{supp}(\sigma)$ contains all myopically optimal actions against each degenerate belief δ_ω concentrated on $\omega \in \operatorname{supp}(\pi)$. In addition, $\operatorname{supp}(\sigma)$ must be optimal against δ_ω for an agent who maximizes discounted utility, because the dynamic programming problem described by (6.2) reduces to a static maximization problem when the belief is degenerate. This implies that $\operatorname{supp}(\sigma)$ is also (dynamically) optimal against π. Further, since A^θ is upper hemicontinuous (by Lemma 7), there exists $\tilde{\epsilon} > 0$ small enough such that $\operatorname{supp}(\sigma) = A^\theta(\tilde{\pi})$ for all $\tilde{\pi} \in B_{\tilde{\epsilon}}(\pi)$. The rest of the proof is identical to the proof of Lemma 10. □

B.4 Proof of Theorem 2(i)

Necessity. Suppose θ is locally robust at some full-support prior π_0^θ. It follows from Theorem 1 and identifiability that there exists $\hat{\omega} \in \Omega^\theta$ such that the degenerate belief δ_ω supports a p-absorbing SCE under θ, i.e. $C^\theta \neq \emptyset$.

Sufficiency. Suppose $C^\theta \neq \emptyset$. Take any $\hat{\omega} \in C^\theta$ and any full-support prior π_0^θ. Consider the probability measure $\mathbb{P}_S^{\theta,\hat{\omega}}$, i.e. the probability measure over infinite histories H induced by the switcher if the true DGP is as described by θ and $\hat{\omega}$. By identifiability and Lemma 5, the posterior π_t^θ converges to $\delta_{\hat{\omega}}$ almost surely under $\mathbb{P}_S^{\theta,\hat{\omega}}$. So for any $\mu > 0$, we can find a positive measure of length-T histories $\hat{H}_T$ where the posterior for model θ enters the μ-neighborhood of $\delta_{\hat{\omega}}$, i.e. $\pi_T^\theta \in B_\mu(\delta_{\hat{\omega}})$. Let μ be small enough so that the posterior $\pi_T^\theta(\hat{\omega}) > 1/\sqrt{\alpha}$. By absolute continuity, we know $\hat{H}_T$ is also realized with positive probability under the true measure $\mathbb{P}_S$.

Next I show that when $\mathcal{Y}$ is discrete, we can choose ϵ to be sufficiently small such that for any $\theta' \in N_\epsilon(\theta)$ and prior $\pi_0^{\theta'} \in N_\epsilon(\pi_0^\theta; \theta, \theta')$, the Bayes factor λ_t never exceeds $\sqrt{\alpha}$ before period T.[39] For each $\omega \in \Omega^\theta$, with a slight abuse of notation, denote the set of "nearby" parameters within θ' by $N_\epsilon(\omega; \theta') := \{\omega' \in \Omega^{\theta'} : d(Q^{\theta,\omega}, Q^{\theta',\omega'}) \leq \epsilon\}$. Then we have $q^{\theta'}(y|a,\omega') \leq q^\theta(y|a,\omega) + \epsilon$ for all $y \in \mathcal{Y}$, $a \in \mathcal{A}$, and $\omega' \in N_\epsilon(\omega;\theta')$. Let ϵ be sufficiently small such that $N_\epsilon(\omega;\theta')$ is disjoint across Ω^θ. By construction we have $\pi_0^{\theta'}(N_\epsilon(\omega;\theta')) \leq \pi_0^\theta(\omega) + \epsilon$. Hence,

$$
\begin{aligned}
\lambda_t = \frac{\ell_t(\theta')}{\ell_t(\theta)} &= \frac{\sum_{\omega \in \Omega^\theta} \sum_{\omega' \in N_\epsilon(\omega;\theta')} \pi_0^\theta(\omega') \prod_{\tau=0}^{t-1} q^{\theta'}(y_\tau | a_\tau, \omega')}{\sum_{\omega \in \Omega^\theta} \pi_0^\theta(\omega) \prod_{\tau=0}^{t-1} q^\theta(y_\tau | a_\tau, \omega)} \\
&\leq \frac{\sum_{\omega \in \Omega^\theta} (\pi_0^\theta(\omega) + \epsilon) \prod_{\tau=0}^{t-1} (q^\theta(y_\tau|a_\tau,\omega) + \epsilon)}{\sum_{\omega \in \Omega^\theta} \pi_0^\theta(\omega) \prod_{\tau=0}^{t-1} q^\theta(y_\tau|a_\tau,\omega)} \\
&= \max_{\omega \in \Omega^\theta} \left(1 + \frac{\epsilon}{\pi_0^\theta(\omega)}\right) \prod_{\tau=0}^{t-1} \left(1 + \frac{\epsilon}{q^\theta(y_\tau|a_\tau,\omega)}\right).
\end{aligned}
$$

We can choose ϵ to be sufficiently small so that λ_t does not exceed $\sqrt{\alpha} > 1$ for $t = 0, ..., T$ regardless of the action and outcome history.

Finally, note that for any $t > T$, we can write

$$
\lambda_t = \lambda_T \frac{\sum_{\omega' \in \Omega^{\theta'}} \prod_{\tau=T}^{t-1} \pi_\tau^{\theta'}(\omega') q^{\theta'}(y_\tau|a_\tau,\omega')}{\sum_{\omega \in \Omega^\theta} \prod_{\tau=T}^{t-1} \pi_\tau^\theta(\omega) q^\theta(y_\tau|a_\tau,\omega)} := \lambda_T \lambda_{T,t}.
$$

Recall that on histories $\hat{H}_T$ we have $\pi_T^\theta(\hat{\omega}) > 1/\sqrt{\alpha}$, so we can use the same arguments as in the proof of Theorem 2(ii) to show that $\mathbb{P}_S(\lambda_{T,t} \leq \sqrt{\alpha}, \forall t > T) > 0$. Since on $\hat{H}_T$ we have no switches before period T and ϵ is small enough such that $\lambda_T < \sqrt{\alpha}$, we have $\mathbb{P}_S(\lambda_t \leq \alpha, \forall t \geq 0) \geq \mathbb{P}_S(\hat{H}_T) \cdot \mathbb{P}_S(\lambda_{T,t} \leq \sqrt{\alpha}, \forall t > T) > 0$. $\square$

[39] The proof for the case of a continuous $\mathcal{Y}$ can be found in Appendix C.

B.5 Proof of Theorem 2(ii) and (iii)

Note that part (iii) immediately follows from part (ii). I now prove part (ii) in two steps.

Necessity. Suppose θ is globally robust at prior π_0^θ. By Theorem 1, we know that there must exist a p-absorbing SCE under θ. By identifiability, any SCE can only be supported by a pure belief, and hence $C^\theta \neq \emptyset$. Pick any prior π_0^θ such that $\pi_0^\theta(C^\theta) < 1/\alpha$.

Let us construct a competing model θ' such that it contains the prediction of C^θ and the true DGP. In particular, let $\Omega^{\theta'} = C^\theta \cup \{\omega^*\}$ and let predictions $q^{\theta'}$ satisfy that for all $a \in \mathcal{A}$,

$$q^{\theta'}(\cdot|a,\omega) = \begin{cases} q^\theta(\cdot|a,\omega) & \text{if } \omega \in C^\theta, \\ q^*(\cdot|a) & \text{if } \omega = \omega^*. \end{cases}$$

In addition, pick some $\epsilon \in (0,1)$ and let the prior $\pi_0^{\theta'}$ be such that

$$\pi_0^{\theta'}(\omega) = \begin{cases} (1-\epsilon)\frac{\pi_0^\theta(\omega)}{\pi_0^\theta(C^\theta)} & \text{if } \omega \in C^\theta, \\ \epsilon & \text{if } \omega = \omega^*. \end{cases}$$

Since θ' is correctly specified, by Lemma 2, on the paths where m_t eventually equals θ, the agent eventually only play actions in the support of a SCE almost surely, and her posterior converges to a supporting belief of the SCE, i.e. $\pi_t^\theta(C^\theta) \xrightarrow{\text{a.s.}} 1$. By construction

$$\ell_t(\theta') = (1-\epsilon)\sum_{\omega \in C^\theta} \frac{\pi_0^\theta(\omega)}{\pi_0^\theta(C^\theta)}\ell_t(\theta,\omega) + \epsilon\ell_t(\theta^*),$$

so we have

$$\frac{\ell_t(\theta')}{\ell_t(\theta)} = (1-\epsilon)\frac{\pi_t^\theta(C^\theta)}{\pi_0^\theta(C^\theta)} + \epsilon\frac{\ell_t(\theta^*)}{\ell_t(\theta)}.$$

Since θ' is correctly specified, by Lemma 2, on paths where m_t eventually equals θ, the first term almost surely converges to $(1-\epsilon)\frac{1}{\pi_0^\theta(C^\theta)}$. Since $\pi_0^\theta(C^\theta) < 1/\alpha$, there exists a small enough ϵ such that $\frac{\ell_t(\theta')}{\ell_t(\theta)} > \alpha$ for sufficiently large t, contradicting the assumption that m_t eventually equals θ.

Sufficiency. Suppose $C^\theta \neq \emptyset$ and $\pi_0^\theta(C^\theta) \geq 1/\alpha$. Pick any competing model θ' and a full-support prior $\pi_0^{\theta'}$. We will show that model θ persists against θ' at the given priors.

Define a new probability measure $\hat{\mathbb{P}}$ over the action and outcome histories H such that for any histories $\hat{H} \subset H$,

$$\hat{\mathbb{P}}\left(\hat{H}\right) = \sum_{\omega \in C^\theta} \frac{\pi_0^\theta(\omega)}{\pi_0^\theta(C^\theta)} \mathbb{P}_S^{\theta,\omega}\left(\hat{H}\right),$$

where $\mathbb{P}_S^{\theta,\omega}$ is the probability measure over histories induced by the agent switcher if the true DGP is identical to the DGP prescribed by θ and ω. Define the following process,

$$\hat{\lambda}_t := \frac{1}{\pi_0^\theta(C^\theta)} \frac{\ell_t(\theta')}{\sum_{\omega \in C^\theta} \frac{\pi_0^\theta(\omega)}{\pi_0^\theta(C^\theta)} \ell_t(\theta,\omega)}.$$

Then it is a martingale w.r.t. $\hat{\mathbb{P}}$ with $\mathbb{E}^{\hat{\mathbb{P}}}(\hat{\lambda}_0) = 1/\pi_0^\theta(C^\theta)$. By definition, $\hat{\lambda}_t \geq \lambda_t$, where the equality holds only if $\Omega^\theta = C^\theta$. By Ville's maximum inequality,

$$\hat{\mathbb{P}}(\lambda_t \leq \alpha, \forall t) \geq \hat{\mathbb{P}}(\hat{\lambda}_t \leq \alpha, \forall t) > 1 - \frac{1}{\pi_0^\theta(C^\theta)\alpha} \geq 0.$$

This then implies that there exists $\hat{\omega} \in C^\theta$ such that

$$\mathbb{P}_S^{\theta,\hat{\omega}}(\lambda_t \leq \alpha, \forall t) > 0.$$

Since θ has no traps, it is identifiable and all of its p-absorbing SCEs are quasi-strict. Identifiability implies that $\mathbb{P}_S^{\theta,\hat{\omega}}(\pi_t^\theta(\hat{\omega})) \xrightarrow{\text{a.s.}} 1$. With quasi-strictness, by Lemma 10, there exists $\epsilon > 0$ such that the optimal actions must be in the support of a SCE when $\pi_t^\theta(\hat{\omega}) > 1 - \epsilon$. Taken together, the no-trap conditions imply that there exists $T > 0$ such that with positive probability (measured by $\mathbb{P}_S^{\theta,\hat{\omega}}$), the agent plays only SCE actions after period T and never switches. Denote the set of such histories by $\hat{H}$. Moreover, for any $\hat{h} \in \hat{H}$, denote the observable history for the first T periods by $\hat{h}_{T-}$ and the history after the first T periods by $\hat{h}_{T+}$. Since T is finite, by absolute continuity (Assumption 2), for any $\hat{h} \in \hat{H}$, the history $\hat{h}_{T-}$ also occurs with positive probability under the true measure $\mathbb{P}_S$. Conditional on $\hat{h}_{T-}$, since the agent plays only SCE actions on $\hat{H}$ after the first T periods, the two probability measures $\mathbb{P}_S^{\theta,\hat{\omega}}$ and $\mathbb{P}_S$

over $\hat{H}$ are identical to each other. Therefore,

$$\begin{aligned}\mathbb{P}_S(\hat{H}_T) &= \sum_{\hat{h}\in\hat{H}_T} \mathbb{P}_S(\hat{h}_{T-})\mathbb{P}_S(\hat{h}_{T+}|\hat{h}_{T-}) \\ &= \sum_{\hat{h}\in\hat{H}_T} \mathbb{P}_S(\hat{h}_{T-})\mathbb{P}_S^{\theta,\hat{\omega}}(\hat{h}_{T+}|\hat{h}_{T-}) \\ &\geq \min_{\tilde{h}\in\hat{H}_T} \frac{\mathbb{P}_S(\tilde{h}_{T-})}{\mathbb{P}_S^{\theta,\hat{\omega}}(\tilde{h}_{T-})}\mathbb{P}_S^{\theta,\hat{\omega}}(\hat{H}_T) > 0.\end{aligned}$$

This means that with positive probability (under the true probability measure $\mathbb{P}_S$), the agent never switches to θ'. Therefore, model θ persists against θ'.

B.6 Proof of Corollary 2

Note that the proof of Theorem 2 does not use the assumption that $\alpha > 1$. Therefore, Corollary 2 is immediately implied by Theorem 2.

B.7 Proof of Theorem 3

To show that Theorem 1 continues to hold when $\alpha > K$, it suffices to show that a model θ is globally robust at some prior by the new definition if θ admits a p-absorbing SCE. Without loss of generality, take any $\Theta' = \{\theta^1, ..., \theta^K\} \subseteq \Theta$ and define for each $k \in \{1, ..., K\}$ a process $\{S_t^k\}_t$ as follows,

$$S_t^k = \frac{\sum_{\omega'\in\Omega^{\theta^k}} \pi_0^{\theta^k}(\omega') \prod_{\tau=0}^t q^{\theta^k}(y_\tau|a_\tau, \omega')}{\prod_{\tau=0}^t q^*(y_\tau|a_\tau)}.$$

Then for any $\eta \in (1, \alpha)$, we have

$$\mathbb{P}_D(S_t^k \leq \eta, \forall t \geq 0) \geq 1 - \frac{\mathbb{E}^{\mathbb{P}_D} S_0^k}{\eta} = 1 - \frac{1}{\eta}.$$

Hence, when η is sufficiently close to α,

$$\begin{aligned}&\mathbb{P}_D(S_t^k \leq \eta, \forall t \geq 0, \forall k \in \{1, ..., K\}) \\ \geq& 1 - \sum_{k=1}^K P_B(S_t^k > \eta \text{ for some } t \geq 0) \\ \geq& 1 - \frac{K}{\eta} > 0.\end{aligned}$$

The rest of the argument is identical to the proof in Appendix B.1. It follows that Theorem 1 also continues to hold when $\alpha > K$.

$\square$

B.8 Proof of Proposition 1

Suppose the agent's action space contains K elements, $a^1 < a^2 < ... < a^K$. Define function $h : [\underline{\omega}, \overline{\omega}] \to [\underline{\omega}, \overline{\omega}]$, such that $h(\omega)$ returns the KL-minimizer evaluated at the largest myopically optimal action against the degenerate belief δ_ω i.e. $h(\omega)$ minimizes $D_{KL}\left(q^*(\cdot|\overline{a}(\omega)) \parallel q(\cdot|\overline{a}(\omega), \hat{b}, \omega)\right)$ where $\overline{a}(\omega) = \max A^\theta(\delta_\omega)$. There exists an increasing sequence of intervals $\{(\omega_k, \omega_{k+1})\}_{k=0}^{K}$ such that $\omega_0 = \underline{\omega}$, $\omega_K = \overline{\omega}$, a^k is the unique myopically optimal action over (ω_{k-1}, ω_k) and both a^{k-1} and a_k are myopically optimal at ω_{k-1}. Function h is flat within each interval. If there exists a pure BN-E under model θ, then it must be supported by a degenerate belief at ω such that $h(\omega) = \omega$. By the assumption of complete parameter space, any pure BN-E must also be self-confirming, and any mixed BN-E cannot be self-confirming.

Suppose $\hat{b} > b^*$, then h jumps up discontinuously at all cutoffs $\{\omega_k\}_{1\leq k\leq K-1}$. Suppose there exists no solution to $h(\omega) = \omega$. Then since $h(\underline{\omega}) \geq \underline{\omega}$ and $h(\overline{\omega}) \leq \overline{\omega}$, we know that there must exist $\hat{k}$ such that $h(\omega) > \omega$ for all $\omega \in (\omega_{k^*-1}, \omega_{k^*})$ and $h(\omega') < \omega'$ for all $\omega' \in (\omega_{k^*}, \omega_{k^*+1})$. But this contradicts the observation that h jumps up at ω_{k^*}. It also immediately follows that there exists a solution $\hat{\omega}$ to $h(\hat{\omega}) = \hat{\omega}$ such that $h(\omega') > \omega'$ for $\omega' < \hat{\omega}$ and $h(\omega'') < \omega''$ for $\omega'' < \hat{\omega}$. Let $\hat{a}$ be the unique myopically optimal action at $\delta_{\hat{\omega}}$. Then $\hat{a}$ is a pure self-confirming equilibrium, supported by the generate belief at $\hat{\omega}$. By the assumption of complete parameter space, $\omega \in \Omega^\theta$, and thus $\hat{a}$ is also a self-confirming equilibrium under θ. Note that $\hat{a}$ is uniformly strict. By Corollary 1, model θ is globally robust.

Now suppose the agent is underconfident, then h jumps down discontinuously at the cutoffs $\{\omega_k\}_{1\leq k\leq K-1}$. Hence, there exists at most one solution to $h(\omega) = \omega$. Suppose there exists a SCE $\sigma^\dagger$ when the agent believes his ability is given by $\tilde{b}$. Then by the upper-hemicontinuity of A^θ, when $\hat{b}$ is lower than but sufficiently close to $\tilde{b}$, there exists some $\hat{\omega} > \omega^*$ such that $g(a^\dagger, \hat{b}, \hat{\omega}) = g(a^\dagger, b^*, \omega^*)$, where $a^\dagger = \max \operatorname{supp}(\sigma^\dagger)$ and is the unique myopically optimal action against $\delta_{\hat{\omega}}$. It follows that $a^\dagger$ is a uniformly strict SCE under θ. Since there always exists a SCE when the agent is correctly specified, i.e. $\tilde{b} = b^*$, we infer that model θ is globally robust when $b^* - \hat{b}$ is sufficiently small.

Suppose instead that there is no solution to $h(\omega) = \omega$ when the agent's self-perception is given by $\hat{b}$. If so, there exists no SCE under model θ. By Theorem 1, θ

is not locally or globally robust. By continuity, $h(\omega) = h(\omega)$ also does not admit any solution at $\tilde{b}$ if it is sufficiently close to $\tilde{b}$. Therefore, there exists an open neighborhood around $\hat{b}$ such that model θ is not locally or globally robust.

B.9 Proof of Proposition 3

It suffices to show that the agent makes a switch to $\hat{\theta}$ with positive probability. It then follows from Proposition 2 that $\hat{\theta}$ is eventually adopted forever with positive probability.

Define a new probability measure $\hat{\mathbb{P}}$ over the action and outcome histories H such that for any histories $\hat{H} \subset H$,

$$\hat{\mathbb{P}}\left(\hat{H}\right) = \pi_0^{\hat{\theta}}(\omega^L)\mathbb{P}_S^{\hat{\theta},\omega^L}\left(\hat{H}\right) + \pi_0^{\hat{\theta}}(\omega^R)\mathbb{P}_S^{\hat{\theta},\omega^R}\left(\hat{H}\right),$$

where $\mathbb{P}_S^{\hat{\theta},\omega}$ is the probability measure over histories induced by the agent switcher if the true DGP is identical to the DGP prescribed by $\hat{\theta}$ and ω. Then $\ell_t(\theta,\omega^M)/\ell_t(\hat{\theta})$ is a martingale w.r.t. $\hat{\mathbb{P}}$ with an expectation of 1. Hence, for any $\eta > 1$, the probability that $\ell_t(\theta,\omega^M)/\ell_t(\hat{\theta}) \le \eta$ for all t is positive (measured by $\hat{\mathbb{P}}$). Since a^M is the only SCE under model θ, by Lemma 2 the agent almost surely eventually play a^M on the paths where the model choice eventually equals θ. If so, the agent's posterior π_t^θ almost surely converges to δ_{ω^M}. In summary, on paths where m_t eventually equals θ, it happens with positive probability (measured by $\hat{\mathbb{P}}$) that $\ell_t(\theta,\omega^M)/\ell_t(\hat{\theta}) \le \eta$ for all t and $\pi_t^\theta \xrightarrow{\text{a.s.}} \delta_{\omega^M}$. This then implies that for any $\epsilon > 0$, we can construct a finite sequence of outcome realizations $(y_0, ..., y_{t-1})$ such that $\ell_t(\theta,\omega^M)/\ell_t(\hat{\theta}) \le \eta$ for all $t \le T$ and $\pi_T^\theta \in B_\epsilon(\delta_{\omega^M})$. Moreover, since T is finite, this sequence of outcomes are also realized with positive probability under the true measure $\mathbb{P}_S$. Notice that

$$\frac{\ell_T(\hat{\theta})}{\ell_T(\theta)} = \pi_T^\theta(\omega^M)\frac{\ell_T(\hat{\theta})}{\pi_0^\theta(\omega^M)\ell_t(\theta,\omega^M)} \ge (1-\epsilon)\frac{\eta}{\pi_0^\theta(\omega^M)},$$

where the right-hand side is strictly larger than α when $\pi_0^\theta(\omega^M) < 1/\alpha$ if ϵ is close enough to 0 and η is close enough to 1. Therefore, the agent makes a switch from θ to $\hat{\theta}$ with positive probability.

C Supplemental Appendix

C.1 Examples Omitted from Section 4

Example 4 (A p-absorbing mixed SCE)**.** Consider a dogmatic modeler's problem, where there are two actions $\mathcal{A} = \{1, 2\}$ and three parameters $\Omega^\theta = \{1, 1.5, 2\}$ inside the parameter space of model θ. The agent's payoff is simply the outcome y_t, with the true DGP being the normal distribution $N(0.25, 1)$ for all actions. Model θ is misspecified, predicting that $y_t \sim N((\omega - a_t)^2, 1)$. Note that every mixed action is a self-confirming equilibrium, with the supporting belief assigning probability 1 to the parameter value of 1.5. Here, every mixed SCE is p-absorbing since its support contains every action that can be played by the agent. But her action sequence may never converge. To see that, notice that a belief that assigns larger probability to $\omega = 1$ than $\omega = 2$ leads to action $a = 2$, but such play in turn induces her to attach lower probability to $\omega = 1$ than $\omega = 2$ and leads to action $a = 1$. Nevertheless, Corollary 1 tells us that the aforementioned SCE is indeed p-absorbing.

Example 5 (A self-confirming equilibrium that fails to be p-absorbing)**.** Consider a dogmatic modeler's problem, where there are two actions $\mathcal{A} = \{1, 3\}$ and three parameters $\Omega^\theta = \{1, 2, 3\}$ inside the parameter space of model θ. The agent's payoff is the absolute value of the outcome, $|y_t|$, with the true DGP of y_t given by a normal distribution $N(1, 1)$ for all actions. Consider a misspecified model θ that predicts $y_t \sim N(\omega - a_t, 1)$. Note that θ admits a single self-confirming equilibrium in which the agent plays $a^* = 1$ with probability 1, supported by a belief that assigns probability 1 to $\omega^* = 2$. However, this SCE is not p-absorbing. To see that, notice that the agent is indifferent between the two actions when the parameter takes the value of 2. When the agent keeps playing $a = 1$, the parameters 1 and 3 fit the data equally well on average, so their log-posterior ratio is a random walk which a.s. crosses 1 infinitely often. However, the high action $a = 3$ is strictly optimal against any belief that assigns a higher probability to $\omega = 1$ than $\omega = 3$. Hence, the high action must be played infinitely often almost surely.

C.2 Micro-Foundation for Application 5.2

In this subsection I specify the payoff structure for the news consumption problem in Application 5.2, which provides a micro-foundation for Assumption 3.

To do this, we first extend the learning framework introduced in Section 3 to allow

$E^{\theta}(\text{payoff}\mid a,\omega)$	ω^L	ω^M	ω^R
a^L	1.4	1.1	1
a^M	1.25	1.15	1.25
a^R	1	1.1	1.4

$E^{\hat{\theta}}(\text{payoff}\mid a,\omega)$	ω^L	ω^R
a^L	1.2	1
a^M	1.15	1.15
a^R	1	1.2

Table 5: Expected payoffs under model θ (left) and expected payoffs under model θ' (right).

for an unobserved payoff that may depend on an unknown state. That is, besides the observable payoff jointly determined by the action and the random outcome $u(a_t, y_t)$, there may exist an unobserved payoff $\tilde{u}(a_t, \omega)$ that depends on the action and a fundamental state $\omega \in \Omega$. Under any subjective model θ, the agent maximizes the sum of the observed and the unobserved payoff given her belief over the fundamental state and possibly other parameters. This maximization gives rise to an optimal-action correspondence $A^{\theta}_m : \Delta\Omega^{\theta} \rightrightarrows \mathcal{A}$, which we can use to define a self-confirming equilibrium. All results in Section 4 remain unchanged.

Subscribing to media outlets provide entertainment value. Media outlets produce higher quality news reports if the story is aligned with their political leaning. If the agent subscribes to media a^L, she earns an emotional utility of 1 iff she receives a l story; similarly, if she subscribes to media a^R, she earns an emotional utility of 1 iff she receives a r story. If she subscribes to the neutral media a^M, she earns a constant emotional payoff of 0.65.

Subscribing to media outlets also provide valuable information. In additional to subscribing to a media outlet a_t, the agent takes an outside action $v_t \in \{v^L, v^M, v^R\}$ upon receiving the story y_t. The agent earns a payoff of 1 if she takes v^L in state ω^L and v^R in state ω^R, but in state ω^M she earns a constant payoff of 0.5 by taking any action. Note that it is optimal for the agent to follow the story she receives in each period.

In Table 5, I summarize the expected total payoffs associated with each action under model θ and model θ'. It is then straightforward to verify Assumption 3.

C.3 Examples Omitted from Section 6

I provide two examples below to substantiate the observation in Footnote 33. Example 6 presents a scenario in which θ persists against θ^1 and θ^2 separately but does not persist against $\{\theta^1, \theta^2\}$, while Example 7 shows an opposite scenario.

Example 6. Let x^1 and x^2 be two i.i.d. normally distributed variables, both with mean 0 and variance 1. Suppose x^3 and x^4 are also i.i.d. normally distributed but with mean 1 and variance 1. Suppose the agent can play one of two actions in each period, $\mathcal{A} = \{1, 2\}$ and uses subjective models to learn about the mean of each element in (x^1, x^2, x^3, x^4). Her flow payoff is given by $a \cdot (x^4 - x^3)$. Hence, she would like to play $a = 2$ if $\overline{x}^4 > \overline{x}^3$ and play $a = 1$ if $\overline{x}^3 > \overline{x}^4$. However, x^1 and x^3 are only observable when $a = 1$, while x^2 and x^4 are only observable when $a = 2$. That is, the outcome y is given by (x^1, x^3) when $a = 1$ and given by (x^2, x^4) when $a = 2$. She entertains an initial model θ and two competing models, $\{\theta^1, \theta^2\}$, each of which is equipped with a binary parameter space. The predictions of each model are summarized by the following table. The predicted means are independent of the actions taken.

θ	ω^1	ω^2
$(\overline{x}^1, \overline{x}^2, \overline{x}^3, \overline{x}^4)$	$(1, 1, 1, 0)$	$(1, 1, 0, 1)$

θ^1	$\omega^{1\prime}$	$\omega^{2\prime}$
$(\overline{x}^1, \overline{x}^2, \overline{x}^3, \overline{x}^4)$	$(1, 0, 1, 0)$	$(1, 0, 0, 1)$

θ^2	$\omega^{1\prime\prime}$	$\omega^{2\prime\prime}$
$(\overline{x}^1, \overline{x}^2, \overline{x}^3, \overline{x}^4)$	$(0, 1, 1, 0)$	$(0, 1, 0, 1)$

Notice that there are two strict (and thus p-absorbing) Berk-Nash equilibria under θ: (1) $a = 1$ is played w.p. 1, supported by the belief that assigns probability 1 to ω^1; (2) $a = 2$ is played w.p. 1 , supported by the belief that assigns probability 1 to ω^2. First observe that θ persists against θ^1 at a prior π_0^θ that assigns sufficiently high belief to ω^1. This follows from the fact that the likelihood ratio between θ and θ^1 is always 1 when $a = 1$ is played, and that the equilibrium is p-absorbing. Analogously, θ persists against θ^2 at a prior π_0^θ that assigns sufficiently high belief to ω^2. However, notice that θ does not persist against $\{\theta^1, \theta^2\}$ at any priors and policies, because regardless of the actions taken by the agent, at least one of θ^1 and θ^2 would fit the data strictly better than θ, prompting the agent to adopt θ^1 and θ^2 infinitely often.

Example 7. Let y be a normally distributed variable with mean 0 and variance 1, whose distribution is independent of actions. The agent can play one of two actions in each period, $\mathcal{A} = \{1, 2\}$ and uses subjective models to learn about the mean of y. Her flow payoff is given by $a \cdot y$. She entertains an initial model θ and two competing models, $\{\theta^1, \theta^2\}$. Model θ^1 has a single parameter and perfectly matches the true DGP, while models θ and θ^2 both have a binary parameter space. The predictions about $\overline{y}$ of each model are summarized by the following table.

θ	ω^1	ω^2
a^1	-1	1
a^2	-2	1

θ^1	ω'
a^1, a^2	-1

θ^2	ω''
a^1, a^2	2

Suppose the agent's prior satisfies that $\pi_0^\theta(\omega^1) = 1-\pi_0^\theta(\omega^0) = \frac{0.5}{\alpha} < \frac{1}{\alpha}$. First consider what happens when the agent has only one competing model, θ^1. By the Law of Large Numbers, the likelihood ratio between θ^1 and θ eventually exceeds α almost surely because

$$\begin{aligned}\frac{\ell_t(\theta^1)}{\ell_t(\theta)} &= \frac{\prod_{\tau=0}^{t-1} q^{\theta^1}\left(y_\tau|a_\tau,\omega'\right)}{\prod_{\tau=0}^{t-1} q^{\theta}\left(y_\tau|a_\tau,\omega^1\right)\pi_0^\theta\left(\omega^1\right)+\prod_{\tau=0}^{t-1} q^{\theta}\left(y_\tau|a_\tau,\omega^2\right)\pi_0^\theta\left(\omega^2\right)} \\ &= \frac{\prod_{\tau=0}^{t-1} q^{*}\left(y_\tau|a_\tau\right)}{\prod_{\tau=0}^{t-1} \mathbf{1}_{(a_\tau=a^1)} q^{*}\left(y_\tau|a_\tau\right)\pi_0^\theta\left(\omega^1\right)+\xi(h_t)} \\ &\geq \frac{\prod_{\tau=0}^{t-1} q^{*}\left(y_\tau|a_\tau\right)}{\prod_{\tau=0}^{t-1} \frac{1}{\alpha} q^{*}\left(y_\tau|a_\tau\right)+\xi(h_t)}\end{aligned}$$

where $\frac{\xi(h_t)}{\prod_{\tau=0}^{t-1} q^*(y_\tau|a_\tau)}$ converges to 0 almost surely. Threrefore, θ does not persist against θ^1 under prior π_0^θ.

However, model θ persists against $\Theta' \coloneqq \{\theta^1, \theta^2\}$ at prior π_0^θ. First notice that for any $a_0 \in \mathcal{A}$, there exists some y_0 sufficiently large such that

$$\ell_1(\theta^2) > \alpha \cdot \max\{\ell_1(\theta), \ell_1(\theta^1)\}$$

and thus the agent switches to θ^2 in the beginning of period 1. As a result, the agent plays $a_1 = a^2$ in period 1 since it is the strictly dominant strategy under θ^2. But then we could find some sufficiently small y_1 such that the following two inequalities hold:

$$\ell_2(\theta) > \alpha \cdot \max\{\ell_2(\theta^1), \ell_2(\theta^2)\},$$

$$\pi_2^\theta(\omega^1) = \frac{\pi_0^\theta(\omega^1) q^\theta(y_0|a_0,\omega^1) q^\theta(y_1|a_1,\omega^1)}{\sum_{\omega\in\{\omega^1,\omega^2\}} \pi_0^\theta(\omega) q^\theta(y_0|a_0,\omega) q^\theta(y_1|a_1,\omega)} > \max\{\frac{1}{\alpha}, c\},$$

where c is chosen such that while adopting θ^1, the agent finds a^1 payoff-maximizing if her belief assigns a probability higher than c to ω^1, i.e. $\pi^\theta(\omega^1) > c$. The first inequality implies that the agent switches back to θ in the beginning of period 2. The second inequality, together with the observation that the pure strategy a^1 is a uniformly strict self-confirming equilibrium supported by the belief that assigns probability 1 to ω^1, ensures that with positive probability, the agent plays action a^1 forever, provided that her decisions are made based on model θ. But notice that on those paths, the agent indeed no longer switches to other models after period 1 because for $t > 2$,

$$\frac{\ell_t(\theta^1)}{\ell_t(\theta)} < \frac{\ell_2(\theta^1)}{\ell_2(\theta)} \frac{\prod_{\tau=2}^{t-1} q^*\left(y_\tau|a_\tau\right)}{\prod_{\tau=2}^{t-1} q^*\left(y_\tau|a_\tau\right) \pi_2^\theta\left(\omega^1\right)} < 1 < \alpha.$$

Since outcomes y_0 and y_1 that satisfy the aformentioned properties are drawn with positive probability, we conclude that θ persists against $\Theta^c := \{\theta^1, \theta^2\}$ at prior π_0^θ.