Misinterpreting Yourself*

Paul Heidhues
DICE, University of Düsseldorf

Botond Kőszegi
University of Bonn

Philipp Strack
Yale University

February 12, 2025

Abstract

We model an agent who stubbornly underestimates his undesirable motives, and misattributes behavior resulting from these motives to other considerations. Applying the framework to partially naive present bias, we show that in many stable situations, the agent eventually predicts his behavior well. This "apparent sophistication" implies that existing empirical tests of sophistication in intertemporal choice can reach incorrect conclusions. The agent's unrealistic self-view does, however, manifest itself in several ways. First, he comes to act in a more present-biased manner than a sophisticated agent. Second, he systematically mispredicts how he will react when circumstances change, such as when incentives for forward-looking behavior increase or he is placed in a new, ex-ante identical environment. Third, even for physically non-addictive products, he follows addiction-like consumption dynamics that he does not anticipate. Fourth, he holds beliefs that — when compared to those of other agents — display puzzling correlations between logically unrelated issues. We also apply our framework to provide a novel perspective on implicit bias, showing that a naive endorsement of egalitarian standards may be harmful for behavior. Consequently, making a person aware of his subsconscious prejudice can mitigate his biased behavior even absent eliminating the prejudice itself.

Keywords: present bias, naivete, sophistication, misspecified learning, apparent sophistication, implicit bias, prejudice

JEL Codes: D91, D83, D11

*We are grateful to Peter Andre, Drew Fudenberg, Jana Gieselmann, Marc Kaufmann, Ricky Li, George Loewenstein, Rani Spiegler, Dmitry Taubinsky, and Sevgi Yuksel for insightful discussions, and to seminar and conference audiences for comments. Heidhues and Kőszegi thank the European Research Council for financial support under Grant #788918.

"Sooner or later, everyone invents a story that they believe to be their life."

—Max Frisch

1 Introduction

Several lines of research raise the possibility that people have stubborn misconceptions about their own motives or inclinations, especially ones that can be seen as flaws. In full or partial naivete regarding present bias, a person underestimates his tendency to underweight the future (e.g., Augenblick and Rabin, 2019). According to a common interpretation, implicit racial bias involves a racist person who thinks of himself as non-racist (e.g., Bertrand and Duflo, 2017, Carlana et al., 2022). And a pervasive theme in academic and non-academic discussions is that many aggressive individuals do not self-identify as such (e.g., Eisikovits and Buchbinder, 1997, Anderson and Umberson, 2001).

A common implication of the above misconceptions is that the person might mispredict his own behavior. A present-biased smoker may predict that he will quit soon, and then not do it. An employer with implicit racial bias may think that he is treating all applicants equally, and then end up with a racially unbalanced team. And an aggressive individual may believe that he will be nicer with his new partner than with his previous one, and then be just as he was before.

In this paper, we ask the natural question: what does a person make of such mispredictions? Building on the quickly growing literature on misspecified learning, we propose that he adjusts his beliefs about other aspects of himself or the environment, which in turn feeds back into his behavior. A person who smokes more than he intended and predicted, for instance, may develop the exaggerated belief that smoking helps him concentrate or socialize, or that it is not as harmful for him as for others. We build a general machinery for analyzing the implications of such misinferences, and then apply our framework to partially naive present bias and implicit bias. We show that the combination of partial naivete and the resulting self-justificatory views (i) can explain empirical patterns that existing theories have difficulty simultaneously accounting for, (ii) identifies a flaw in how economists usually think about sophistication, and (iii) makes additional novel predictions. And we show that in the context of implicit bias, the conscious endorsement of egalitarian attitudes may make behavior more prejudiced.

Section 2 presents our general model. In each period t, the agent observes a period-specific shock $s_t = \Theta + \epsilon_t$, where Θ is an unknown, normally distributed, time-independent fundamental, and the ϵ_t are independent, mean-zero, normally distributed errors. The agent then chooses an action a_t to maximize the expectation of $v(a_t, s_t, \Theta)$. Crucially, his self-knowledge is limited in two ways. First, after period t he remembers a_t but not s_t. This captures the idea that individuals may not remember or even have direct access to all the reasons behind their actions. Second, when interpreting his past behavior, the agent is misspecified regarding his motives, hoping and thinking that he had been maximizing the expectation of $\tilde{v}(a_t, s_t, \Theta)$. This captures the mistaken self-conceptions motivating our paper. Beyond these limitations, however, the agent is rational: he has a correct prior about Θ, and updates his beliefs using Bayes' rule.

We define an intuitive notion of a stable belief about the fundamental that we call a *self-observation equilibrium (SOE)*. Suppose that the agent believes the fundamental to be $\tilde{\theta}$ and acts upon this belief for a long time. Based on $\tilde{\theta}$ and his perceived utility function $\tilde{v}$, he can infer the shocks $\tilde{s}_t$ that he believes have driven his actions. He can then ask: what fundamental best explains the distribution of $\tilde{s}_t$? If it is $\tilde{\theta}$, then $\tilde{\theta}$ is a coherent belief that the agent sees no reason to change, so it is an SOE.[1] An easy-to-check SOE arises when $\tilde{\theta}$ *perfectly* explains the distribution of $\tilde{s}_t$, and hence it perfectly explains — and predicts — the distribution of actions it generates. Going further, in the appendix we identify conditions under which the long-run limit of the beliefs the agent forms based on his actions is an SOE. In a stable environment, therefore, one can understand long-run beliefs and behavior by studying SOE's. And in such an environment, the agent often learns to predict his behavior well, or even perfectly.

For much of our paper, we restrict attention to problems where the fundamental Θ and the shock s_t affect the optimal action in the same direction. Such an "equi-directional" specification is natural when uncertainty pertains to how much one should consume, so that information affects optimal actions in different periods in the same direction. If the agent learns that exercise is beneficial, for instance, then he should exercise more in all periods. We establish a general property of SOE with equi-directional problems: for any v and $\tilde{v}$ that have different implications for optimal actions, the agent's misspecification lowers his utility according to his wished-for preferences $\tilde{v}$. In other words, whenever the agent thinks that his motives are less flawed, his behavior actually becomes more flawed.

[1] This definition adapts the spirit of Berk-Nash equilibrium (Esponda and Pouzo, 2016) to our different setting where the agent misinterprets his own actions.

In Section 3, we turn to our main application, partially naive present bias. Here, the agent's true objective v discounts the future consequences of today's action by the factor β, but his perceived objective $\tilde{v}$ instead applies the discount factor $\tilde{\beta} > \beta$. Several new insights emerge.

First, we identify a novel type of harm associated with partial naivete. Namely, as a manifestation of our general result on the self-defeating nature of misspecification, in equi-directional problems partial naivete is always welfare-decreasing. To illustrate, suppose that the agent chooses the level of harmful consumption a_t in each period t, and the fundamental Θ and signal s_t pertain to his instantaneous marginal utility of consumption. A basic implication of naive present bias is that the agent starts off consuming more than he wants or expects. To explain his high consumption, he eventually overestimates his marginal utility. A college student who goes to too many parties due to his present bias, for instance, comes to exaggerate how fun the average party is. This false belief increases his consumption, moving it even further from that of the more patient person he thinks he is. Alternatively, suppose that Θ and s_t pertain to the future harm from consumption. Then, to explain his high consumption, the agent comes to believe that the product is not very harmful. A smoker may, for instance, believe that alternative activities are just as harmful as smoking, or that smoking helps him concentrate. This again exacerbates overconsumption and is therefore welfare-decreasing. In contrast to this unambiguously harmful mechanism, in existing models naivete acts only through the (often weak) intertemporal interdependence of consumption decisions, and its effect on the agent's welfare can be positive or negative. Hence, the main harm from naivete may stem from its impact on other beliefs through misspecified learning.

Second, although in a stable environment the agent may learn to predict his behavior well, he also displays patterns that distinguish him from a realistic agent. In terms of predictions, he is generally incorrect about how he will react to a change in the environment. As a case in point, because he continues to overestimate the weight he will put on the future, he overestimates his response to future incentives. For instance, a smoker may understand that he will continue to smoke at the same rate during the on-going high-pressure period at his job, but also incorrectly believe that he will quit once the stressful period is over. And because the agent does not draw conclusions about his present bias, he mispredicts what he will do when a new fundamental applies, e.g., when he starts a new diet after previous ones have failed. In terms of beliefs, the agent's naivete can lead to multiple types of incorrect beliefs, such as the simultaneous beliefs that smoking is beneficial for socialization and that alternative activities are risky. Hence, across a population of agents with different levels of naivete, beliefs about such logically and factually unrelated issues will be correlated.

As a third insight, our theory has implications for the substantial empirical literature on whether individuals are sophisticated in intertemporal choice. This literature — and everyday thinking in the profession — overwhelmingly presumes that a person can be considered sophisticated if he correctly predicts his behavior in the situation at hand (see Ericson and Laibson, 2019, for a review). Such a prediction test, however, is vulnerable to the "apparent sophistication" of a naive agent who acts suboptimally yet predicts his behavior perfectly in an SOE.

While the possibility of apparent sophistication has never been explicitly tested, some patterns suggest that it is empirically relevant. Notably, we argue that our model better explains the behavior of experienced payday-loan borrowers in Allcott et al. (2022) than true sophistication does, and conclusions drawn from estimates of these borrowers' present bias are overly optimistic. Correctly specified learning about present bias implies that as a person learns, his beliefs adjust downwards to his true present bias; whereas our model predicts that *measures of* his present bias adjust *upwards* to his belief. The latter is closer to what Allcott et al. find than the former. Furthermore, exactly as our model predicts, Allcott et al. document that borrowers mispredict their responses to a change in incentives. Finally, our model's prediction that the agent learns to forecast his behavior in a stable environment but does not transfer his knowledge to other environments naturally accounts for a conspicuous general pattern in the literature. Namely, while individuals appear to quickly become sophisticated in some specific settings, in almost all experiments and other studies on less familiar choices, they are quite naive.

Fourth, although our model presumes a physically non-addictive product, it generates dynamic behavior with some features reminiscent of those in existing addiction models. Because the agent accounts for higher past consumption using the belief that his marginal utility is high, as in intertemporal-complementarities models (e.g., Becker and Murphy, 1988, Gruber and Kőszegi, 2001) his current consumption is increasing in shocks to past consumption. For the same reason, his consumption profile is often increasing over time. Unlike in the intertemporal-complementarities approach, however, the agent's consumption does not respond to future prices. In addition, because the agent comes to overestimate his marginal utility, his consumption eventually becomes too high even from the perspective of his current self. Similarly to mistake-based models (e.g., Bernheim and Rangel, 2004) and clinical descriptions of addiction, therefore, consumption is on average not worth it. Finally, our model predicts an intertemporal pattern in the response to news that is unlike in either previous approach. If a smoker receives new negative information regarding the health effects of smoking, he cuts back, but by less than he expects. To account for his lackluster response, he comes to believe in higher benefits from smoking, diminishing or reversing the effect of the news. We argue that some empirical findings provide tentative support for our combination of predictions. In particular, our model offers one explanation for the phenomenon of "behavioral" — as opposed to physiological — addiction described by psychologists.

In Section 4, we use our framework to provide a novel perspective on implicit bias or prejudice. In our model, implicit bias arises when a person consciously endorses and believes he holds egalitarian preferences ($\tilde{v}$), but he acts on subconscious negative attitudes toward another group (v). To explain why he acts in a biased way, he develops the belief that the other group is less deserving. Our theory therefore predicts that among people who view themselves as egalitarian, inaccurate beliefs and taste-based discrimination are positively correlated. Due to this combination, the agent acts in a more biased manner than if he was honest to himself about his bias. Accordingly, the transformation of prejudice from explicit to implicit forms may have been detrimental. Conversely, however, making a person aware of his subsconscious prejudice can mitigate his biased behavior even absent eliminating the prejudice itself.

We conclude the paper in Section 5 by briefly studying a non-equi-directional setting with present bias. Such specifications are natural if uncertainty pertains to the optimal timing of actions. Suppose that the agent repeatedly allocates work between now and later, s_t describes how busy he is right now, and Θ describes how busy he is on average. Then, a higher s_t means that he should work less now, but a higher Θ means that he should work more now. We show that the agent still becomes apparently sophisticated, but the welfare implications of naivete are now ambiguous. At one end, naivete can benefit the agent in an extreme sense: he may come to behave as if he was time-consistent and correctly learned the fundamental, despite neither of these being true. Since he underweights his future busyness when deciding how much to work, he tends to put off working. Given his naivete, he interprets this behavior as indicating that he is often rather busy. On an average day, therefore, he believes that he is less busy than usual, and is willing to work a decent amount — exactly as a time-consistent agent would. This, however, is only a possibility result. Depending on the specification, naivete can also harm the agent by inducing him to allocate too much work to early periods.

Related Literature As we have indicated, our paper builds on and belongs to a growing literature on learning with "misspecified" models. The core assumption of this literature is that individuals update their beliefs using an incorrect understanding of the situation. Researchers have studied misspecifications about the laws of Bayesian inference (e.g., Rabin, 2002, Rabin and Vayanos, 2010, Benjamin et al., 2016), the causal structure of outcomes (Spiegler, 2016, 2020, Levy et al., 2022), the distribution of others' types (Ettinger and Jehiel, 2010, Levy and Razin, 2017, Bohren and Hauser, 2019, Frick et al., 2020, 2022), statistical correlations (He, 2021), individual ability (Heidhues et al., 2018, Bohren et al., 2019, Murooka and Yamamoto, 2021, Ba and Gindin, 2023), market or technological parameters (Nyarko, 1991, Esponda and Pouzo, 2016, Fudenberg et al., 2017, Heidhues et al., 2021), and memory (Fudenberg et al., forthcoming).[2] It has been known at least since the prophet Matthew, however, that people are often most miscalibrated about their own flawed inclinations,[3] and our paper analyzes consequences of this important type of misspecification. Formally, our theory — in which the agent interprets actions that he does not know the precise reasons for — is most similar to those on social learning with misspecification (Bohren, 2016, Bohren and Hauser, 2019, forthcoming, Frick et al., 2020, 2021b), but the particular model and questions are very different. And at the technical level, our paper contributes to the literature studying the convergence of belief processes with incorrect inferences and endogenous actions (Esponda et al., 2021, Fudenberg et al., 2021b, in addition to papers cited above), which in general remains an unsolved problem.

Our paper also builds on empirical and experimental research on the general hypothesis that people may use misspecified models in making inferences. For instance, Benjamin (2019) reviews a large literature on inferential mistakes, Bohren et al. (2019) and Bohren et al. (forthcoming)

[2] In studying implications of imperfect memory, our paper is also related to other recent economics research (e.g., Mullainathan, 2002, Bodoh-Creed, 2019, Bordalo et al., 2020, 2021, Wachter and Kahana, 2021, Kőszegi et al., 2022). All of this previous research posits that recalled memories are sensitive to the current context or decision, whereas our agent's recollections are not.

[3] Matthew 7:3 reads: "Why do you see the speck that is in your brother's eye, but don't consider the beam that is in your own eye?" This is an attack on hypocrites who notice a small flaw in others but are blind to a large flaw in themselves. More recently, Fedyk (2021) documents that the average subject understands the present bias of other subjects almost perfectly, but is completely naive about his own present bias.

document patterns suggesting inaccurate inferences about groups in discrimination settings, and Goette and Kozakiewicz (2018) find that subjects update according to a misspecified model when their own ability is involved. More directly related to our model, two papers document misspecified learning about one's own preferences, albeit not from actions. Haggag et al. (2018) find that a subject who was experimentally made more thirsty at the time of trying a new drink has higher demand for the drink on future occasions. They also document that consumers who visited an amusement park during a nice weather shock are more likely to return later and to recommend the park to others. In both cases, individuals misattribute their temporary tastes to a permanent preference. In related work, Bushong and Gagnon-Bartsch (2020) find that a subject for whom a task was surprisingly difficult is less likely to want to do it again than a subject who expected the task to be difficult. Here, subjects misattribute the unpleasant surprise to a permanent dislike of the task.

2 Framework and General Results

2.1 Model

Basics There are infinitely many periods, $t = 1, 2, \ldots$. At the beginning of each period t, the agent observes a temporary shock or signal $s_t = \Theta + \epsilon_t$, where $\Theta \in \mathbb{R}$ is an unknown time-invariant fundamental and the ϵ_t are i.i.d. random variables with mean 0. Afterwards, the agent chooses an action $a_t \in A \subseteq \mathbb{R}$ aiming to maximize the expectation of $v(a_t, s_t, \Theta)$. Hence, the utility function v captures his true motives. Both Θ and the ϵ_t are normally distributed, and the agent's prior about Θ is correct (i.e., it equals the distribution from which Θ is drawn). Assuming normally distributed shocks simplifies the analysis of convergence, and while inconsequential for long-run behavior, imposing a correct prior simplifies our presentation of results on learning dynamics.

Self-Knowledge The agent's self-knowledge is limited in two ways. First, he has limited memory: after period t, he remembers the action a_t he took but not the shock s_t he observed. He also does not remember, or does not observe, the realized utility $v(a_t, s_t, \Theta)$. Second, he has an incorrect self-view: when interpreting his past behavior or predicting his future behavior, he believes that his objective had been or will be to maximize the expectation of $\tilde{v}(a_t, s_t, \Theta)$. He does not update his belief $\tilde{v}$ about his preferences.

Beyond his biased self-view, the agent understands the decision problem correctly, and in each period computes his belief about Θ according to Bayes' rule. We denote his belief at the beginning of period t, when he has not yet observed the shock s_t, by μ_t. By definition, μ_1 is his prior, and μ_t is obtained from μ_1 by conditioning on past actions $(a_1, \ldots, a_{t-1})$ while assuming that those actions were chosen to maximize the expectation of $\tilde{v}$. Before choosing his action a_t to maximize the expectation of $v(a_t, s_t, \Theta)$, the agent updates μ_t by conditioning also on the signal s_t.[4]

We assume that v and $\tilde{v}$ are twice differentiable and single-peaked in a_t. The latter implies that for any θ and s, there are unique optimal actions

$$\pi_\theta(s) = \arg\max_{a \in A} v(a, s, \theta) \qquad \text{and} \qquad \tilde{\pi}_\theta(s) = \arg\max_{a \in A} \tilde{v}(a, s, \theta)\,.$$

We impose that the functions π_θ and $\tilde{\pi}_\theta$ are either strictly increasing in s for all θ or strictly decreasing in s for all θ, and hence invertible.[5] Invertibility implies that our results do not derive solely from a lack of information: remembering the past actions a_t is sufficient for inferring the shocks s_t and therefore learning the fundamental Θ in the long run.[6] This rules out slow learning as defined by Frick et al. (2021b), where the actions of an agent who is confident in the fundamental reveal no information.

Comments We think of $\tilde{v}$ as capturing motives that the agent not only believes to have, but also wishes to have instead of his true motives v. In the case of intertemporal choice, for instance, a partially naive present-biased agent not only expects, but also prefers to be less present-biased than he actually is. Much research observes that individuals often maintain such positive biases in ego-relevant domains. Furthermore, the individuals in question are typically adults who have had plenty of opportunities to learn about themselves. Hence, the biases must be stubborn in that they

[4] The assumption that the agent does not care about future actions implies that decisions in different periods are linked only through beliefs about the fundamental — and not through strategic interdependence — allowing us to focus on the novel, learning implications of our theory. The same assumption also implies that the agent does not use his action to influence the beliefs of future selves, as in the literature on self-signaling. While self-signaling is important in situations with uncertainty and limited memory, it is orthogonal to the issues we study. Indeed, our main focus is on degenerate limiting beliefs about θ, where self-signaling is irrelevant even if the agent does care about future choices.

[5] A simple sufficient condition is that the functions v and $\tilde{v}$ are strictly supermodular (or submodular) in (a, s) for every θ and the maximizer is always interior.

[6] Another direct implication of invertibility is that if storing the s_t in memory has any (arbitrarily small) effort cost, then the agent perceives it as optimal not to store the s_t. This is because he believes that he can retrieve the s_t based on his memory of his actions a_t.

are not eliminated, or only very slowly eliminated, by learning. Our assumption of deterministic incorrect beliefs provides a tractable way to study the implications of these stubborn biases.[7]

The central feature above, that the agent uses his past actions a_t as evidence of the reasons s_t behind his actions, is extremely natural from many perspectives in psychology as well as economics. In psychology, the hypothesis that a person consciously or subconsciously "concocts the beliefs and desires" consistent with a past action forms the basis for the large literature on rationalization (Cushman, 2020). In behavioral economics, the idea that people do not fully understand their preferences underlies the possibility of various forms of naivete, including naive present bias and projection bias; and the idea that they can learn from their own behavior underlies arguments that they may become sophisticated over time. Similarly, the notion that a person uses past behavior as a guide to understand himself is a basic assumption in various papers on self-signaling (e.g., Bodner and Prelec, 1996, Prelec and Bodner, 2003, Grossman and van der Weele, 2017) and memory (e.g., Bénabou and Tirole, 2002, 2011, Bernheim and Thomadsen, 2005).[8]

Self-Observation Equilibrium We now define our notion of a stable belief about the fundamental Θ. For our definition, we denote the probability density function of the shocks ϵ_t by f.

Definition 1. A fundamental $\tilde{\theta}$ is a *self-observation equilibrium* (SOE) given the true Θ if

$$\tilde{\theta} = \arg\max_{z} \int \log f\left(\tilde{\pi}_{\tilde{\theta}}^{-1}\left(\pi_{\tilde{\theta}}(\Theta+\epsilon)\right) - z\right) f(\epsilon) d\epsilon\,. \tag{1}$$

Roughly, an SOE is a fundamental such that if the agent both chooses his actions and later interprets the reasons behind his actions based on a point belief on this fundamental, then he has no inducement to change his belief. To understand the definition more precisely, consider what happens as a function of the realized error ϵ if the agent believes that the fundamental is $\tilde{\theta}$. Given the true fundamental Θ, he observes the shock $s_t = \Theta + \epsilon$, and hence takes the action $a_t = \pi_{\tilde{\theta}}(\Theta + \epsilon)$. But he believes that he is using the policy function $\tilde{\pi}_{\tilde{\theta}}$, so when he remembers

[7] Nevertheless, the assumption of deterministic beliefs is not literally accurate because individuals do update these beliefs. In fact, one mechanism through which optimistic beliefs might arise and persist is asymmetric updating (e.g., Sharot and Garrett, 2016, Zimmermann, 2020, Drobner, 2022).

[8] Of course, it is also possible that a person forgets what he did in the past, not only why he did it. Consistent with at least partial forgetting, Carrera et al. (2022) find that providing information about past gym attendance, and making this information salient, changes individuals' expectations about future behavior. Our model applies only to the extent that individuals recall their past behavior, or are reminded of it, non-selectively. Selective forgetting introduces a bias that is quite different from the one studied in our paper (Fudenberg et al., forthcoming).

his action, he infers that the shock must have been $\tilde{s}_t = \tilde{\pi}_{\tilde{\theta}}^{-1}(a_t) = \tilde{\pi}_{\tilde{\theta}}^{-1}\left(\pi_{\tilde{\theta}}(\Theta+\epsilon)\right)$. Over time, he accumulates a distribution of such $\tilde{s}_t$ due to different realizations of ϵ. Now we can think of the agent as performing a simple sanity check by asking himself: is $\tilde{\theta}$ the best explanation for the signals $\tilde{s}_t$ I have observed? Capturing such a requirement, Equation (1) says that $\tilde{\theta}$ maximizes, over all possible fundamentals z, the expected (log) likelihood of the signals $\tilde{s}_t$ that the agent infers from his past actions. If the equation is satisfied, $\tilde{\theta}$ passes the sanity check, and the agent has no inducement to change his belief.

In most of our applications, we have chosen functional forms so that an SOE's required consistency between beliefs and behavior takes an extreme form:

Observation 1. The belief $\tilde{\theta}$ is an SOE if for all $\epsilon \in \mathbb{R}$ we have

$$\pi_{\tilde{\theta}}(\Theta+\epsilon) = \tilde{\pi}_{\tilde{\theta}}(\tilde{\theta}+\epsilon). \tag{2}$$

Since in this case $\tilde{\pi}_{\tilde{\theta}}^{-1}\left(\pi_{\tilde{\theta}}(\Theta+\epsilon)\right) = \tilde{\theta}+\epsilon$, $\tilde{\theta}$ is a perfect explanation for the distribution of $\tilde{s}_t$, so — passing the above sanity check with flying colors — $\tilde{\theta}$ is an SOE.[9] Intuitively, Condition (2) says that the action the agent thinks he chooses coincides with the action he actually chooses for any realized ϵ. This implies that the distribution of actions the agent expects coincides with his actual distribution of actions, so despite his incorrect self-view, he perfectly accounts for his behavior.[10] Although Condition (2) is demanding, it is especially easy to confirm in specific settings, so it is useful for understanding the logic of our model.

In Appendix A, we introduce notation for studying the agent's dynamic learning and optimization problem, and lay out further, technical assumptions to make it tractable. We also identify conditions — satisfied in all our main applications — under which the agent's beliefs converge to an SOE. This implies that to understand long-run behavior, it is sufficient to analyze SOE's. Since convergence is typically difficult to establish in models of misspecified learning, we view these

[9]This result is an implication of Gibbs' inequality, which shows that the negative of the entropy (which up to a constant equals the expected log-likelihood) is always maximized at the true data generating process. In our model, this can be established by verifying the first-order condition $\frac{\partial}{\partial z}\int \log f\left(\tilde{\pi}_{\tilde{\theta}}^{-1}\left(\pi_{\tilde{\theta}}(\Theta+\epsilon)\right)-z\right) f(\epsilon)d\epsilon|_{z=\tilde{\theta}} = \frac{\partial}{\partial z}\int \log f\left(\tilde{\theta}-z+\epsilon\right) f(\epsilon)d\epsilon|_{z=\tilde{\theta}} = -\int \frac{f'\left(\tilde{\theta}-z+\epsilon\right)}{f\left(\tilde{\theta}-z+\epsilon\right)} f(\epsilon)d\epsilon|_{z=\tilde{\theta}} = -\int f'(\epsilon)\, d\epsilon = 0$ and observing that the expected log-likelihood is strictly concave.

[10] Although the settings are different, the perfect-explanation SOE in Observation 1 is similar in spirit to a self-confirming equilibrium defined, for instance, in Fudenberg and Levine (1993), Ba (2022), and Battigalli et al. (2022). In both cases, the agent's equilibrium observations must be perfectly consistent with his beliefs, and he must act optimally given his beliefs.

formal arguments as a useful contribution of our paper.[11] But they are unnecessary for understanding the gist of our applications, so they can be skipped by readers not interested in the technical contribution.

Unawareness and the Impossibility of Disconfirming Incorrect Self-Perception We briefly discuss an alternative to our assumption that the agent forgets s_t: that he is unaware of s_t in the first place. This variant generates essentially the same behavior, and may yield a situation in which the agent never observes information that contradicts his self-view $\tilde{v}$.

Suppose that the agent learns (and remembers) the function $a \mapsto v(a, s_t, \theta)$ for each realized s_t, but he neither learns the s_t themselves, nor understands the full function v. In the case of harmful consumption, for example, he may feel an inclination to consume, but not have direct access to the precise reasons for his urge; and in the case of racial bias, he may feel that he should hire the majority applicant, but not have direct access to the precise reasons behind his intuition. As before, the agent believes that his utility function is $\tilde{v}$. These possibilities are consistent with a long line of research in psychology arguing that people do not understand their mental processes, and interpret them using a-priori fixed theories (e.g., Nisbett and Wilson, 1977). To make behavior simple to define, we impose that (as in most of our applications) for any s_t there is a $\tilde{s}_t$ such that $v(a_t, s_t, \theta) = \tilde{v}(a_t, \tilde{s}_t, \theta)$ for all a_t and θ. Then, there is never — not even at the moment of choice — an explicit contradiction between v and $\tilde{v}$, so it is especially plausible that the agent does not question his belief $\tilde{v}$. For the same reason, the implications remain unchanged even if the agent remembers his realized utility.

Under the above variant, the definition of and motivation for SOE, and hence the analysis of steady-state behavior, remain unchanged. At the same time, for non-degenerate beliefs the appropriate specification of behavior is slightly different from that in our main model. Since the agent does not observe s_t in period t, he must infer it from his feeling. Hence, before maximizing the expectation of v, he updates his belief using the $\tilde{s}_t$ defined above, not using s_t.

[11] Existing convergence results to Berk Nash Equilibrium (e.g., Esponda and Pouzo, 2016, Fudenberg et al., 2017, Bohren and Hauser, 2019, Frick et al., 2021a) do not apply to our setting as there is no direct way of rephrasing the question of convergence under misspecified and incomplete memory to an SOE into the question of convergence under misspecification about the signal structure and perfect memory. We discuss the relationship between our convergence proof and the literature following Proposition 8.

2.2 Equi-Directional Problems

For much of the paper, we restrict attention to an important class of basic environments:

Definition 2. A problem is equi-directional if v and $\tilde{v}$ are either both supermodular or both submodular in (a, θ) and (a, s).[12]

In equi-directional problems, increases in the signal s_t and the fundamental Θ change the optimal action in the same direction. Equivalently, an increase in s_t changes the optimal action in the current period and — through its effect on beliefs about Θ — in future periods in the same direction. This assumption is natural in settings where uncertainty pertains to the benefit or harm of decisions that are not (strong) substitutes, i.e., the main question is how much to consume. For instance, if the agent learns that exercise is beneficial, then he should exercise more, and do so in all periods. On the other hand, equi-directionality rules out situations where uncertainty pertains to both sides of a tradeoff, e.g., the main question is when to consume. We discuss non-equi-directional problems briefly in Section 5.

To identify a general property of equi-directional problems, we define:

Definition 3. The agent's self-view is self-defeating if for any state Θ, any corresponding SOE $\tilde{\theta}$, and any signal s,

$$\tilde{v}(\pi_{\tilde{\theta}}(s), s, \Theta) < \tilde{v}(\pi_{\Theta}(s), s, \Theta). \tag{3}$$

Inequality (3) compares the welfare of two agents according to the wished-for preferences $\tilde{v}$. One agent knows that he acts according to v, so being correctly specified, he learns Θ and takes the action $\pi_{\Theta}(s)$ for each s.[13] The other agent falsely thinks that he acts according to $\tilde{v}$, so he comes to believe that the fundamental is $\tilde{\theta}$ and takes the action $\pi_{\tilde{\theta}}(s)$ for each s. The inequality says that the latter agent always acts less in accordance with his wished-for preference $\tilde{v}$ than the former agent. Note that by the definition of π, the weak version of Inequality (3) trivially holds when $\tilde{v}$ is replaced by v (i.e., $v(\pi_{\tilde{\theta}}(s), s, \Theta) \leq v(\pi_{\Theta}(s), s, \Theta)$). Hence, a self-defeating self-view makes the agent worse off according to both v and $\tilde{v}$.

Proposition 1. *For any equi-directional problem in which v and $\tilde{v}$ imply different optimal actions for all s and θ, the agent's self-view is self-defeating.*

[12] Formally, sgn v_{as} = sgn $v_{a\theta}$ = sgn $\tilde{v}_{as}$ = sgn $\tilde{v}_{a\theta}$ for all a, θ, s, and these cross-derivatives do not change sign.

[13] Observation 2 in Appendix A establishes that an agent who is correctly specified ($v = \tilde{v}$) almost surely learns the state from his own actions, even if he does not remember the signals he observed.

Proposition 1 says that within the equi-directional class of environments, self-defeating learning occurs in any choice situation and for any type of decision-relevant naivete the agent may have. Hence, for instance, an incorrect self-view regarding intertemporal choice is self-defeating if the agent wants to be more patient than he is, and also if he wants to be less patient than he is. Intuitively, given that the agent does something different from what he expects, he must come to believe in reasons for exactly that kind of different behavior. This pushes his behavior further in the same unwanted direction.[14]

3 Present Bias

3.1 Setup

The agent's period-t incarnation, self t, chooses consumption $a_t \in \mathbb{R}$. His decision utility is

$$v(a_t, s_t, \Theta) = u(a_t) + \phi_t a_t - \beta\kappa a_t\,,$$

where $u : \mathbb{R} \to \mathbb{R}$ is a strictly concave function satisfying $\lim_{a\to-\infty} u'(a) = \infty$ and $\lim_{a\to\infty} u'(a) = -\infty$, $\phi_t = l\Theta + (1-l)s_t$ for a constant $l \in (0,1)$, $\kappa \in \mathbb{R}$, and $\beta \in (0,1]$. At other times, the agent thinks that self t's utility function is

$$\tilde{v}(a_t, s_t, \Theta) = u(a_t) + \phi_t a_t - \tilde{\beta}\kappa a_t,$$

where $\tilde{\beta} \in [\beta, 1]$. Just before period t, the agent would prefer to set $\beta = 1$, i.e., to maximize $u(a_t) + \phi_t a_t - \kappa a_t$. Our focus will be on the case $\beta < 1$ and $\tilde{\beta} > \beta$.

The above is a standard formulation of partially naive present bias for a situation where $u(a_t) + \phi_t a_t$ is the current utility from consumption and κa_t is the future harm. The agent discounts future consequences by the factor β (Laibson, 1997, O'Donoghue and Rabin, 1999), but believes that he is using the discount factor $\tilde{\beta}$ (O'Donoghue and Rabin, 2001). For $\beta < 1$, he is present-biased in that

[14] In an earlier paper, Heidhues et al. (2018), we also investigate conditions under which a misspecification about oneself is self-defeating. But there, the nature of the problem is completely different: the agent is misspecified about his production function rather than his preferences, he learns based on observing output rather than actions, he learns about the favorability of the state rather than directly about the effect of his actions, he has perfect rather than imperfect memory, and he has a single objective function. Consequently, the condition for his misspecification to be self-defeating is different, and arguably less economically interpretable than in the current model.

he discounts more than he wishes. And for $\tilde{\beta} > \beta$, he is partially naive in that he overoptimistically thinks that he is less present-biased — or closer to his wished-for preference — than he actually is.

But beyond adopting the core structure of existing present-bias models, our formulation also incorporates the natural idea that instantaneous marginal utility is subject to uncertainty and shocks.[15] A simple interpretation is that marginal utility depends on permanent (Θ) and temporary (s_t) factors. In a decision of whether to take a payday loan, for instance, s_t could represent expenses that have come up so far this month, and Θ could represent the expectation of further expenses that will come up in the near future according to long-run trends. Alternatively, our specification captures, in reduced form, signal-extraction situations in which the agent does not perfectly observe shocks to his utility. When deciding whether to go to a party, for instance, the agent may combine his prior about the types of parties he is invited to with a signal about the specific party.[16]

3.2 Self-Defeating Naivete

We begin by demonstrating that naivete exacerbates present-biased behavior and is therefore self-defeating. To do so, we look for an SOE that satisfies Condition (2) in Observation 1. If the agent believes that the fundamental is $\tilde{\theta}$, then he chooses a_t satisfying the first-order condition

$$-u'(a_t) = l\tilde{\theta} + (1-l)s_t - \beta\kappa = l\tilde{\theta} + (1-l)(\Theta + \epsilon_t) - \beta\kappa. \tag{4}$$

In contrast, he believes that if he receives the signal $\tilde{s}_t$, he chooses $\tilde{a}_t$ to satisfy

$$-u'(\tilde{a}_t) = l\tilde{\theta} + (1-l)\tilde{s}_t - \tilde{\beta}\kappa. \tag{5}$$

Furthermore, he believes that $\tilde{s}_t = \tilde{\theta} + \epsilon_t$, so he believes the above equals

$$-u'(a_t) = l\tilde{\theta} + (1-l)(\tilde{\theta} + \epsilon_t) - \tilde{\beta}\kappa\,. \tag{6}$$

[15] Fudenberg and Levine (2006) allow for utility shocks in the context of a dual-self model, but they do not incorporate misspecification or learning.

[16] To formalize such a signal-extraction situation, suppose that d_t is the period-t benefit of consumption, the agent is attempting to maximize the expectation of $u(a_t) + d_t a_t - \beta\kappa a_t$, and he thinks he is attempting to maximize the expectation of $u(a_t) + d_t a_t - \tilde{\beta}\kappa a_t$. We let $d_t = \Theta + \epsilon_{d,t}$, where the $\epsilon_{d,t}$ are iid mean-zero normal random variables with variance σ_d^2, and $s_t = d_t + \epsilon_{s,t}$, where the $\epsilon_{s,t}$ are iid mean-zero normal random variables with variance σ_s^2. Then, defining $l = \sigma_d^{-2}/(\sigma_d^{-2} + \sigma_s^{-2})$, we get $\mathbb{E}[d_t|s_t, \Theta] = l\Theta + (1-l)s_t = \phi_t$, reducing the problem to the above formulation.

Equating the right-hand sides of (4) and (6), we obtain an SOE belief $\tilde{\theta}$ at which the agent perfectly predicts his behavior.

Proposition 2. *The agent's beliefs converge a.s. to the unique SOE*

$$\tilde{\theta} = \Theta + \frac{\tilde{\beta} - \beta}{1 - l}\kappa. \tag{7}$$

The SOE satisfies Condition (2)*, allowing the agent to perfectly predict his behavior. At the SOE, the agent chooses consumption satisfying*

$$\frac{\partial v(a_t, s_t, \Theta)}{\partial a_t} = u'(a_t) + l\Theta + (1 - l)s_t - \beta\kappa = -\frac{l(\tilde{\beta} - \beta)}{1 - l}\kappa\,. \tag{8}$$

Proposition 2 implies that for situations such as harmful consumption and spending ($\kappa > 0$), the agent ends up overestimating the average marginal utility from consumption ($\tilde{\theta} > \Theta$). Intuitively, a most basic implication of partial naivete is that the agent starts out underestimating his average consumption. To explain the — to him surprisingly high — consumption levels he actually chooses, he revises his beliefs regarding the average marginal utility of consumption upwards. In addition, the more naive he is or the greater is the harm (i.e., the higher is $\tilde{\beta}$ or κ), the more he mispredicts his behavior initially, so the more he revises his beliefs.

These revisions are detrimental for his behavior. By Equation (8), the agent's long-run consumption is higher than it would be if he was correctly calibrated ($\tilde{\beta} = \beta$), it is higher than optimal from the perspective of the decisionmaking self's actual preferences v, and it is increasing in his naivete. Given his pre-existing tendency to overconsume — optimal consumption from the perspective of v is already higher than what the agent wishes just before period t — his mislearning therefore exacerbates his suboptimal behavior. This is a manifestation of the self-defeating learning we have identified in Proposition 1. As an example, suppose that in each period the agent chooses how many parties to go to and how long to stay, and he views parties as fun in the present but costly for the future. He starts off going to more parties than he expects, so he develops the self-view that he enjoys parties. As a result, he goes to too many parties, and stays too long, even from the perspective of his short-run, party-going self.

Reflecting the general message of Proposition 1, Proposition 2 implies that the agent's inferences are self-defeating not only for harmful, but also for beneficial activities. If $\kappa < 0$, then the agent underestimates the instantaneous utility, i.e., he overestimates the instantaenous disutility, of consumption. He may, for instance, develop the belief that exercise has large personal costs. This in turn exacerbates his tendency to exercise too little.

The above self-defeating inferences reflect a novel type of harm associated with naive present bias. Whereas in previous models (e.g., O'Donoghue and Rabin, 1999, Gruber and Kőszegi, 2001, Heidhues and Kőszegi, 2010, Ericson and Laibson, 2019) naivete affects current behavior due to mispredictions regarding future behavior, in ours it does so due to mispredictions regarding the instantaneous utility function. As a consequence, in our setting the agent acts too impatiently from the perspective of his present-biased preferences v despite perfectly predicting his future behavior. Furthermore, while the previously known effect of mispredicting future behavior on current behavior is often small and can also be beneficial, in our equi-directional settings naivete is always detrimental. Hence, in real-life settings our model identifies an arguably more robustly harmful effect of naivete than do previous models.

3.3 Beliefs About Future Harm and Offsetting Benefits

In the above model of present-biased consumption, the uncertainty in the agent's utility function pertains to current marginal utility. We now consider the obvious alternative that the uncertainty is about the future impact of consumption. For harmful products, the agent's true and perceived utility functions are $v(a_t, s_t, \Theta) = u(a_t) - \beta e^{\phi_t} a_t$ and $\tilde{v}(a_t, s_t, \Theta) = u(a_t) - \tilde{\beta} e^{\phi_t} a_t$, respectively, where $e^{\phi_t} a_t$ is now the harm from consumption, and as before $\phi_t = l\Theta + (1-l)s_t$ for some $l \in (0,1)$. In a consumption-savings setting, for example, ϕ_t could be a measure of the likelihood or seriousness of a future contingency, such as a large expense or a period of unemployment, for which it is harmful to be unprepared. Similarly, ϕ_t could measure the impact of smoking on future utility, which combines negative health consequences as well as other effects, e.g., the benefit of concentrating on work. For beneficial products, we define $v(a_t, s_t, \Theta) = u(a_t) + \beta e^{\phi_t} a_t$ and $\tilde{v}(a_t, s_t, \Theta) = u(a_t) + \tilde{\beta} e^{\phi_t} a_t$. In the context of exercise, for instance, e^{ϕ_t} could capture the future benefit of exercising now, which may include health benefits as well as the value of feeling or looking better.

Proposition 3. *For both harmful and beneficial products, the agent's beliefs converge a.s. to the unique SOE*

$$\tilde{\theta} = \Theta - \frac{\ln \tilde{\beta} - \ln \beta}{1-l}.$$

The SOE satisfies Condition (2)*, allowing the agent to perfectly predict his behavior.*

Proposition 3 implies that the agent's long-run belief about the future impact of increasing current consumption is biased toward zero. The agent may come to believe, for instance, that contingencies for which he might need to save rarely happen. Similarly, he may believe that smoking is not as harmful for him as for others, or that it has a benefit for concentration that offsets its negative health consequences. This underestimation of future consequences exacerbates the agent's underweighting of the same consequences due to β, so for any $l > 0$ and both harmful and beneficial products, naivete is again self-defeating. And again, suboptimal behavior occurs despite the agent perfectly predicting his behavior.

Despite a semblance of similarity, the agent's long-run bias about the future is different from an optimistic bias due to anticipatory utility as modeled for instance by Brunnermeier and Parker (2005) and documented for instance by Oster et al. (2013). Whereas the predictions of anticipatory utility center on beliefs about the *level* of future utility, our predictions center on beliefs about the *effect of consumption* on the level of future utility. To see the contrast especially clearly, consider the implications of adding a term $h(\Theta, s_t)$ to both utility functions v and $\tilde{v}$. This shifts the level of utility under different fundamentals, and hence which beliefs are more or less optimistic. But it leaves the agent's real or perceived decisionmaking problem, and hence his behavior and his inferences from it, unchanged. In particular, suppose that $h(\theta, s_t) = -\tilde{v}(\tilde{\pi}_{\tilde{\theta}}(s_t), s_t, \theta)$, where $\tilde{\theta}$ is the SOE belief corresponding to the true fundamental Θ. Then, the agent perceives his utility from his SOE action to be fixed, so he does not update about this utility. In the context of smoking, for instance, he may not update his view that smoking is harmful. To explain his behavior, then, he comes to believe that stopping to smoke will also entail harmful consequences. Anticipatory utility does not predict such sad beliefs.

3.4 Behavior and Beliefs that Reveal Partial Naivete

The previous sections have shown that in a stable environment, a partially naive present-biased agent can learn to account for and predict his behavior well — potentially perfectly — so in this first-pass sense he is realistic. We now show that upon closer inspection, his behavior and beliefs exhibit patterns that are distinct from those of a realistic agent.

Misprediction: Change in Future Incentives Consider the basic setup of Section 3.1, with the agent having converged to SOE beliefs. Suppose that just before period t, the agent learns about a surprise (temporary or permanent) change in the future harm of consumption κ. This does not affect his belief about the fundamental. Hence, from Equations (6) and (4), the responsivenesses of perceived and true consumption to the news equal

$$\partial \tilde{a}_t/\partial \kappa = \tilde{\beta}/u''(\tilde{a}_t) \qquad \text{and} \qquad \partial a_t/\partial \kappa = \beta/u''(a_t), \tag{9}$$

respectively. Because $\tilde{a}_t = a_t$ for any ϵ_t, the former is greater than the latter at any point in the consumption distribution. Intuitively, despite predicting his behavior, the agent continues to overestimate how much weight he will put on the future, so he continues to overestimate how sensitive he will be to incentives stemming from a change in future consequences.[17]

Misprediction: Change in Current Incentives Suppose, slightly extending our model, that current utility is $u(a_t) + \phi_t \kappa' a_t$ for a constant $\kappa' > 0$, and the future harm is still κa_t. For instance, ϕ_t might denote the effectiveness of smoking in reducing stress, and κ' the importance of reducing stress in the agent's life. Then, because the agent overestimates Θ and therefore also ϕ_t, he overestimates his response to a change in κ'.[18]

As an example, consider an employee who has been working on a high-pressure project, and who smokes a lot. We can think of this as a situation with a combination of a high κ' and a relatively low κ. A high κ' captures that the benefit of stress reduction is high. A low κ captures that there is an incentive to perform, so the benefit of concentration on one's work is high, offsetting the negative health effects of smoking. According to our model, the agent correctly forecasts that he will continue to smoke a lot while the project lasts. At the same time, he overestimates how much he will cut back once the project comes to an end, i.e., when κ' decreases and κ increases. Existing models of partial naivete, in contrast, do not imply such a connection between overoptimism about cutting back and improvements in circumstances. In those models, the agent thinks that even in an unchanged environment, from the next period on he will be better-behaved.

[17] In Equation (9), the ratio of the agent's perceived to his true reaction is exactly $\tilde{\beta}/\beta$, so that his misprediction accurately reveals his naivete. This is not a general prediction of our model. It is easy to construct examples in which the agent's misprediction is more severe or less severe.

[18] Precisely, at the median consumption level, the responsivenesses of the agent's perceived and true consumption equal $\partial \tilde{a}_t/\partial \kappa' = -\tilde{\theta}/u''(a_t)$ and $\partial a_t/\partial \kappa' = -(l\tilde{\theta} + (1-l)\Theta)/u''(a_t)$, respectively.

Misprediction: New Environment Suppose that the agent is placed in a new, ex-ante identical environment in which another fundamental Θ is drawn independently. Then, he starts off mispredicting his behavior just like previously, as if he had not learned anything. In this sense, he learns to understand himself in a fixed environment, but his understanding does not transfer to other environments. He might repeatedly learn, for instance, that his cravings under the current diet are too strong to carry through with the regimen, but perpetually believe that the next type of diet will work. In contrast, a correctly specified agent should learn his present-bias parameter β and transfer this knowledge to other settings.

Patterns in Beliefs Our model also predicts patterns in beliefs that would seem puzzling from the perspective of a correctly specified model. At a basic level, the agent's belief about the benefit of consuming a harmful product ($\tilde{\theta}$ in Equation (7)) responds positively to changes in the future cost of consumption (κ), although there is no logical relationship between the two parameters. More subtly, beliefs about logically unrelated issues can exhibit systematic correlations. We have shown that a partially naive agent might come to overestimate the benefit or underestimate the relative harm of consuming a harmful product. Simple extensions of our methods imply that he can hold both biases at the same time (see Appendix D). He might, for example, believe both that smoking is beneficial for concentration and that other activities are harmful. Hence, across a population of agents with heterogeneity in naivete (say, due to heterogeneity in $\tilde{\beta}$ given β), such different types of beliefs will be correlated. Evidence by Oakes et al. (2004) and Fotuhi et al. (2013) provides support for this prediction.[19]

3.5 Apparent Sophistication

In this subsection, we identify implications of our theory for the empirical literature on sophistication in intertemporal choice. The models above raise immediate doubt about the general approach in the literature, which assesses a person's level of sophistication by his ability to predict his own behavior. At the SOE, our agent passes this prediction test — he predicts his behavior perfectly — yet he is only *apparently sophisticated*: he neither understands himself, nor acts optimally given

[19] Smokers are more likely than ex-smokers to think that smoking is joyful or helps with relaxation, socialization, or concentration. They also think that the health risks of smoking — which they generally understand and sometimes even overestimate (Viscusi, 1998) — do not apply specifically to them, or other activities are just as risky. Further, these beliefs exhibit a significant positive correlation across individuals.

his time inconsistency. We develop the case that apparent sophistication is not only a theoretical possibility, but also empirically relevant.

Most importantly, we argue that findings on the sophistication of payday-loan borrowers by Allcott et al. (2022) are better explained by apparent than by real sophistication. To capture their setting in our framework, we reinterpret and augment the model in Section 3.1. Each period t is an episode of financial distress for the agent, in anticipation of which he takes out a payday loan. Then, he observes the severity s_t of his situation, and chooses consumption a_t, determining whether he rolls over his loan. This yields instantaneous utility $u(a_t) + \phi_t a_t$, and has a loan-repayment cost of κa_t with $\kappa > 0$ at the end of the distress episode. Furthermore, as part of Allcott et al.'s experiment, the agent makes two choices going into period t (hence, before observing s_t). He (i) reports his subjective expectation of a_t, and (ii) reveals how much he values a marginal uniform decrease of a_t, where the latter value is derived from the goal to maximize the expectation of the undiscounted utility $u(a_t) + \phi_t a_t - \kappa a_t$.[20] We call the agent inexperienced if $t = 1$, and experienced if his beliefs have converged and are thus given by the SOE.

Suppose that an observer (she) has access to mean consumption a_t as well as (i) and (ii) for infinite samples of inexperienced and experienced agents. In reality, all agents have the same u, β, $\tilde{\beta}$, prior distribution of Θ, variance σ_ϵ^2 of s_t, and κ, so there is no sampling error in estimation. The observer knows u and κ, and assumes that the agents also know Θ and σ_ϵ^2. The latter corresponds to the usual assumption that agents know their instantaneous utility functions. The observer is interested in inferring each group's true and perceived discount factors β and $\tilde{\beta}$, allowing for the possibility that these (as well as the distribution of Θ and σ_ϵ^2) differ across groups.[21] If she can find a unique β and $\tilde{\beta}$ that (along with other parameters) perfectly explain her data for a group, then she infers that these are the group's parameters; otherwise, she rejects her model.

Proposition 4. *The observer does not reject her model. For inexperienced agents, she correctly*

[20] Formally, suppose that before observing s_t, the agent expects to choose the action $\tilde{\pi}_{\mu_t}(s_t)$. Then, (i) equals $\mathbb{E}_{\mu_t}[\tilde{\pi}_{\mu_t}(s_t)]$, and (ii) equals $C'(0)$, where $C(\Delta)$ is the agent's willingness to pay to choose $\tilde{\pi}_{\mu_t}(s_t) - \Delta$ instead of $\tilde{\pi}_{\mu_t}(s_t)$. Empirical methods for eliciting (i) and (ii) are detailed in Allcott et al. (2022). The more tricky valuation (ii) can be elicited by measuring the agent's willingness to pay for incentives to decrease consumption.

Note also that the predicted mean consumption and value of a marginal decrease in consumption at time t only depend on the "public" belief μ_t at the beginning of the period, which is common knowledge among all later selves. Thus, even if these additional "actions" are remembered by the agent, they will not influence his beliefs.

[21] In practice, the observer would have to infer the shape of u from behavior. This is irrelevant for our analysis, so we ignore it. Further, since we can normalize the marginal utility of money in the future to be 1, in the payday setting we can assume that the observer knows κ. In other settings where the future impact of current choices is not financial (e.g., the benefit of exercise), the money equivalent of κ would also have to be estimated.

estimates the true parameters β, $\tilde{\beta}$. For experienced agents, she correctly estimates the true $\tilde{\beta}$, but incorrectly infers that the agent is sophisticated, i.e., that $\beta = \tilde{\beta}$.

When the agent is inexperienced, the observer understands him correctly. But when he is experienced, she thinks of him exactly what he thinks of himself: that he is sophisticated with a present-bias parameter of $\tilde{\beta}$. The logic for the second result is in two parts. First, the observer can infer $\tilde{\beta}$ from the agent's willingness to pay to decrease consumption a_t. Since this depends only on the agent's belief about how he will treat the future, and hence not on his belief about the fundamental, the inference is correct. Second, the observer can infer the agent's degree of naivete from his bias in predicting his own behavior. Since he predicts his behavior perfectly, the observer concludes that he is sophisticated, so that his present-bias parameter must be $\tilde{\beta}$.

A model of correct learning about present bias and our model of misspecified learning therefore have distinctly different predictions as to what happens as learning unfolds. In the former case, the agent's belief adjusts downwards to his true present bias. In the latter case, in contrast, the agent's measured level of present bias adjusts upward to his belief. Supporting our model, the central structural finding in Allcott et al. (2022, Table 3) is much closer to the latter prediction than to the former one: inexperienced borrowers' measured β and $\tilde{\beta}$ are 0.73 and 0.92, respectively, whereas for experienced borrowers both measures equal 0.86.[22]

Consider also the finding of Le Yaouanq and Schwardmann (2022) that subjects in a laboratory experiment quickly become better at predicting their own willingness to work. The authors find strong evidence that subjects learn about the cost of effort, and weaker evidence that they learn about present bias. This is consistent with an explanation Allcott et al. propose for their own findings within the realm of correctly specified learning: that subjects were sophisticated all along, and the reduced misprediction of future behavior is due to learning about the utility cost of repayment. But analogously to the payday example, it is entirely possible that the combination of facts is at least in part due to our explanation: that subjects *remained partially naive* all along, and they *mislearned* about their cost of effort.[23]

In addition, further patterns are inconsistent with models of correctly specified learning and naturally predicted by ours. First, Allcott et al. (2022) report that much like their inexperienced

[22] These are estimates for risk-neutral borrowers, which according to the paper's definition our agent is.

[23] Our model, however, does not explain another puzzling fact in Le Yaouanq and Schwardmann (2022), that subjects underestimate their future change in beliefs after taking an unexpected action.

counterparts, experienced borrowers overestimate their response to a monetary incentive to reduce borrowing in the future. Since a new incentive is equivalent to an increase in the harm κ of consumption, our model predicts exactly such an overestimation (see Expressions (9)).

Second, our model's prediction regarding the non-transferability of learning across environments (Section 3.4) helps interpret the literature on sophistication as a whole. In particular, this prediction reconciles the above theme that in some isolated settings individuals' predictions about their own behavior improve quickly (see also Allcott et al., 2021, Carrera et al., 2022) with another major theme: that in a variety of other, especially unfamiliar situations many individuals are quite naive (Skiba and Tobacman, 2008, Acland and Levy, 2015, Fang and Wang, 2015, Fedyk, 2021, Augenblick and Rabin, 2019, Chaloupka et al., 2019, Carrera et al., 2022, John, 2020, Bai et al., forthcoming, Kuchler and Pagel, 2021). In reconciling these findings, our model also says that the former ones may be due to apparent sophistication, making the latter ones more informative about individuals' naivete. From the perspective of a correctly specified model, in contrast, the two themes are arguably contradictory. Based on such conventional logic, if a person quickly learns about his present bias in many situations, then he should quickly develop an accurate understanding of his average present bias, so by adulthood he should not be systematically biased in new situations.[24]

Importantly, the observer's misunderstanding of an experienced agent in Proposition 4 leads her to overly optimistic conclusions about multiple aspects of his behavior. She thinks that his level of a_t (in the payday example, his loan renewal) reflects sophisticated present bias with a parameter of $\tilde{\beta}$. In contrast, self-defeating naivete means that a_t is actually higher than that implied by sophisticated present bias with a parameter of β. The observer also thinks that the agent's willingness to pay to lower a_t (i.e., to repay the loan early) is correct. In contrast, it remains too low. Finally, the observer would presumably guess that an earlier consumption decision (e.g., a decision of whether to take a payday loan in the first place) also reflects a present bias of $\tilde{\beta}$. In Appendix B, we extend our model to allow for such a decision, and show that here too the agent acts more impatiently.

[24] This is true even if present bias varies from situation to situation. It may be the case, for instance, that a person's β is drawn from a distribution independently for each situation. Then, the type of quick learner suggested by the above experiments should arguably learn about the distribution, and come to an accurate understanding of his average behavior.

3.6 Learning Process and Addiction-Like Behavior

Having analyzed properties of the agent's long-run behavior, we now study his learning process and behavior away from the limit. Although we assume that consumption is not addictive — it does not feature physiological or other history dependence of instantaneous utility — our model generates some consumption patterns that resemble those in models of addiction. Furthermore, our theory makes some distinct predictions that complement existing modeling approaches.

We take our model of harmful consumption from Section 3.1 (where $\kappa > 0$). To keep our analysis tractable, we assume, like much of the literature on addiction starting from Becker and Murphy (1988), that utility is quadratic: $u(a) = -a^2/2$. We denote the consumption level the agent chooses by $a_t = \arg\max_a \mathbb{E}_{\mu_t}[v(a, s_t, \theta)|s_t]$, and in slight abuse of notation write $a_t(s_t, a^{t-1})$ for the action as a function of past actions $a^{t-1} = (a_1, \ldots, a_{t-1})$ and the current signal s_t. Analogously, we denote the optimal consumption level from the perspective of self t's utility function v by $a_t^* = \arg\max_a \mathbb{E}\left[v(a, s_t, \theta)\middle|s_t, a^{t-1}\right]$. Proposition 5 characterizes the agent's behavior:

Proposition 5. *Suppose that $\tilde{\beta} > \beta$.*

I. Given any signal s_t,

$$\frac{\partial a_t(s_t, a^{t-1})}{\partial a_{t-1}} > \frac{\partial a_t(s_t, a^{t-1})}{\partial a_{t-2}} > \cdots > \frac{a_t(s_t, a^{t-1})}{\partial a_1} > 0\,.$$

II. For any $t \geq 2$, self t overestimates the state ($\mathbb{E}_{\mu_t}[\Theta] > \mathbb{E}\left[\Theta\middle|s^t\right]$), and overconsumes ($a_t > a_t^*$). Furthermore, his overconsumption $a_t - a_t^*$ is strictly increasing in κ.

III. The agent's ex-ante expected consumption $\mathbb{E}_{\mu_1}[a_t]$ is strictly increasing in t, with $\mathbb{E}_{\mu_1}[a_t] - \mathbb{E}_{\mu_1}[a_{t-1}]$ strictly increasing in κ. His ex-ante belief regarding the expected a_t is constant in t.

Part I says that just like in the intertemporal-complementarities approach to addiction (e.g., Becker and Murphy, 1988, Orphanides and Zervos, 1995, Gruber and Kőszegi, 2001), the agent's current consumption is increasing in his past consumption. In previous models, this "backward-looking intertemporal complementarity" occurs because higher past consumption raises the marginal utility of current consumption. In our model, it occurs because higher past consumption indicates to the agent that his marginal utility had been higher in the past, raising his *belief* about the marginal utility of current consumption. Since future circumstances are not informative about current marginal utility, however, our theory does not generate the prediction — tightly connected under the intertemporal-complementarities approach — that consumption in a period is sensitive to prices in *future* periods. Consistent with our model, backward-looking intertemporal complementarity is widely documented, but the latter "forward-looking intertemporal complementarity" has received only mixed support.[25]

While Part I also holds for a correctly specified agent, Parts II and III only hold for a misspecified agent. Part II generalizes what we have seen in Proposition 2 for long-run behavior: in every period $t \geq 2$, the agent overestimates his marginal utility of consumption and therefore overconsumes. This property is consistent with clinical definitions of harmful addictions, which suggest that addicts' high consumption is not worth the cost. It is also similar to the mistaken overconsumption in the cue-triggered consumption model of Bernheim and Rangel (2004). In Bernheim and Rangel's model, however, mistaken overconsumption is exogenously assumed rather than derived.

Part III says that the misspecified agent's average consumption is increasing over time, but he does not anticipate this addiction-like process. Increasing consumption also occurs naturally in models of intertemporal complementarities, but the mechanism is again different. In previous models, the agent accumulates a stock of past consumption that increases the marginal utility of consumption — which is a process he understands unless he has another bias (such as projection bias, as in Loewenstein et al., 2003). In our setting, instead, the agent accumulates misinferences, so it is only his beliefs about the marginal utility of consumption that keep rising — which is a process he fails to understand even without any additional biases.

Finally, the last two patterns are increasing in the harm of consumption κ. If the product is more harmful, then the agent mispredicts his behavior by more, so he makes greater mistakes in inferring his instantaneous utility. As his inferential mistakes induce greater consumption, he ends up overconsuming by more, and exhibits a more sharply increasing consumption profile.

Our model therefore says that a person may display patterns of addictive behavior even for harmful products that do not feature intertemporal complementarities in utility. Indeed, researchers have long described the phenomenon of "behavioral addiction" for many activities that are not physiologically addictive, such as gambling, internet use, video gaming, and shopping (see Grant et al., 2010, for an introduction). While models of intertemporal complementarities are not inconsistent

[25] Becker et al. (1994) and Gruber and Kőszegi (2001) document evidence consistent with forward-looking intertemporal complementarity, but Rees-Jones and Rozema (2020) argue that this could be due to other factors, and Liu et al. (1999), Petruzzello (2019), and Allcott et al. (2021) do not find evidence of forward-looking intertemporal complementarity.

with such behavior, nor do they predict why complementarities should arise in these cases. Similarly, models of intertemporal complementarities do not explain why behavioral addictions are less common for beneficial goods.

In addition, our model makes a prediction that is distinct from the predictions of all other models. Suppose that the agent has been making choices under the assumption that the harm from consumption κ would remain unchanged forever. But just before period t, κ suddenly rises, and the agent expects it to remain at the new level forever. This may capture, for example, an instance when new and definitive scientific evidence regarding the harm of smoking becomes public. Then:

Corollary 1. *The agent's short-run (period-t) response to an increase in the harm κ is greater than his long-run (SOE) response.*

The intuition derives from the agent's failure to predict his response to a change in incentives. Since his reaction to the increase in κ is smaller than he expects, he gradually becomes more convinced that consumption is beneficial, so his consumption rebounds. This also implies that informational interventions regarding the harm of consumption are at least partially offset by mislearning. In models of intertemporal complementarities, in contrast, the short-run decrease in consumption lowers marginal utility and hence leads to a greater long-run response.

While we have not found precise evidence on the above prediction, a puzzling combination of findings does seem suggestive. Based on smokers' long-run smoking responses to health information, Ippolito and Ippolito (1984, Table 3) estimate that (depending on the discount rate) smokers value their lives at \$0.46-2.44 million on average; but based on labor-market data, Viscusi and Hersch (2008) report that smokers value their lives at \$7.32 million on average (all numbers in \$2006). The offset effect behind Corollary 1 explains why these numbers are so different: the agent responds to health information by developing beliefs that lower the effect of the information, so he displays too little sensitivity to the information. Since labor-market choices largely trade off future compensation with future risks, there is no mislearning of the same type.[26,27]

[26] Note that present bias is in itself insufficient to explain the lower value of life estimated from smoking decisions. Ippolito and Ippolito (1984) estimate the value from comparing consumers' sensitivities to health information and to prices. To the extent that consumers evaluate a payment as future disutility — a plausible assumption for consumers who are not budget constrained — present bias drops out in such a comparison. If — implausibly — consumers evaluate payments entirely as current utility, the value of life estimated from responses to information should be β times the value estimated from labor-market tradeoffs. The proportion, however, is much lower than existing estimates of β.

[27] Empirically, long-run and short-run responses are often compared by looking at the effects of a period's price

4 Implicit Bias

Contemporary accounts view much discrimination as implicit: the discriminating person endorses egalitarian standards, but acts on hidden negative attitudes (e.g., Dovidio and Gaertner, 2000, Payne et al., 2010, Bertrand and Duflo, 2017). We provide a novel perspective on this type of prejudice that centers around the agent's unawareness of his biased motives.

Suppose that in each period, the agent distributes resources between an individual from the majority and an individual from a minority, choosing the extra amount $a_t \in \mathbb{R}$ to give to the latter. This decision may represent, for instance, labor-market hiring, university admission, or loan approval. The agent's true and perceived utility functions are

$$v(a_t, s_t, \Theta) = (\phi_t - b)a_t - \frac{a_t^2}{2} \qquad \text{and} \qquad \tilde{v}(a_t, s_t, \Theta) = \phi_t a_t - \frac{a_t^2}{2}, \tag{10}$$

where $b > 0$. We define Θ as the average ability, deservingness, or need of the minority relative to the majority, s_t as a signal of the relative deservingness of the particular minority individual in period t, and $\phi_t = l\Theta + (1-l)s_t$ as the correct posterior regarding the relative deservingness of that minority individual.[28]

The utility functions in (10) incorporate three key assumptions. First, the agent is sensitive to deservingness. Second, his true preference v is biased against the minority by the amount $b > 0$. Third, because his perceived preference $\tilde{v}$ treats individuals symmetrically, he is unaware of his bias. These assumptions best describe a subjective decision in which the agent is not aware of the signal s_t he has observed. Instead, he experiences $\phi_t - b$ (or $(1-l)s_t - b$) as a subjective judgment or intuition regarding the particular case at hand. If the agent was employing clear objective criteria, it would be easier for him to identify and keep track of s_t over time, and it might be easier for him to eliminate the influence of the bias b on his decisions.

change over time. Supporting the intertemporal-complementarities model, research based on this approach tends to find that long-run responses are greater than short-run responses. Whether these findings are consistent with our model depends on details of how the agent treats a price change. One possibility is that he does not fully account for past prices when interpreting past consumption (formally, the effect of a price change is treated as part of the shock s_t). Then, by Proposition 5 our model makes the same prediction as the intertemporal-complementarities approach, so it is consistent with the evidence. Another possibility is that the agent fully accounts for past prices when interpreting past consumption. If in addition he evaluates payments as future disutility, then a permanent price change is equivalent to a change in κ. This means that our model makes the opposite prediction than the intertemporal-complementarities approach, and is therefore inconsistent with the empirical findings.

[28] See Section 2 and Footnote 16 for arguments that our model captures such signal-extraction situations.

Proposition 6. *The agent's beliefs converge a.s. to the unique SOE*

$$\tilde{\theta} = \Theta - \frac{b}{1-l}\,. \tag{11}$$

We interpret Proposition 6 by assuming that $\Theta = 0$, i.e., the groups are on average equally deserving and should on average receive the same amounts. Then, the proposition says that the agent adopts the incorrect belief that the minority is less deserving than the majority. As a result, he engages in two types of discrimination against the minority: implicit taste-based discrimination due to b, and "inaccurate" or "misspecified" statistical discrimination (England and Lewin, 1989, Bohren et al., forthcoming, Heidhues et al., 2022) due to his incorrect beliefs. The latter reinforces the former, so the agent acts in an overly prejudiced way not just relative to his egalitarian explicit standards, but also relative his negative implicit attitudes; and his behavior becomes more prejudiced than it would be if he was aware of his negative attitudes. Indeed, the bias in the agent's belief ($b/(1-l)$) is increasing in the bias b in his preferences. Our model therefore predicts that among individuals who consider themselves egalitarian, taste-based discrimination and statistical discrimination are positively correlated.

While we have assumed that b is a taste parameter, there is a simple reinterpretation in which it is an implicit bias in beliefs. For instance, the agent may consciously endorse the view that men and women are equally qualified, while subconsciously believing that women are less able. Then, to account for his behavior, he develops other types of beliefs that further disfavor women. He may, for instance, come to believe that women are less committed to their careers, or come up with criteria for choosing a_t that in the relevant population men are more likely to satisfy (Rudman and Glick, 2001, Uhlmann and Cohen, 2005). In this case, our model predicts a correlation between the agent's implicit bias and his other views.[29]

The self-defeating nature of implicit bias provides a pessimistic interpretation of the transition from explicit to implicit discrimination. Since people endorse egalitarian standards at least on a conscious level, one may interpret the development as a step in the right direction, or at least not a

[29] Formally, suppose that the agent's objective function is $u(a_t, s_t, \theta, \theta')$, where θ and θ' are two relevant fundamentals. For instance, θ could correspond to women's average commitment and θ' to their average ability. Consciously, the agent endorses the view that $\theta' = \Theta'$, so he thinks he is maximizing $\tilde{v}(a_t, s_t, \theta) = u(a_t, s_t, \theta, \Theta')$. But subconsciously, he believes that $\theta' = \tilde{\theta}' < \Theta'$, so he acts according to $v(a_t, s_t, \theta) = u(a_t, s_t, \theta, \tilde{\theta}')$. Then, the agent's implicit negative belief regarding women's ability leads to the view that women are less committed, reinforcing his implicit bias.

step backwards. In our model, however, the agent's inaccurate self-view results in even more biased behavior.

Relatedly, our model reinforces the view that in situations of potential implicit bias, basing decisions on subjective criteria may be harmful. First, our model formalizes the common concern behind this view, that subjective criteria allow implicit biases to infiltrate decisions. Second, our theory adds that repeated subjective decisionmaking not only facilitates prejudices in the short run, but also exacerbates them in the long run. Hence, using solely objective criteria can lead to better decisions, even if in principle subjective judgments contain valuable additional information.

Conversely to the harmfulness of unawareness, our framework says that making the agent aware of his bias (i.e., setting $\tilde{v} = v$) improves his behavior through two potential channels. First, even without changing the motives driving his behavior (v), self-awareness improves behavior by making his beliefs $\tilde{\theta}$ unbiased. Second, less surprisingly and analogously to commitment in the context of present bias, an agent who is aware of but does not agree with his biased motives may want to take steps to consciously counteract them. To do so, he may use internal mechanisms such as paying attention to and analyzing his instinctual choices, or external mechanisms such as publicly committing to objective decision criteria. Consistent with these predictions, some research suggests that awareness-raising interventions do improve behavior (Son Hing et al., 2002, Pope et al., 2018, Boring and Philippe, 2021).

It is noteworthy that the bias in the agent's SOE belief is greater if l is higher, i.e., if the signals s_t about relative deservingness are less accurate. Hence, our model predicts not only that biases are more likely to be translated into behavior when a situation is less clear (e.g., Bertrand and Duflo, 2017), but also that the biases are greater when a person frequently encounters such situations. Intuitively, when l is greater, the agent responds less to his signal s_t, so to explain his surprisingly pro-majority behavior, he must believe that he has observed better reasons to disfavor the minority individual. As a result, his SOE belief must be more biased.

The model in this section is related to those of Ridout (2021) and Cunningham and de Quidt (2022). In both previous models, the decisionmaker has prejudiced preferences or intuitions that he only wants to act upon selectively. In particular, he does not want to act upon them when the only possible reason to do so is prejudice (e.g., because two candidates differ only in gender), but he does want to act upon them when there are more acceptable reasons (e.g., because the candidates have non-comparable qualifications). This context sensitivity generates intransitivity in prejudicial behavior that has been documented for instance by DeSante (2013) and Barron et al. (2020). Without additional assumptions analogous to those in the previous papers, our model does not explain such evidence. At the same time, these previous models do not make predictions regarding the role of awareness in behavior or the effect of implicit biases on beliefs. And most importantly, in these models as well as approaches to cognitive dissonance in economics (e.g., Rabin, 1994), the egalitarian standards the agent wishes to uphold have a beneficial effect in moderating biased behavior. In contrast, our model identifies a reason why such standards may exacerbate biased behavior.

5 Non-Equi-Directional Problems

Our insight above that the agent's misspecification is always self-defeating is derived for equi-directional problems. We conclude the paper by showing that outside this class, misdirected learning can be beneficial as well as harmful.

Suppose that in each period t, the agent divides work over two subperiods of length 1 each. He decides the amount $a_t \in (0,1)$ to do in subperiod 1 (leaving leisure $1-a_t$), and then must complete the work by doing the amount $1-a_t$ in subperiod 2 (leaving leisure a_t). The agent is present biased and partially naive, with true and perceived objective functions

$$v(a_t,s_t,\theta) = e^{s_t}u(1-a_t) + \beta h(\theta)u(a_t) \quad \text{and} \quad \tilde{v}(a_t,s_t,\theta) = e^{s_t}u(1-a_t) + \tilde{\beta} h(\theta)u(a_t).$$

We assume that u is a differentiable, strictly increasing, and strictly concave utility function over leisure satisfying $\lim_{a\searrow 0} u'(a) = \infty$, and h is a strictly increasing positive-valued function. This means that the agent faces a log-normal shock to his utility from leisure in subperiod 1, and has an (expected) value of leisure of $h(\theta)$ in subperiod 2. Note that increases in s_t and θ move the optimal action in opposite directions, so the decisionmaking problem is now not equi-directional.

Proposition 7. *There is a unique SOE belief $\tilde{\theta} = \Theta + \ln(\tilde{\beta}/\beta)$, with which the agent predicts the distribution of his future actions perfectly. Furthermore, his SOE policy function satisfies*

$$\frac{u'(1-\pi_{\tilde{\theta}}(s))}{u'(\pi_{\tilde{\theta}}(s))} = \frac{u'(1-\tilde{\pi}_{\Theta}(s))}{u'(\tilde{\pi}_{\Theta}(s))} \cdot \frac{\beta}{\tilde{\beta}} \cdot \frac{h(\tilde{\theta})}{h(\Theta)}. \tag{12}$$

To understand the proposition, notice that the ratio on the left-hand side of Equation (12) is a measure of how much work the agent allocates early. The first ratio on the right-hand side, in turn, is the same measure of how much he would allocate early if he behaved according to his perceived preference — that is, he had present bias $\tilde{\beta}$ — and knew the fundamental. If the agent is completely naive, for instance, this is time-consistent behavior with perfect information. The rest of the right-hand side is therefore how much the agent's behavior deviates from his perceived and preferred behavior. This is composed of two factors. The first factor, $\beta/\tilde{\beta}$, is standard: due to his present bias, the agent allocates too little work early. The second factor, $h(\tilde{\theta})/h(\Theta)$, is a new term due to misspecified learning.

As a natural starting point, suppose that $h(\theta) = \mathbb{E}[e^{s'_t}]$, where $s'_t \sim N(\theta, \sigma_\epsilon^2)$ and s'_t is independent of other random variables. This captures a situation where the agent faces iid shocks to his value of leisure, and he does not know the shock in the second subperiod when deciding how much to work in the first subperiod. Then, Proposition 7 says that $\pi_{\tilde{\theta}}(s) = \tilde{\pi}_\Theta(s)$: the agent acts exactly as he perceives and wishes, so that misdirected learning is not only not self-defeating, but beneficial. For instance, if the agent is completely naive, then he comes to behave in a time-consistent manner. Intuitively, as the agent observes that he tends to delay work, he incorrectly infers that his value of time is high. This misinference induces him to do more work early, counteracting his present bias. In the case of iid shocks, the mislearning exactly offsets his present bias.

The agent's policy function in the SOE for iid shocks contrasts in an interesting way with that of a present-biased agent who is correctly specified about his instantaneous utility function. In the latter case, the agent's behavior depends on his true discount factor β, and not at all on his perceived discount factor $\tilde{\beta}$. Hence, his welfare does not depend on his naivete. In the SOE of our model, the agent's behavior depends on his perceived discount factor $\tilde{\beta}$, and not at all on his true discount factor β. Furthermore, the more naive he is (the higher is $\tilde{\beta}$), the higher is his welfare.

The logic that mislearning counteracts the agent's present bias holds for any strictly increasing h. If $h(\tilde{\theta})/h(\Theta) > \tilde{\beta}/\beta$, however, the agent's behavior overshoots the optimal action, and he ends up working too much early relative to his wished-for preferences. Given that $h(\tilde{\theta})/h(\Theta)$ can be arbitrarily large, mislearning can be arbitrarily harmful.

As did our earlier models, this model also satisfies apparent sophistication — at the SOE, the agent predicts his future behavior perfectly, so he would pass prediction tests of sophistication commonly used in economics. In the case of a fully naive decisionmaker, in particular, a prediction test combined with a measure of the agent's willingness to pay for commitment would conclude that he has no present bias and is rational. His welfare, however, can be equal to as well as arbitrarily lower than a rational time-consistent decisionmaker's.

References

Acland, Dan and Matthew R. Levy, "Naiveté, Projection Bias, and Habit Formation in Gym Attendance," *Management Science*, 2015, *61* (1), 146–160.

Allcott, Hunt, Joshua Kim, Dmitry Taubinsky, and Jonathan Zinman, "Are High-Interest Loans Predatory? Theory and Evidence from Payday Lending," *Review of Economic Studies*, 2022, *89* (3), 1041–1084.

_ , Matthew Gentzkow, and Lena Song, "Digital Addiction," 2021. Working Paper.

Anderson, Kristin L. and Debra Umberson, "Gendering Violence: Masculinity and Power in Men's Accounts of Domestic Violence," *Gender and Society*, 2001, *15* (3), 358–380.

Augenblick, Ned and Matthew Rabin, "An Experiment on Time Preference and Misprediction in Unpleasant Tasks," *Review of Economic Studies*, 2019, *86* (3), 941–975.

Ba, Cuimin, "Robust Misspecified Models and Paradigm Shift," 2022. Working Paper.

_ and Alice Gindin, "A Multi-Agent Model of Misspecified Learning with Overconfidence," *Games and Economic Behavior*, November 2023, *142*, 315–338.

Bai, Liang, Benjamin Handel, Edward Miguel, and Gautam Rao, "Self-Control and Demand for Preventive Health: Evidence from Hypertension in India," *Review of Economics and Statistics*, forthcoming.

Barron, Kai, Ruth Ditlmann, Stefan Gehrig, and Sebastian Schweighofer-Kodritsch, "Explicit and Implicit Belief-Based Gender Discrimination: A Hiring Experiment," 2020. Working Paper.

Battigalli, Pierpaolo, Simone Cerreia-Vioglio, Fabio Maccheroni, Massimo Marinacci, and Thomas Sargent, "A Framework for the Analysis of Self-Confirming Policies," *Theory and Decision*, 2022, *92* (3), 455–512.

Becker, Gary S. and Kevin M. Murphy, "A Theory of Rational Addiction," *Journal of Political Economy*, 1988, *96* (4), 675–700.

_ , Michael Grossman, and Kevin M. Murphy, "An Empirical Analysis of Cigarette Addiction," *American Economic Review*, 1994, *84* (3), 396–418.

Bénabou, Roland and Jean Tirole, "Self-Confidence and Personal Motivation," *Quarterly Journal of Economics*, 2002, *117* (3), 871–915.

— **and** — , "Identity, Morals, and Taboos: Beliefs as Assets," *Quarterly Journal of Economics*, 2011, *126* (2), 805–855.

Benjamin, Daniel J., "Errors in Probabilistic Reasoning and Judgmental Biases," in Douglas B. Bernheim, Stefano DellaVigna, and David I. Laibson, eds., *Handbook of Behavioral Economics*, Vol. 2, Elsevier, 2019.

— **, Matthew Rabin, and Collin Raymond**, "A Model of Non-Belief in the Law of Large Numbers," *Journal of the European Economic Association*, 2016, *14* (2), 515–544.

Bernheim, B. Douglas and Antonio Rangel, "Addiction and Cue-Triggered Decision Processes," *American Economic Review*, 2004, *94* (5), 1558–1590.

— **and Raphael Thomadsen**, "Memory and Anticipation," *Economic Journal*, 2005, *115* (503), 271–304.

Bertrand, Marianne and Esther Duflo, "Field Experiments on Discrimination," in Esther Duflo and Abhijit Banerjee, eds., *Handbook of Field Experiments*, Vol. 1, Elsevier, 2017, chapter 8, pp. 309–394.

Bodner, Ronit and Drazen Prelec, "The Emergence of Private Rules in a Self-Signaling Model," *International Journal of Psychology*, 1996, *31*, 3652–3653.

Bodoh-Creed, Aaron L., "Mood, Memory, and the Evaluation of Asset Prices," *Review of Finance*, 2019, *24* (1), 227–262.

Bohren, J. Aislinn, "Informational Herding with Model Mispecification," *Journal of Economic Theory*, 2016, *163*, 222–247.

— **, Alex Imas, and Michael Rosenberg**, "The Dynamics of Discrimination: Theory and Evidence," *American Economic Review*, 2019, *109* (10), 3395–3436.

— **and Daniel Hauser**, "Misinterpreting Social Outcomes and Information Campaigns," 2019. Working Paper.

— **and** — , "Learning with Heterogeneous Misspecified Models: Characterization and Robustness," *Econometrica*, forthcoming.

— **, Kareem Haggag, Alex Imas, and Devin G. Pope**, "Inaccurate Statistical Discrimination: An Identification Problem," *Review of Economics and Statistics*, forthcoming.

Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, Frederik Schwerter, and Andrei Shleifer, "Memory and Representativeness," *Psychological Review*, 2021, *128* (1), 71–85.

— **, Nicola Gennaioli, and Andrei Shleifer**, "Memory, Attention, and Choice," *Quarterly Journal of Economics*, 2020, *135* (3), 1399–1442.

Boring, Anne and Arnaud Philippe, "Reducing Discrimination in the Field: Evidence from an Awareness Raising Intervention Targeting Gender Biases in Student Evaluations of Teaching," *Journal of Public Economics*, 2021, *193*, 1–17.

Brunnermeier, Markus and Jonathan Parker, "Optimal Expectations," *American Economic Review*, 2005, *95* (4), 1092–1118.

Bushong, Benjamin and Tristan Gagnon-Bartsch, "Reference Dependence and Attribution Bias: Evidence from Real-Effort Experiments," 2020. Working Paper.

Carlana, Michela, Eliana La Ferrara, and Paolo Pinotti, "Implicit Stereotypes in Teachers' Track Recommendations," *AEA Papers and Proceedings*, 2022, *112*, 409–412.

Carrera, Mariana, Heather Royer, Mark Stehr, Justin Sydnor, and Dmitry Taubinsky, "Who Chooses Commitment? Evidence and Welfare Implications," *Review of Economic Studies*, 2022, *89* (3), 1205–1244.

Chaloupka, Frank J., Matthew R. Levy, and Justin S. White, "Estimating Biases in Smoking Cessation: Evidence from a Field Experiment," 2019. Working Paper.

Cunningham, Tom and Jonathan de Quidt, "Implicit Preferences," 2022. Working Paper.

Cushman, Fiery, "Rationalization Is Rational," *Behavioral and Brain Sciences*, 2020, *43* (e28), 1–59.

DeSante, Christopher D., "Working Twice as Hard to Get Half as Far: Race, Work Ethic, and America's Deserving Poor," *American Journal of Political Science*, 2013, *57* (2), 342–356.

Dovidio, John F. and Samuel L. Gaertner, "Aversive Racism and Selection Decisions: 1989 and 1999," *Psychological Science*, 2000, *11* (4), 315–319.

Drobner, Christoph, "Motivated Beliefs and Anticipation of Uncertainty Resolution," *American Economic Review: Insights*, 2022, *4* (1), 89–105.

Eisikovits, Zvi and Eli Buchbinder, "Talking Violent: A Phenomenological Study of Metaphors Battering Men Use," *Violence Against Women*, 1997, *3* (5), 482–498.

England, Paula and Peter Lewin, "Economic and Sociological Views of Discrimination in Labor Markets: Persistence or Demise?," *Sociological Spectrum*, 1989, *9* (3), 239–257.

Ericson, Keith M. Marzilli and David I. Laibson, "Intertemporal Choice," in B. Douglas Bernheim, Stefano DellaVigna, and David I. Laibson, eds., *Handbook of Behavioral Economics - Foundations and Applications 2*, Elsevier, 2019.

Esponda, Ignacio and Demian Pouzo, "Berk-Nash Equilibrium: A Framework for Modeling Agents with Misspecified Models," *Econometrica*, 2016, *84* (3), 1093–1130.

_ , _ , and Yuichi Yamamoto, "Asymptotic Behavior of Bayesian Learners with Misspecified Models," 2021. Working Paper.

Ettinger, David and Philippe Jehiel, "A Theory of Deception," *American Economic Journal: Microeconomics*, 2010, *2* (1), 1–20.

Fang, Hanming and Yang Wang, "Estimating Dynamic Discrete Choice Models with Hyperbolic Discounting, with an Application to Mammography Decisions," *International Economic Review*, 2015, *56* (2), 565–596.

Fedyk, Anastassia, "Asymmetric Naivete: Beliefs About Self-Control," 2021. Working Paper.

Fotuhi, Omid, Geoffrey T. Fong, Mark P. Zanna, Ron Borland, Hua-Hie Yong, and K. Michael Cummings, "Patterns of Cognitive Dissonance-Reducing Beliefs Among Smokers: A Longitudinal Analysis from the International Tobacco Control (ITC) Four Country Survey," *Tobacco Control*, 2013, *22* (1), 52–58.

Frick, Mira, Ryota Iijima, and Yuhta Ishii, "Misinterpreting Others and the Fragility of Social Learning," *Econometrica*, 2020, *88* (6), 2281–2328.

_ , _ , and _ , "Belief Convergence under Misspecified Learning: A Martingale Approach," 2021. Working Paper.

_ , _ , and _ , "Welfare Comparisons for Biased Learning," 2021. Working Paper.

_ , _ , and _ , "Dispersed Behavior and Perceptions in Assortative Societies," *American Economic Review*, 2022, *112* (9), 3063–3105.

Fudenberg, Drew and David K. Levine, "Self-Confirming Equilibrium," *Econometrica*, 1993, *61* (3), 523–545.

_ and _ , "A Dual-Self Model of Impulse Control," *American Economic Review*, 2006, *96* (5), 1449–1476.

_ , Giacomo Lanzani, and Philipp Strack, "Limit Points of Endogenous Misspecified Learning," *Econometrica*, 2021, *89* (3), 1065–1098.

_ , _ , and _ , "Pathwise Concentration Bounds for Misspecified Bayesian Beliefs," 2021. Working Paper.

_ , _ , and _ , "Selective Memory Equilibrium," *Journal of Political Economy*, forthcoming.

_ , Gleb Romanyuk, and Philipp Strack, "Active Learning with a Misspecified Prior," *Theoretical Economics*, 2017, *12* (3), 1155–1189.

Goette, Lorenz and Marta Kozakiewicz, "Experimental Evidence on Misguided Learning," 2018. Working Paper.

Grant, Jon, Marc Potenza, Aviv Weinstein, and David Gorelick, "Introduction to Behavioral Addictions," *American Journal of Drug and Alcohol Abuse*, 2010, *36* (5), 233–241.

Grossman, Zachary and Joël J. van der Weele, "Self-Image and Willful Ignorance in Social Decisions," *Journal of the European Economic Association*, 2017, *15* (1), 173–217.

Gruber, Jonathan and Botond Kőszegi, "Is Addiction 'Rational?' Theory and Evidence," *Quarterly Journal of Economics*, 2001, *116* (4), 1261–1305.

Haggag, Kareem, Devin G. Pope, Kinsey B. Bryant-Lees, and Maarten W. Bos, "Attribution Bias in Consumer Choice," *Review of Economic Studies*, 2018, *86* (5), 2136–2183.

He, Kevin, "Mislearning from Censored Data: The Gambler's Fallacy and Other Correlational Mistakes in Optimal-Stopping Problems," 2021. Working Paper.

Heidhues, Paul and Botond Kőszegi, "Exploiting Naivete about Self-Control in the Credit Market," *American Economic Review*, 2010, *100* (5), 2279–2303.

_ , _ , and Philipp Strack, "Unrealistic Expectations and Misguided Learning," *Econometrica*, 2018, *86* (4), 1159–1214.

_ , _ , and _ , "Convergence in Models of Misspecified Learning," *Theoretical Economics*, 2021, *16* (1), 73–99.

_ , _ , and _ , "Overconfidence and Prejudice," 2022. Working Paper.

Ippolito, Pauline M. and Richard A. Ippolito, "Measuring the Value of Life Saving from Consumer Reactions to New Information," *Journal of Public Economics*, 1984, *25* (1-2), 53–81.

John, Anett, "When Commitment Fails: Evidence from a Field Experiment," *Management Science*, 2020, *66* (2), 503–529.

Kőszegi, Botond, George Loewenstein, and Takeshi Murooka, "Fragile Self-Esteem," *Review of Economic Studies*, 2022, *89* (4), 2026–2060.

Kuchler, Theresa and Michaela Pagel, "Sticking to Your Plan: Empirical Evidence on the Role of Present Bias for Credit Card Debt Paydown," *Journal of Financial Economics*, 2021, *139* (2), 359–388.

Kushner, Harold J. and G. George Yin, *Stochastic Approximation and Recursive Algorithms and Applications*, Springer, 2003.

Laibson, David I., "Golden Eggs and Hyperbolic Discounting," *Quarterly Journal of Economics*, 1997, *112* (2), 443–477.

Levy, Gilat and Ronny Razin, "The Coevolution of Segregation, Polarized Beliefs, and Discrimination: The Case of Private versus State Education," *American Economic Journal: Microeconomics*, 2017, *9* (4), 141–170.

_ , _ , and Alwyn Young, "Misspecified Politics and the Recurrence of Populism," *American Economic Review*, 2022, *112* (3), 928–962.

Liu, Jin-Long, Jin-Tan Liu, James K. Hammitt, and Shin-Yi Chou, "The Price Elasticity of Opium in Taiwan, 1914–1942," *Journal of Health Economics*, 1999, *18* (6), 795–810.

Loewenstein, George, Ted O'Donoghue, and Matthew Rabin, "Projection Bias in Predicting Future Utility," *Quarterly Journal of Economics*, 2003, *118* (4), 81–123.

Mullainathan, Sendhil, "A Memory-Based Model of Bounded Rationality," *Quarterly Journal of Economics*, 2002, *117* (3), 735–774.

Murooka, Takeshi and Yuichi Yamamoto, "Misspecified Bayesian Learning by Strategic Players: First-Order Misspecification and Higher-Order Misspecification," *Working Paper*, 2021.

Nisbett, Richard E. and Timothy DeCamp Wilson, "Telling More Than We Can Know: Verbal Reports on Mental Processes," *Psychological Review*, 1977, *84* (3), 231–259.

Nyarko, Yaw, "Learning in Mis-Specified Models and the Possibility of Cycles," *Journal of Economic Theory*, 1991, *55* (2), 416–427.

Oakes, Wendy, Simon Chapman, Ron Borland, James Balmford, and Lisa Trotter, ""Bulletproof Skeptics in Life's Jungle": Which Self-Exempting Beliefs About Smoking Most Predict Lack of Progression Towards Quitting?," *Preventive Medicine*, 2004, *39* (4), 776–782.

O'Donoghue, Ted and Matthew Rabin, "Doing It Now or Later," *American Economic Review*, 1999, *89* (1), 103–124.

— **and** — , "Choice and Procrastination," *Quarterly Journal of Economics*, 2001, *116* (1), 121–160.

Orphanides, Athanasios and David Zervos, "Rational Addiction with Learning and Regret," *Journal of Political Economy*, 1995, *103* (4), 739–758.

Oster, Emily, Ira Shoulson, and E. Ray Dorsey, "Optimal Expectations and Limited Medical Testing: Evidence from Huntington Disease," *American Economic Review*, 2013, *103* (2), 804–30.

Payne, B. Keith, Jon A. Krosnick, Josh Pasek, Yphtach Lelkes, Omair Akhtar, and Trevor Tompson, "Implicit and Explicit Prejudice in the 2008 American Presidential Election," *Journal of Experimental Social Psychology*, 2010, *46* (2), 367–374.

Petruzzello, Esteban, "Testing for Forward-Looking Behaviour: Evidence from the Enactment of Smoking Restrictions," *Applied Economics*, 2019, *51* (19), 2061–2069.

Pope, Devin G., Joseph Price, and Justin Wolfers, "Awareness Reduces Racial Bias," *Management Science*, 2018, *64* (11), 4988–4995.

Prelec, Drazen and Ronit Bodner, "Self-Signaling and Self-Control," in "Time and Decision: Economic and Psychological Perspectives on Intertemporal Choice," New York, NY, US: Russell Sage Foundation, 2003, pp. 277–298.

Rabin, Matthew, "Cognitive Dissonance and Social Change," *Journal of Economic Behavior & Organization*, 1994, *23* (2), 177–194.

— , "Inference by Believers in the Law of Small Numbers," *Quarterly Journal of Economics*, 2002, *117* (3), 775–816.

— **and Dimitri Vayanos**, "The Gambler's and Hot-Hand Fallacies: Theory and Applications," *Review of Economic Studies*, 2010, *77* (2), 730–778.

Rees-Jones, Alex and Kyle Rozema, "Price Isn't Everything: Behavioral Response around Changes in Sin Taxes," 2020. Working Paper.

Ridout, Sarah, "Choosing for the Right Reasons," 2021. Working Paper.

Rudman, Laurie A. and Peter Glick, "Prescriptive Gender Stereotypes and Backlash Toward Agentic Women," *Journal of Social Issues*, 2001, *57* (4), 743–762.

Sharot, Tali and Neil Garrett, "Forming Beliefs: Why Valence Matters," *Trends in Cognitive Sciences*, 2016, *20* (1), 25–33.

Skiba, Paige Marta and Jeremy Tobacman, "Payday Loans, Uncertainty, and Discounting: Explaining Patterns of Borrowing, Repayment, and Default," 2008. Vanderbilt Law and Economics Research Paper No. 08-33.

Son Hing, Leanne S., Winnie Li, and Mark P. Zanna, "Inducing Hypocrisy to Reduce Prejudicial Responses among Aversive Racists," *Journal of Experimental Social Psychology*, 2002, *38* (1), 71–78.

Spiegler, Ran, "Bayesian Networks and Boundedly Rational Expectations," *Quarterly Journal of Economics*, 2016, *131* (3), 1243–1290.

_ , "Behavioral Implications of Causal Misperceptions," *Annual Review of Economics*, 2020, *12*, 81–106.

Uhlmann, Eric Luis and Geoffrey L. Cohen, "Constructed Criteria: Redefining Merit to Justify Discrimination," *Psychological Science*, 2005, *16* (6), 474–480.

Viscusi, W. Kip, "Public Perception of Smoking Risks," *International Conference on the Social Costs of Tobacco*, 1998.

_ **and Joni Hersch**, "The Mortality Cost to Smokers," *Journal of Health Economics*, 2008, *27*, 943–958.

Wachter, Jessica A. and Michael J. Kahana, "A Retrieved-Context Theory of Financial Decisions," 2021. Working Paper.

Yaouanq, Yves Le and Peter Schwardmann, "Learning about One's Self," *Journal of the European Economic Association*, 2022, *20* (5), 1791–1828.

Zimmermann, Florian, "The Dynamics of Motivated Beliefs," *American Economic Review*, 2020, *110* (2), 337–361.

A Updating Problem, Existence, and Convergence

This section establishes convergence of the sequence $(\mu_t)_t$ of the agent's beliefs. We start with some definitions and assumptions. First, we extend the optimal and perceived-optimal action functions $\pi_\theta(s)$ and $\tilde{\pi}_\theta(s)$ to general beliefs μ with which the agent may enter the period:[30]

$$\pi_\mu(s) = \arg\max_{a \in A} \int v(a, s, z) f(s - z) d\mu(z) \tag{13}$$

$$\tilde{\pi}_\mu(s) = \arg\max_{a \in A} \int \tilde{v}(a, s, z) f(s - z) d\mu(z)\,. \tag{14}$$

We also extend our monotonicity assumption guaranteeing that the policy functions are invertible in s_t:

Assumption 1. For every sequence of actions $(a_1, \ldots, a_{t-1})$ and the corresponding induced posterior $\mu_t \equiv \mu_t(\cdot; a_1, \ldots, a_{t-1})$, the policies $\pi_{\mu_t}, \tilde{\pi}_{\mu_t}$ are both either strictly increasing or strictly decreasing in s_t.

While Assumption 1 amounts to a joint assumption on beliefs and utility functions, we verify below that it is satisfied in all of our applications.

Belief Updating Given the agent's perceived strategy $\tilde{\pi}$, he believes to have observed the signal

$$\tilde{s}_t = \tilde{\pi}_{\mu_t}^{-1}(a_t)\,.$$

Hence, if he took the action a_t, by Bayes' Rule for any $C \subseteq \mathbb{R}$ he updates his beliefs according to

$$\mu_{t+1}(C) = \frac{\int_{z \in C} f(\tilde{\pi}_{\mu_t}^{-1}(a_t) - z) d\mu(z)}{\int_{z \in \mathbb{R}} f(\tilde{\pi}_{\mu_t}^{-1}(a_t) - z) d\mu(z)}\,. \tag{15}$$

Using that the agent's true strategy is given by $a_t = \pi_{\mu_t}(s_t)$, we can thus express the dynamics of the agent's belief process only in terms of the sequence of signals $s_1, \ldots, s_t$:

$$\mu_{t+1}(C) = \frac{\int_{z \in C} f\left(\tilde{\pi}_{\mu_t}^{-1}\left(\pi_{\mu_t}(s_t)\right) - z\right) d\mu(z)}{\int_{z \in \mathbb{R}} f\left(\tilde{\pi}_{\mu_t}^{-1}\left(\pi_{\mu_t}(s_t)\right) - z\right) d\mu(z)}\,. \tag{16}$$

[30] In this case, the agent updates μ using s before choosing his action. Formally, the posterior belief distribution after observing s is given by $\frac{f(s-z)d\mu(z)}{\int f(s-z')d\mu(z')}$, but as the maximizer remains unchanged when multiplying the objective by the constant $\int f(s - z')d\mu(z')$, we can state $\pi_\mu, \tilde{\pi}_\mu$ in this simpler form.

As conditional on the true fundamental signals are independent, (16) establishes that the agent's beliefs follow a Markov process.

We next introduce an assumption to ensure that the agent's misspecification is not "too large" and well-behaved:

Assumption 2. There exists a constant $k > 0$ such that for every time t, every sequence of past actions $(a_1, \ldots, a_{t-1})$, and the corresponding induced posterior $\mu_t \equiv \mu_t(\cdot; a_1, \ldots, a_{t-1})$, the function $s \mapsto \tilde{\pi}_{\mu_t}^{-1}(\pi_{\mu_t}(s))$ is continuous; and for every action a

$$\left|\pi_{\mu_t}^{-1}(a) - \tilde{\pi}_{\mu_t}^{-1}(a)\right| \leq k\,.$$

Assumption 2 requires that the difference between the signal the agent believes to have observed and the true signal is bounded. Absent this assumption, the agent's beliefs might diverge, as his changing actions might lead his misinterpretation of the signals, and thus misinference about Θ, to keep increasing.

We now turn to analyzing the agent's long-run beliefs and actions. As a useful benchmark, we first observe that in our setting a correctly specified agent (for whom $v \equiv \tilde{v}$) learns the true fundamental Θ despite his incomplete memory. By Assumption 1, such an agent correctly infers his past signals from his past actions, and because signals are i.i.d. conditional on Θ, by the law of large numbers his beliefs converge to Θ. Thus, his action converges to the optimal action given Θ:

Observation 2. If the agent is correctly specified, i.e. $\tilde{v} = v$, and Assumption 1 is satisfied, then the agent's belief $(\mu_t)_t$ a.s. concentrate on the true fundamental Θ and the agent's actions $(a_t)_t$ a.s. converge to the action $\pi_\Theta(s)$ that is optimal given Θ.

Even more simply, if the agent knew the s_t and updated based on them, then despite his incorrect self-view he would learn Θ. It is therefore the *combination* of an incorrect self-view and limited memory or self-awareness that leads to the mislearning results of our paper.

Consider now a misspecified agent (i.e., $\tilde{v} \neq v$). We define the subjective log-likelihood maximizer (which equals the posterior mean) given past beliefs $(\mu_1, \ldots, \mu_{t-1})$ and actions $(a_1, \ldots, a_{t-1})$ as:[31]

$$\tilde{\theta}_t = \arg\max_{z\in\mathbb{R}} \left[\sum_{r=1}^{t-1} \log f\left(\tilde{\pi}_{\mu_r}^{-1}(a_r) - z\right) + \mu_1(z)\right]. \tag{17}$$

[31] Since the density of the normal distribution is log-concave, the argmax is unique and $\tilde{\theta}_t$ is well defined.

Intuitively, since the fundamental $\tilde{\theta}_t$ maximizes the log-likelihood given past actions, it best explains the agent's past actions. Now because the agent starts with a normal prior and believes to see independent draws from a normal distribution, the updating formula for conjugate priors implies that his beliefs concentrate around this log-likelihood maximizer in the long-run: there exists a constant $c > 0$ such that for every t and sequence of signals s, we have

$$\int_{\mathbb{R}} (z - \tilde{\theta}_t)^2 d\mu_t(z) \leq \frac{c}{t}\,.$$

Given that the agent becomes subjectively certain that the fundamental is the log-likelihood maximizer, it is intuitive that — for a reasonably behaved payoff functions v and $\tilde{v}$ — his action can be approximated by the action that is optimal when having a point belief at $\tilde{\theta}_t$. Indeed, this is the case in all our applications below, and to prove convergence we henceforth assume that such an approximation is possible.

Assumption 3. There exist constants $c_1, c_2 > 0$ such that for every sequence of past actions $(a_1, \ldots, a_{t-1})$ and the corresponding induced posteriors $\mu_t \equiv \mu_t(\cdot; a_1, \ldots, a_{t-1})$ and every signal s, we have

$$\left|\tilde{\pi}_{\mu_t}^{-1}\left(\pi_{\mu_t}(s)\right) - \tilde{\pi}_{\tilde{\theta}_t}^{-1}\left(\pi_{\tilde{\theta}_t}(s)\right)\right| \leq c_1 t^{-c_2}\,.$$

Despite the actions converging to those that are optimal given point beliefs on the log-likelihood-maximizing fundamental, it remains unclear whether the log-likelihood-maximizing beliefs converge. The agent's misspecification of his own payoff function implies that how he interprets past actions depends on his own past beliefs going into the period. Hence, the signals he infers from his actions are not iid conditional on the true fundamental, so that off-the-shelf convergence results do not apply.[32] We overcome this difficulty by analyzing the dynamics of an auxiliary process where we only keep track of the subjective log-likelihood-maximizing state and assume that the agent's actions are optimal given point beliefs on that state. This belief process is real-valued and tractable. Furthermore, we show that in the long run this process approximates the agent's beliefs well, which leads to the following result.

Proposition 8. *An SOE exists. If there are finitely many SOEs, the agent's beliefs $(\mu_t)_t$ almost*

[32] Examples of misspecified learning settings where there is no belief convergence include Nyarko (1991) and Fudenberg et al. (2017).

surely converge to an SOE.

Proposition 8 allows us to determine the agent's long-run beliefs and behavior not by analyzing the dynamics of the belief process — which lives in the space of distributions over the reals — but by solving (17) — a *static* fixed-point equation over the reals.

Relation to Other Belief Convergence Results Due to the different nature of the decision-making problem, we cannot directly apply existing results on convergence to Berk-Nash Equilibrium (Esponda and Pouzo, 2016, Fudenberg et al., 2017, Heidhues et al., 2018, Bohren and Hauser, 2019, Heidhues et al., 2021, Frick et al., 2021a, Fudenberg et al., 2021a, Esponda et al., 2021) in our setting. In this literature (with exceptions discussed below), the dynamics of the agent's beliefs — defined as the distribution over changes in his subjective log-likelihood ratios — depend only on his action; in our model, in contrast, the dynamics also depend on the belief he held at the time of taking the action.[33] In addition, although our formal model is a social-learning model with misspecification about others' preferences, most papers studying such models (Bohren, 2016, Bohren and Hauser, 2019, Frick et al., 2020) do not consider continuous states and actions. As the lone exception, Frick et al.'s (2021a) appendix develops a framework with general states that subsumes our model, but we see no obvious way of verifying their convergence-ensuring iterated-dominance condition in our setting. At the same time, although Heidhues et al. (2021) and Esponda et al. (2021) analyze formally and economically very different problems, they use related stochastic-approximation arguments to prove convergence to Berk-Nash equilibrium.

B Appendix: Short-Horizon Consumption-Savings Example

Consider an agent who repeatedly faces a short-horizon consumption-savings problem. Each period t is divided into three subperiods, $t.0$, $t.1$, $t.2$. The agent is a partially naive quasi-hyperbolic discounter who discount future utility by β while thinking that all other selves use $\tilde{\beta}$.

At $t.0$, the agent receives income I and decides how much to spend on consumption c_t, which yields instantaneous utility $u_0(c_t)$. In subperiod $t.1$, the agent chooses consumption a_t, which gives instantaneous utility $u_1(a_t) = u(a_t) + \phi_t a_t$. We assume that ϕ_t is determined exactly as in our main

[33] One may be tempted to resolve the problem by thinking of the policy functions as the actions, because knowing these and the true state makes observed distributions over actions iid. Given the true fundamental, however, how much the agent misperceives his policy function depends on his beliefs.

model in Section 3.1, that the agent observes the signal s_t at the beginning of subperiod $t.1$, and that $u_0(c_t)$ and $u_1(a_t)$ are twice-differentiable, strictly concave utility functions with a marginal utility approaching ∞ (respectively $-\infty$) as its argument goes to $-\infty$ (respectively ∞). Finally, in period $t.2$ the agent enjoys utility $u_2(I - c_t - a_t) = I - c_t - a_t$.[34] This quasi-linear specification of utility in subperiod $t.2$ closely resembles our basic model with linear harm.

In subperiod t.1, for a given signal s_t and posterior belief μ_t, the agent hence maximizes

$$\max_{a_t} u_1(a_t) + \bar{\phi}(s_t, \mu_t)a_t + \beta(I - c_t - a_t),$$

where the agent's posterior (mean) belief $\bar{\phi}(s_t, \mu_t)$ as a function of s_t and μ_t is formally stated in (25). In any other (sub-)period the agent thinks he has or will use $\tilde{\beta}$ instead of β in the above maximization. Solving the first-order conditions yields the exact same policy functions as in the model of Section 3.1, i.e.:

$$\pi_{\mu_t}(s_t) = (u')^{-1}(\beta\kappa - \bar{\phi}(s_t, \mu_t))$$
$$\tilde{\pi}_{\mu_t}(s_t) = (u')^{-1}(\tilde{\beta}\kappa - \bar{\phi}(s_t, \mu_t)) \, .$$

The policy functions are independent of c_t, and because c_t reveals no information about the fundamental, the agent updates his beliefs about Θ exactly as in the model of Section 3.1.

In subperiod $t.0$, the agent maximizes

$$\max_{c_t} u_0(c_t) + \beta\,\tilde{\mathbb{E}}_{\mu_t}[u_1(a_t) + \phi_t a_t + \beta(I - c_t - a_t)],$$

where the subjective expectation is based on the agent's current belief μ_t regarding the fundamental and the presumption that his future self uses policy function $\tilde{\pi}_{\mu_t}(s_t)$. Because this policy function, and hence the anticipated choice of a_t, is independent of c_t, the agent chooses $c_t = (u_0')^{-1}(\beta)$. Moreover, it follows from the equivalence of choices and beliefs in subperiod $t.1$ to those of our model in Section 3.1 that the agent overconsumes even from self $t.1$'s perspective.

Because the outside observer's econometric estimates are derived from choices at $t.1$ and predictions and commitment choices about them, the fact that there is a decision at $t.0$ does not change the observer's estimates. The outside observer, hence, incorrectly concludes that the agent is sophisticated with a present-bias parameter $\tilde{\beta}$. The observer, thus, incorrectly anticipates the

[34] The specification of consumption in subperiod $t.2$ amounts to assuming that the gross interest rate is 1.

agent to — from a long-run perspective — (over-)consume less in subperiod $t.0$; she anticipates an experienced agent to choose a consumption level $(u_0')^{-1}(\tilde{\beta})$ instead of $c_t = (u_0')^{-1}(\beta)$. Furthermore, despite correctly predicting her subperiod-$t.1$ behavior, the agent overconsumes in $t.0$ and $t.1$ also from the $\tilde{\beta}$-preference perspective.

Consider next an extension in which the $u_0(c_t) = u(c_t) + \phi_t c_t$, so that the agent's learning about the marginal utility of consumption applies equally to subperiods $t.0$ and $t.1$. We suppose the remainder of the model remains specified as above. In particular, we still suppose that the agent observes s_t at the beginning of subperiod $t.1$. This implies that his choice at subperiod $t.0$ contains no information about the fundamental, so learning about the fundamental continues to take place as above. Furthermore, since the agent's $t.1$-policy functions above are independent of c_t, behavior in $t.1$ remains unaltered. In subperiod $t.0$, the agent now chooses consumption such that $\mathbb{E}_{\mu_t}[u_0'(c_t)] = u'(c_t) + \mathbb{E}_{\mu_t}[\phi_t] = \beta$. Because an agent with SOE beliefs overestimates the marginal utility of consumption, she now overconsumes in subperiod $t.0$ (as well as in subperiod $t.1$) even from self $t.0$'s perspective; that is she chooses $c_t = (u')^{-1}(\beta - \tilde{\theta}) > (u')^{-1}(\beta - \Theta)$. Our above outside observer with access to period $t.1$ choices as well as predictions and commitments regarding these, would underpredict this overconsumption.

C Appendix: Proofs

We denote by σ_ϵ^2 the variance of the shocks $(s_t)_t$. Throughout of the appendix, we use the notation $(\epsilon_t)_t$ for the sequence of shocks and $\epsilon^t = (\epsilon_1, \ldots, \epsilon_t)$ for the vector of length t of past realizations of shocks. We label the sequence and past realizations of actions, signals, perceived signals, and beliefs analogously.

Below, we denote by $\mathbb{P}[\cdot]$ the (objective) probability measure over beliefs induced by (16) given the fundamental Θ, and by $\mathbb{E}[\cdot]$ the corresponding (objective) expectation.

C.1 Belief Dynamics

We begin by describing the dynamics of the agent's belief. For the sake of doing so and establishing the conversion argument, it is convenient to normalize the policy functions as being increasing in the signal (respectively perceived signal). If, as in some of our applications, the policy functions are decreasing we can just re-normalize the signal by multiplying it with minus one.

Since the agent's prior is Normally distributed and the signals are Normally distributed, if the agent believes that the sequence of past signals was equal to $\tilde{s}^{t-1}$ his posterior belief μ_t is also Normally distributed. We denote by $\tilde{\theta}_t$ the posterior mean and by σ_t^2 the posterior variance of the posterior belief μ_t. It is well known that[35]

$$\tilde{\theta}_t = \frac{\sigma_1^{-2}\theta_1 + \sigma_\epsilon^{-2}\sum_{r=1}^{t-1}\tilde{s}_r}{\sigma_1^{-2} + \sigma_\epsilon^{-2}(t-1)} \qquad\qquad \sigma_t^2 = \frac{1}{\sigma_1^{-2} + \sigma_\epsilon^{-2}(t-1)}\,. \tag{18}$$

Furthermore, as the posterior belief is Normal the posterior mean equals the log-likelihood maximizer defined in (17), which is why we denote them with the same symbol. The dynamics of the posterior mean can be rewritten as

$$\begin{aligned}\tilde{\theta}_t &= \frac{\sigma_1^{-2}\theta_1 + \sigma_\epsilon^{-2}\sum_{r=1}^{t-2}\tilde{s}_r}{\sigma_1^{-2} + \sigma_\epsilon^{-2}(t-1)} + \frac{\sigma_\epsilon^{-2}\tilde{s}_{t-1}}{\sigma_1^{-2} + \sigma_\epsilon^{-2}(t-1)} = \frac{\sigma_1^{-2} + \sigma_\epsilon^{-2}(t-2)}{\sigma_1^{-2} + \sigma_\epsilon^{-2}(t-1)}\tilde{\theta}_{t-1} + \frac{\sigma_\epsilon^{-2}}{\sigma_1^{-2} + \sigma_\epsilon^{-2}(t-1)}\tilde{s}_{t-1} \\ &= \tilde{\theta}_{t-1} + \frac{\sigma_t^2}{\sigma_\epsilon^2}(\tilde{s}_{t-1} - \tilde{\theta}_{t-1})\,. \end{aligned} \tag{19}$$

To prove Proposition 8, we first prove several auxiliary results about the behavior of the agent's beliefs. We define θ_t to be the mean of the belief an outside observer would hold after observing the true signals s^{t-1} when starting from the same prior μ_1 as the agent, i.e.

$$\theta_t = \frac{\sigma_1^{-2}\theta_1 + \sigma_\epsilon^{-2}\sum_{r=1}^{t-1}s_r}{\sigma_1^{-2} + \sigma_\epsilon^{-2}(t-1)}\,.$$

Our next result establishes that the distance between the mean of the agent's belief and that of an outside observer is uniformly bounded.

Lemma 1. *If $|s_r - \tilde{s}_r| \le c$ for all r then $|\theta_t - \tilde{\theta}_t| < c$.*

Proof. We get that the difference between $\tilde{\theta}_t$ and θ_t is uniformly bounded:

$$|\theta_t - \tilde{\theta}_t| = \left|\frac{\sigma_\epsilon^{-2}\sum_{r=1}^{t-1}(s_r - \tilde{s}_r)}{\sigma_1^{-2} + \sigma_\epsilon^{-2}(t-1)}\right| \le \frac{\sigma_\epsilon^{-2}\sum_{r=1}^{t-1}|s_r - \tilde{s}_r|}{\sigma_1^{-2} + \sigma_\epsilon^{-2}(t-1)} \le c\frac{\sigma_\epsilon^{-2}(t-1)}{\sigma_1^{-2} + \sigma_\epsilon^{-2}(t-1)} < c\,. \qquad \square$$

We next use this insight to establish that the distance between the subjective mean belief of the agent $\tilde{\theta}_t$ and the true fundamental Θ will be a.s. bounded in the long-run.

[35] See e.g. `https://en.wikipedia.org/wiki/Conjugate_prior#When_likelihood_function_is_a_continuous_distribution`.

Lemma 2. *If* $|s_r - \tilde{s}_r| \le c$ *for all* r *then* $\limsup_{t\to\infty} |\Theta - \tilde{\theta}_t| \le c$ *with probability* 1.

Proof. The strong law of large numbers implies that the mean belief $(\theta_t)_t$ of an outside observer who observes the signals $s_1, s_2, \ldots$ almost surely converges to the true fundamental Θ, and the result then follows directly from Lemma 1. $\square$

To simplify notation, we define $\tilde{y}_t = \sigma_\epsilon^{-2}[\tilde{s}_t - \tilde{\theta}_t]$ and note that given this definition, the dynamics of $\tilde{\theta}$ are given by

$$\tilde{\theta}_{t+1} = \tilde{\theta}_t + \sigma_{t+1}^2 \tilde{y}_t \,. \tag{20}$$

Our next result establishes that the variance of $\tilde{y}_t$ can be uniformly bounded.

Lemma 3. *If* $|s_r - \tilde{s}_r| \le c$ *for all* r *then* $\sup_t \mathbb{E}\left[\tilde{y}_t^2\right] \le \sigma_\epsilon^{-4}\left(2c + \sigma_\epsilon + \sigma_1\right)^2 < \infty$.

Proof. We have that

$$\sigma_\epsilon^2 |\tilde{y}_t| = |\tilde{s}_t - \tilde{\theta}_t| \le c + |s_t - \tilde{\theta}_t| \le 2c + |s_t - \theta_t| \le 2c + |s_t - \Theta| + |\Theta - \theta_t| \,.$$

By the triangle inequality for the L^2 norm, we thus have that

$$\sigma_\epsilon^2 \sqrt{\mathbb{E}\left[|\tilde{y}_t|^2\right]} \le 2c + \sqrt{\mathbb{E}\left[|s_t - \Theta|^2\right]} + \sqrt{\mathbb{E}\left[|\Theta - \theta_t|^2\right]} = 2c + \sigma_\epsilon + \sqrt{\mathbb{E}\left[|\Theta - \theta_t|^2\right]} \le 2c + \sigma_\epsilon + \sqrt{\sigma_1^2}\,,$$

where the equality follows as s_t is Normally distributed with mean Θ and variance σ_ϵ^2; and the final inequality follows since the expected squared distance between Θ and the posterior mean of an outside observer θ_t, equals the posterior variance given in (18) and is monotone decreasing in t and consequently maximized at time 1 when it equals the prior variance σ_1^2. $\square$

Recall that future selves of the agent believe that the signal $\tilde{s}_t$ the time t self has observed is given by $\tilde{s}_t = \tilde{\pi}_{\mu_t}^{-1}(\pi_{\mu_t}(s_t))$. We next define the function $\bar{g} : \mathbb{R} \to \mathbb{R}$ as the objective expectation of an outside observer of $\tilde{y}_t$ if $\mu_t = \delta_{\tilde{\theta}}$, i.e. the agent is subjectively certain the fundamental equals $\tilde{\theta}$,

$$\bar{g}(\tilde{\theta}) = \sigma_\epsilon^{-2}\left(\mathbb{E}\left[\tilde{\pi}_{\tilde{\theta}}^{-1}(\pi_{\tilde{\theta}}(s_t))\right] - \tilde{\theta}\right) \,. \tag{21}$$

We furthermore denote by γ_t the difference between the true expectation of $\tilde{y}_t$ and $\bar{g}(\tilde{\theta}_t)$

$$\gamma_t = \mathbb{E}\left[\tilde{y}_t | s^{t-1}\right] - \bar{g}(\tilde{\theta}_t) = \sigma_\epsilon^{-2}\mathbb{E}\left[\tilde{\pi}_{\mu_t}^{-1}(\pi_{\mu_t}(s_t)) - \tilde{\pi}_{\tilde{\theta}}^{-1}(\pi_{\tilde{\theta}}(s_t))\right] .$$

Lemma 4. *We have that* $\sum_{t=1}^{\infty} \sigma_t^2 |\gamma_t| < \infty$ *.*

Proof. By Assumption 3 there exists constants $c_1, c_2 > 0$ such that $|\gamma_t| \le c_1 t^{-c_2}$. Hence,

$$\sum_{t=1}^{\infty} \sigma_t^2 |\gamma_t| \le \sum_{t=1}^{\infty} \frac{c_1 t^{-c_2}}{\sigma_1^{-2} + \sigma_\epsilon^{-2}(t-1)} \le \frac{c_1}{\sigma_1^{-2}} + \frac{c_1}{\sigma_\epsilon^{-2}} \sum_{t=2}^{\infty} \frac{1}{t-1} t^{-c_2} \le \frac{c_1}{\sigma_\epsilon^{-2}} \sum_{t=2}^{\infty} (t-1)^{-(1+c_2)} < \infty . \quad \square$$

Proof of Proposition 8. Recall that the agent's posterior mean equals his subjective log-likelihood maximizer $\tilde{\theta}_t$. Note also that by Assumption 2 the difference between the true signal s_r and the signal the agent beliefs to have observed $\tilde{s}_r$ is uniformly bounded by c. The dynamic of $\tilde{\theta}$ is given by

$$\tilde{\theta}_{t+1} = \tilde{\theta}_t + \sigma_{t+1}^2 \tilde{y}_t . \tag{22}$$

We will next use a result on the limit behavior of processes with the above dynamic by Kushner and Yin (2003, page 126-128) and begin by verifying the conditions necessary to apply their theorem. By Lemma 3, Condition A2.1 of Kushner and Yin is satisfied. As σ_t^2 vanishes of the order $1/t$, Condition A2.4 of Kushner and Yin is satisfied. As $\bar{g}$ is continuous and $\sum_{t=1}^{\infty} \sigma_t^2 |\gamma_t| < \infty$, Condition A2.3 and A2.5 of Kushner and Yin are satisfied. By Lemma 2, $\tilde{\theta}_t$ is eventually in $[\Theta - c, \Theta + c]$ and $(\theta_t)_t$ is a.s. bounded. Furthermore, as $\bar{g}$ is continuous and $\tilde{\theta} \in \mathbb{R}$, Condition A2.6 of Kushner and Yin holds. We can apply Theorem 2.1 from Kushner and Yin (2003, page 127), which yields that $\tilde{\theta}_t$ converges a.s. to a connected subset S for which $\bar{g}(\tilde{\theta}) = 0$ for all $\tilde{\theta} \in S$.

We next argue that this condition characterizes exactly the set of self-observation equilibria. Note that by Definition 1 and as $f(\epsilon) = \frac{1}{\sqrt{2\pi\sigma_\epsilon^2}} e^{-z^2/(2\sigma_\epsilon^2)}$ an SOE $\tilde{\theta}$ is characterized by

$$\begin{aligned} \tilde{\theta} &= \arg\max_z \int_{-\infty}^{\infty} \log f\left(\tilde{\pi}_{\tilde{\theta}}^{-1}(\pi_{\tilde{\theta}}(s)) - z\right) f(s-\theta) ds \\ &= \arg\max_z \int_{-\infty}^{\infty} -\frac{1}{2\sigma_\epsilon^2} \left(\tilde{\pi}_{\tilde{\theta}}^{-1}(\pi_{\tilde{\theta}}(s)) - z\right)^2 f(s-\theta) ds . \end{aligned}$$

Taking the first-order condition with respect to z (which is necessary and sufficient for a maximum as the objective is strictly concave) yields that for any SOE $\tilde{\theta}$ satisfies

$$0 = \int_{-\infty}^{\infty} \left(\tilde{\pi}_{\tilde{\theta}}^{-1}(\pi_{\tilde{\theta}}(s)) - \tilde{\theta}\right) f(s-\theta) ds = \bar{g}(\tilde{\theta}) .$$

Thus, a fundamental is an SOE if and only if it is root of $\bar{g}$. Since there are only finitely many

SOEs, every connected subset where $\bar{g}$ equals 0 is just a single point and thus beliefs converge to an SOE.

We are left to show that an SOE exists. As we argued above, a state $\tilde{\theta}$ is an SOE if and only if $\bar{g}(\tilde{\theta}) = 0$. By setting $a = \pi_\theta(s)$ in Assumption 2, we get that

$$g(\Theta - c) = \mathbb{E}\left[\tilde{\pi}^{-1}_{\Theta-c}(\pi_{\Theta-c}(s_1)) - (\Theta - c)\right] \geq \mathbb{E}\left[s_1 - c\right] - (\Theta - c) = 0\,.$$

By the same argument $g(\Theta + c) \leq 0$ and hence g must cross zero at least once in the interval $[\Theta - c, \Theta + c]$. $\square$

C.2 Properties of SOE in Equi-directional Problems

Proof of Proposition 1. Since v and $\tilde{v}$ are twice differentiable, the perceived optimal action $\tilde{\pi}_{\tilde{\theta}}(\cdot)$ and the optimal action $\pi_{\tilde{\theta}}(\cdot)$ are continuous functions. Because they are invertible, they are either strictly increasing or strictly decreasing. We focus on the strictly increasing case; if the functions are strictly decreasing we can re-normalize the signals. Fix a signal s. Because v and $\tilde{v}$ do not give rise to the same optimal action and are continuous, either $\pi_{\tilde{\theta}}(s) > \tilde{\pi}_{\tilde{\theta}}(s)$ everywhere or $\pi_{\tilde{\theta}}(s) < \tilde{\pi}_{\tilde{\theta}}(s)$ everywhere. We focus on the former case, the latter is analogous. Hence, for all s, the perceived signal $\tilde{\pi}^{-1}_{\tilde{\theta}}(\pi_{\tilde{\theta}}(s)) > s$.

We now argue that this implies that $\tilde{\theta} > \Theta$. The first-order condition corresponding to (1) yields

$$0 = \frac{\partial}{\partial z} \int \log f\left(\tilde{\pi}^{-1}_{\tilde{\theta}}\left(\pi_{\tilde{\theta}}(\Theta + \epsilon)\right) - z\right) f(\epsilon) d\epsilon \Big|_{z=\tilde{\theta}} = - \int \frac{f'\left(\tilde{\pi}^{-1}_{\tilde{\theta}}\left(\pi_{\tilde{\theta}}(\Theta + \epsilon)\right) - z\right)}{f\left(\tilde{\pi}^{-1}_{\tilde{\theta}}\left(\pi_{\tilde{\theta}}(\Theta + \epsilon)\right) - z\right)}\Bigg|_{z=\tilde{\theta}} f(\epsilon) d\epsilon$$

$$= - \int \frac{\left(\tilde{\pi}^{-1}_{\tilde{\theta}}\left(\pi_{\tilde{\theta}}(\Theta + \epsilon)\right) - \tilde{\theta}\right)}{\sigma^2_\epsilon} f(\epsilon) d\epsilon\,.$$

Since an agent who is correctly specified (i.e. for whom $\tilde{v} = v$) perceives the signal correctly (i.e. for such an agent $(\tilde{\pi}_{\tilde{\theta}})^{-1}(\pi_{\tilde{\theta}}(\Theta + \epsilon)) = \Theta + \epsilon$ for any $\tilde{\theta}$), $\tilde{\theta} = \Theta$ solves the analogous first order condition absent misspecification. Since for our misspecified agent $\tilde{\pi}^{-1}_{\tilde{\theta}}\left(\pi_{\tilde{\theta}}(\Theta + \epsilon)\right) > \Theta + \epsilon$, the first order condition can thus only hold when $\tilde{\theta} > \Theta$.

Because the perceived optimal action is increasing in s and θ, and because $\pi_{\tilde{\theta}}(s) > \tilde{\pi}_{\tilde{\theta}}(s)$, we have that $\pi_{\tilde{\theta}}(s) > \tilde{\pi}_{\tilde{\theta}}(s) \geq \tilde{\pi}_\Theta(s)$. As $\tilde{\pi}_\Theta(s)$ maximizes $\tilde{v}(a, s, \Theta)$ and $\tilde{v}$ is single-peaked, we thus have

$$\tilde{v}(\tilde{\pi}_{\Theta}(s), s, \Theta) > \tilde{v}(\tilde{\pi}_{\tilde{\theta}}(s), s, \Theta) \geq \tilde{v}(\pi_{\tilde{\theta}}(s), s, \Theta). \qquad \square$$

C.3 Present-Bias Model with Uncertain Benefits Specified in Section 3.1

We consider the setting from Section 3.1 in which the agent's decision utility and perceived decision utility is given by

$$\begin{aligned} v(a_t, s_t, \theta) &= u(a_t) + \phi_t a_t - \beta\kappa a_t \\ \tilde{v}(a_t, s_t, \theta) &= u(a_t) + \phi_t a_t - \tilde{\beta}\kappa a_t \\ \phi_t &= l\Theta + (1-l)s_t\,. \end{aligned} \tag{23}$$

Taking the first-order condition of (23) yields that the objectively and subjectively optimal policies $\pi_\mu(s), \tilde{\pi}_\mu(s)$ given a normally distributed belief $\mu = \mathcal{N}(\mathbb{E}_\mu[\Theta], \sigma_t^2)$ and signal s observed by the agent satisfy

$$\begin{aligned} \beta\kappa &= u'(\pi_\mu(s)) + \bar{\phi}(s,\mu) \\ \tilde{\beta}\kappa &= u'(\tilde{\pi}_\mu(s)) + \bar{\phi}(s,\mu)\,. \end{aligned} \tag{24}$$

where

$$\bar{\phi}(s,\mu) = (l\mathbb{E}_\mu\left[\Theta|s\right] + (1-l)s) = l\frac{\sigma_t^{-2}\mathbb{E}_\mu[\Theta] + \sigma_\epsilon^{-2}s}{\sigma_t^{-2} + \sigma_\epsilon^{-2}} + (1-l)s\,. \tag{25}$$

Rearranging yields that

$$\begin{aligned} \pi_\mu(s) &= (u')^{-1}(\beta\kappa - \bar{\phi}(s,\mu)) \\ \tilde{\pi}_\mu(s) &= (u')^{-1}(\tilde{\beta}\kappa - \bar{\phi}(s,\mu))\,. \end{aligned} \tag{26}$$

The following lemmas verify the convergence conditions of Proposition 8 for this model.

Lemma 5. *If the agent's utility satisfies* (23) *then Assumption 1 is satisfied.*

Proof. We first note that $\bar{\phi}(s,\mu)$ increases in s for every μ. As u is strictly concave, u' is strictly de- and $(u')^{-1}$ strictly increasing, and thus $\pi_\mu, \tilde{\pi}_\mu$ are strictly increasing. $\square$

Lemma 6. *If the agent's utility satisfies* (23) *then Assumption 2 is satisfied.*

Proof. It follows from (26) that $\pi_\mu(s)$ and $\tilde{\pi}_\mu(s)$ are continuous, and thus the function $s \mapsto \tilde{\pi}_{\mu_t}^{-1}(\pi_{\mu_t}(s))$ is also continuous.

Furthermore, by (24), we have that

$$\beta\kappa = u'(a) + \bar{\phi}(\pi_\mu^{-1}(a), \mu)$$
$$\tilde{\beta}\kappa = u'(a) + \bar{\phi}(\tilde{\pi}_\mu^{-1}(a), \mu)\,.$$

Subtracting the second from the first equation yields

$$(\beta - \tilde{\beta})\kappa = \bar{\phi}(\pi_\mu^{-1}(a), \mu) - \bar{\phi}(\tilde{\pi}_\mu^{-1}(a), \mu) = \left(\pi_\mu^{-1}(a) - \tilde{\pi}_\mu^{-1}(a)\right)\left[l\frac{\sigma_\epsilon^{-2}}{\sigma_t^{-2} + \sigma_\epsilon^{-2}} + (1-l)\right].$$

Taking absolute values implies that the distance between the true and perceived signal if the action equals a is bounded

$$\left|\pi_\mu^{-1}(a) - \tilde{\pi}_\mu^{-1}(a)\right| = \frac{|\beta - \tilde{\beta}|\kappa}{l\frac{\sigma_\epsilon^{-2}}{\sigma_t^{-2}+\sigma_\epsilon^{-2}} + (1-l)} \leq \frac{|\beta - \tilde{\beta}|\kappa}{1-l}\,. \qquad \square$$

Lemma 7. *If the agent's utility satisfies* (23) *then Assumption 3 is satisfied.*

Proof. Recall that for every normally distributed belief ν with posterior variance σ^2

$$\bar{\phi}(s, \nu) = \begin{cases} l\frac{\sigma^{-2}\mathbb{E}_\nu[\Theta] + \sigma_\epsilon^{-2}s}{\sigma^{-2}+\sigma_\epsilon^{-2}} + (1-l)s & \text{if } \sigma^2 > 0 \\ l\mathbb{E}_\nu[\Theta] + (1-l)s & \text{else} \end{cases}.$$

As $l < 1$, the function $\bar{\phi}$ is invertible in it's first argument and we denote it's inverse by $\bar{\phi}^{-1}(\cdot, \nu)$. By (26) we have that for any belief ν

$$\begin{aligned}
\tilde{\pi}_\nu^{-1}(\pi_\nu(s)) &= \tilde{\pi}_\nu^{-1}\left((u')^{-1}(\beta\kappa - \bar{\phi}(s,\nu))\right) = \bar{\phi}^{-1}(\tilde{\beta}\kappa - u'\left(\left[(u')^{-1}(\beta\kappa - \bar{\phi}(s,\nu)\right]\right), \nu) \\
&= \bar{\phi}^{-1}(\tilde{\beta}\kappa - \beta\kappa + \bar{\phi}(s,\nu), \nu) = \frac{\partial\bar{\phi}^{-1}(\bar{\phi}, \nu)}{\partial\bar{\phi}}(\tilde{\beta} - \beta)\kappa + \bar{\phi}^{-1}(\bar{\phi}(s,\nu), \nu) \\
&= \frac{\sigma_\epsilon^2 + \sigma^2}{(1-l)\sigma_\epsilon^2 + \sigma^2}(\tilde{\beta} - \beta)\kappa + s\,.
\end{aligned}$$

Here the second to last equation follows as $\bar{\phi}$ and thus also $\bar{\phi}^{-1}$ are linear, and the final follows from the fact that ${\partial\bar{\phi}^{-1}(\bar{\phi},\nu)}/{\partial\bar{\phi}} = \left({\partial\bar{\phi}(s,\nu)}/{\partial s}\right)^{-1}$. As σ_t is the variance of the belief μ_t, we get that

$$\begin{aligned}\left|\tilde{\pi}_{\mu_t}^{-1}\left(\pi_{\mu_t}(s)\right) - \tilde{\pi}_{\tilde{\theta}_t}^{-1}\left(\pi_{\tilde{\theta}_t}(s)\right)\right| &= \left|\frac{\sigma_\epsilon^2+\sigma_t^2}{(1-l)\sigma_\epsilon^2+\sigma_t^2} - \frac{\sigma_\epsilon^2}{(1-l)\sigma_\epsilon^2}\right|(\tilde{\beta}-\beta)\kappa \\ &\leq \frac{l\sigma_t^2}{(1-l)^2\sigma_\epsilon^2+(1-l)\sigma_t^2}(\tilde{\beta}-\beta)\kappa \leq \frac{l}{(1-l)^2\sigma_\epsilon^2}(\tilde{\beta}-\beta)\kappa\sigma_t^2 \\ &= \frac{l}{(1-l)^2\sigma_\epsilon^2}(\tilde{\beta}-\beta)\kappa\frac{1}{\sigma_1^{-2}+\sigma_\epsilon^{-2}(t-1)} \\ &\leq \frac{l}{(1-l)^2}(\tilde{\beta}-\beta)\kappa(t-1)^{-1} \leq \frac{3l(\tilde{\beta}-\beta)\kappa}{(1-l)^2}\, t^{-1}\,. \qquad \square\end{aligned}$$

Proof of Proposition 2. It follows from (4) and (5) that

$$\pi_{\tilde{\theta}}(s_t) = (u')^{-1}[\beta\kappa - l\tilde{\theta} - (1-l)s_t]$$

and

$$\tilde{\pi}_{\tilde{\theta}}(s_t) = (u')^{-1}[\tilde{\beta}\kappa - l\tilde{\theta} - (1-l)s_t].$$

Thus,

$$\tilde{\pi}_{\tilde{\theta}}^{-1}\left(\pi_{\tilde{\theta}}(s_t)\right) = \frac{\tilde{\beta}-\beta}{1-l}\kappa + s_t.$$

This implies that

$$\tilde{\theta} = \Theta + \frac{\tilde{\beta}-\beta}{1-l}\kappa$$

is an SOE that satisfies Condition (2) in Observation 1, allowing the agent to perfectly predict his behavior.

We now show that this is the unique SOE. By (1), any SOE needs to solve the following first-order condition:

$$\begin{aligned}0 &= \frac{\partial}{\partial z}\int \log f\left(\tilde{\pi}_{\tilde{\theta}}^{-1}\left(\pi_{\tilde{\theta}}(\Theta+\epsilon)\right) - z\right) f(\epsilon)d\epsilon\bigg|_{z=\tilde{\theta}} = -\int \frac{f'\left(\tilde{\pi}_{\tilde{\theta}}^{-1}\left(\pi_{\tilde{\theta}}(\Theta+\epsilon)\right)-z\right)}{f\left(\tilde{\pi}_{\tilde{\theta}}^{-1}\left(\pi_{\tilde{\theta}}(\Theta+\epsilon)\right)-z\right)}\Bigg|_{z=\tilde{\theta}} f(\epsilon)d\epsilon \\ &= -\int \frac{\left(\tilde{\pi}_{\tilde{\theta}}^{-1}\left(\pi_{\tilde{\theta}}(\Theta+\epsilon)\right)-\tilde{\theta}\right)}{\sigma_\epsilon^2} f(\epsilon)d\epsilon = -\frac{1}{\sigma_\epsilon^2}\left[\Theta + \frac{\tilde{\beta}-\beta}{1-l}\kappa - \tilde{\theta} + \underbrace{\int \epsilon f(\epsilon)d\epsilon}_{=0}\right].\end{aligned}$$

Thus, the SOE is unique.

Substituting into (4) yields the corresponding consumption choice of the agent. Finally — by Lemmas 5, 6, and 7 — we can apply Proposition 8 to conclude that beliefs converge with probability 1 to the unique SOE. □

C.4 Present Bias Model with Uncertain Harm Specified in Section 3.3

We consider the setting from Section 3.3 in which the agent's objective and subjective utility is given by

$$\begin{aligned} v(a_t, s_t, \theta) &= u(a_t) - \beta e^{\phi_t} a_t \\ \tilde{v}(a_t, s_t, \theta) &= u(a_t) - \tilde{\beta} e^{\phi_t} a_t \\ \phi_t &= l\Theta + (1-l)s_t \,. \end{aligned} \tag{27}$$

Taking the first-order condition of (27) yields that the objectively and subjectively optimal action $\pi_\mu(s), \tilde{\pi}_\mu(s)$ given a belief μ and signal s taken by the agent satisfy

$$\begin{aligned} \beta \, \mathbb{E}_\mu \left[e^{\phi(s,\mu)} | s \right] &= u'(\pi_\mu(s)) \\ \tilde{\beta} \, \mathbb{E}_\mu \left[e^{\phi(s,\mu)} | s \right] &= u'(\tilde{\pi}_\mu(s)) \,. \end{aligned} \tag{28}$$

where the posterior given prior μ about Θ and signal s is normally distributed and has mean

$$\bar{\phi}(s, \mu) = (l \mathbb{E}_\mu \left[\theta | s \right] + (1-l)s) = l \frac{\sigma_t^{-2} \mathbb{E}_\mu[\theta] + \sigma_\epsilon^{-2} s}{\sigma_t^{-2} + \sigma_\epsilon^{-2}} + (1-l)s$$

and variance

$$\mathrm{Var}[\phi(s, \mu)] = l^2 \sigma_{t+1}^2 \,.$$

Rearranging (28), using that $\phi(s, \mu)$ is normally distributed to take the expectation, yields that

$$\begin{aligned} \pi_\mu(s) &= (u')^{-1} \left(\beta e^{\bar{\phi}(s,\mu) + \frac{1}{2} \mathrm{Var}[\phi(s,\mu)]} \right) \\ \tilde{\pi}_\mu(s) &= (u')^{-1} \left(\tilde{\beta} e^{\bar{\phi}(s,\mu) + \frac{1}{2} \mathrm{Var}[\phi(s,\mu)]} \right) \,. \end{aligned} \tag{29}$$

Lemma 8. *If the agent's utility satisfies* (27) *then Assumption 1 is satisfied.*

Proof. We first note that $\bar{\phi}(s, \mu)$ increases in s for every μ and that $\mathrm{Var}[\phi(s, \mu)]$ is independent of s. Thus $e^{\bar{\phi}(s,\mu) + \frac{1}{2} \mathrm{Var}[\phi(s,\mu)]}$ is increasing in s. As u is strictly concave, u' is strictly de- and $(u')^{-1}$ strictly increasing, and thus $\pi_\mu, \tilde{\pi}_\mu$ are strictly increasing. □

Lemma 9. *If the agent's utility satisfies* (27) *then Assumption 2 is satisfied.*

Proof. It follows from (29) that $\pi_\mu(s)$ and $\tilde{\pi}_\mu(s)$ are continuous, and thus the function $s \mapsto \tilde{\pi}_{\mu_t}^{-1}(\pi_{\mu_t}(s))$ is also continuous. Furthermore, by (29) and the fact that $\mathrm{Var}[\phi(s,\mu)]$ is independent of s, we have that $\beta e^{\bar{\phi}(\pi_\mu^{-1}(a),\mu)} = \tilde{\beta} e^{\bar{\phi}(\tilde{\pi}_\mu^{-1}(a),\mu)}$. Taking logarithm and rewriting yields

$$(\ln \tilde{\beta} - \ln \beta) = \bar{\phi}(\pi_\mu^{-1}(a),\mu) - \bar{\phi}(\tilde{\pi}_\mu^{-1}(a),\mu) = \left(\pi_\mu^{-1}(a) - \tilde{\pi}_\mu^{-1}(a)\right)\left[l\frac{\sigma_\epsilon^{-2}}{\sigma_t^{-2}+\sigma_\epsilon^{-2}} + (1-l)\right].$$

Taking absolute values implies that

$$\left|\pi_\mu^{-1}(a) - \tilde{\pi}_\mu^{-1}(a)\right| = \frac{|\ln\tilde{\beta} - \ln\beta|}{l\frac{\sigma_\epsilon^{-2}}{\sigma_t^{-2}+\sigma_\epsilon^{-2}} + (1-l)} \leq \frac{|\ln\tilde{\beta} - \ln\beta|}{1-l}. \qquad \square$$

Lemma 10. *If the agent's utility satisfies* (27) *then Assumption 3 is satisfied.*

Proof. Recall that for every normally distributed belief ν with variance σ^2

$$\bar{\phi}(s,\nu) = \begin{cases} l\frac{\sigma^{-2}\mathbb{E}_\nu[\Theta]+\sigma_\epsilon^{-2}s}{\sigma^{-2}+\sigma_\epsilon^{-2}} + (1-l)s & \text{if } \sigma^2 > 0 \\ l\mathbb{E}_\nu[\Theta] + (1-l)s & \text{else} \end{cases}.$$

As $l < 1$, the function $\bar{\phi}$ is invertible in it's first argument and we denote it's inverse by $\bar{\phi}^{-1}(\cdot,\nu)$. By (29) we have that for any belief ν

$$\begin{aligned} \tilde{\pi}_\nu^{-1}(\pi_\nu(s)) &= \tilde{\pi}_\nu^{-1}\left[(u')^{-1}\left(\beta e^{\bar{\phi}(s,\nu)+\frac{1}{2}\mathrm{Var}[\phi(s,\mu)]}\right)\right] \\ &= \bar{\phi}^{-1}\left(\ln\left[\beta e^{\bar{\phi}(s,\nu)}\right] - \ln\tilde{\beta}, \nu\right) \\ &= \bar{\phi}^{-1}(\ln\beta - \ln\tilde{\beta} + \bar{\phi}(s,\nu),\nu) \\ &= \frac{\partial\bar{\phi}^{-1}(\bar{\phi},\nu)}{\partial\bar{\phi}}(\ln\beta - \ln\tilde{\beta}) + \bar{\phi}^{-1}(\bar{\phi}(s,\nu),\nu) \\ &= \frac{\sigma_\epsilon^2+\sigma^2}{(1-l)\sigma_\epsilon^2+\sigma^2}(\ln\beta - \ln\tilde{\beta}) + s. \end{aligned} \tag{30}$$

Here the second line exploits that (29) implies that $\bar{\phi}(\tilde{\pi}_\nu^{-1}(a),\nu) = \ln[u'(a)] - \ln\tilde{\beta} - \frac{1}{2}\mathrm{Var}[\phi(\tilde{\pi}_\nu^{-1}(a),\mu)]$; the second to last equation follows as $\bar{\phi}$ and thus also $\bar{\phi}^{-1}$ are linear; and the final follows from the fact that ${\partial\bar{\phi}^{-1}(\bar{\phi},\nu)}/{\partial\bar{\phi}} = \left({\partial\bar{\phi}(s,\nu)}/{\partial s}\right)^{-1}$. As σ_t is the variance of the belief μ_t, we get that

$$\begin{aligned}
\left|\tilde{\pi}_{\mu_t}^{-1}\left(\pi_{\mu_t}(s)\right) - \tilde{\pi}_{\tilde{\theta}_t}^{-1}\left(\pi_{\tilde{\theta}_t}(s)\right)\right| &= \left|\frac{\sigma_\epsilon^2 + \sigma_t^2}{(1-l)\sigma_\epsilon^2 + \sigma_t^2} - \frac{\sigma_\epsilon^2}{(1-l)\sigma_\epsilon^2}\right|(\ln\tilde{\beta} - \ln\beta) \\
&\leq \frac{l\sigma_t^2}{(1-l)^2\sigma_\epsilon^2 + (1-l)\sigma_t^2}(\ln\tilde{\beta} - \ln\beta) \\
&\leq \frac{l}{(1-l)^2\sigma_\epsilon^2}(\ln\tilde{\beta} - \ln\beta)\sigma_t^2 \\
&= \frac{l}{(1-l)^2\sigma_\epsilon^2}(\ln\tilde{\beta} - \ln\beta)\frac{1}{\sigma_1^{-2} + \sigma_\epsilon^{-2}(t-1)} \\
&\leq \frac{l}{(1-l)^2}(\ln\tilde{\beta} - \ln\beta)(t-1)^{-1} \\
&\leq \frac{3l(\ln\tilde{\beta} - \ln\beta)}{(1-l)^2}\; t^{-1}\,.
\end{aligned}$$

□

Proof of Proposition 3. By Equation 30,

$$\tilde{\pi}_{\tilde{\theta}}^{-1}\left(\pi_{\tilde{\theta}}(\Theta + \epsilon)\right) = \Theta - \frac{\ln\tilde{\beta} - \ln\beta}{1-l} + \epsilon.$$

Since by (1), any SOE needs to solve the following first-order condition, the above fact yields

$$0 = \frac{\partial}{\partial z}\int \log f\left(\tilde{\pi}_{\tilde{\theta}}^{-1}\left(\pi_{\tilde{\theta}}(\Theta+\epsilon)\right) - z\right) f(\epsilon)d\epsilon\bigg|_{z=\tilde{\theta}} = -\int \frac{f'\left(\tilde{\pi}_{\tilde{\theta}}^{-1}\left(\pi_{\tilde{\theta}}(\Theta+\epsilon)\right) - z\right)}{f\left(\tilde{\pi}_{\tilde{\theta}}^{-1}\left(\pi_{\tilde{\theta}}(\Theta+\epsilon)\right) - z\right)}\Bigg|_{z=\tilde{\theta}} f(\epsilon)d\epsilon$$

$$= -\int \frac{\left(\tilde{\pi}_{\tilde{\theta}}^{-1}\left(\pi_{\tilde{\theta}}(\Theta+\epsilon)\right) - \tilde{\theta}\right)}{\sigma_\epsilon^2} f(\epsilon)d\epsilon = -\frac{1}{\sigma_\epsilon^2}\left[\Theta - \frac{\ln\tilde{\beta} - \ln\beta}{1-l} - \tilde{\theta} + \underbrace{\int \epsilon f(\epsilon)d\epsilon}_{=0}\right].$$

Thus there is a unique SOE $\tilde{\theta}$ given by

$$\tilde{\theta} = \Theta - \frac{\ln\tilde{\beta} - \ln\beta}{1-l},$$

which satisfies Condition (2). Furthermore, Lemmas 8, 9, and 10 verify that Assumptions 1, 2, and 3 hold. Thus, Proposition 8 implies that beliefs converge with probability one to the unique SOE above. □

C.5 Proofs for Subsections 3.5 and 3.6

Proof of Proposition 4. Recall that we denote by $\mathbb{E}_{\mu_t}[\cdot|s_t]$ the expectation of the agent in period t before taking an action. Given belief μ_t and signal s_t, an agent's action at time t maximizes

$$\max_{a_t} \ \mathbb{E}_{\mu_t}\left[u(a_t) + (l\theta_t + (1-l)s_t) - \beta\kappa a_t \mid s_t\right] . \tag{31}$$

(i) Consider an inexperienced agent, i.e. an agent at time $t=1$, whose prior mean equals the state (i.e., $\theta_1 = \Theta$). We prove that the outside observer correctly infers the agent's time-inconsistency and sophistication, i.e. estimates the pair $\beta, \tilde{\beta}$ correctly from observing the distribution of actions. Throughout, we denote by $\hat{\beta}, \hat{\tilde{\beta}}, \hat{\Theta}$, etc the estimates formed by the outside observer of the respective quantities.

We first establish that an outsider will infer $\hat{\beta}\kappa - \hat{\Theta} = \beta\kappa - \Theta$ from the distribution of the agent's action at time $t=1$. Using (18), the fact that $\theta_1 = \Theta$, and that the agent choose the action after having observed s_1, the agent's policy function hence is

$$\begin{aligned}
\pi_{\mu_1}(s_1) &= (u')^{-1}\left[\beta\kappa - l\frac{\sigma_1^{-2}}{\sigma_1^{-2} + \sigma_\epsilon^{-2}t}\Theta - \left(1 - l\frac{\sigma_1^{-2}}{\sigma_1^{-2} + \sigma_\epsilon^{-2}t}\right)s_t\right] \\
&= (u')^{-1}\left[\beta\kappa - \Theta - \left(1 - l\frac{\sigma_1^{-2}}{\sigma_1^{-2} + \sigma_\epsilon^{-2}t}\right)\epsilon_1\right] .
\end{aligned}$$

For brevity, define

$$l_1 \equiv l\frac{\sigma_1^{-2}}{\sigma_1^{-2} + \sigma_\epsilon^{-2}t}.$$

An inexperienced agent's action as a function of the signal $s_1 = \Theta + \epsilon_1$ equals

$$\pi_{\mu_1}(s_1) = (u')^{-1}\left[\beta\kappa - \Theta - (1 - l_1)\,\epsilon_1\right] .$$

Since the observer knows the functional form of u and supposes that the inexperienced agents also knows Θ, she incorrectly believes the inexperienced agent uses the policy function

$$\pi_{\hat{\Theta}}(s_1) = (u')^{-1}\left[\hat{\beta}\kappa - \hat{\Theta} - (1-l)\,\epsilon_1\right] .$$

As a result, she infers

$$\hat{\beta}\kappa - \hat{\Theta} - (1-l)\epsilon_1 = \beta\kappa - \Theta - (1-l_1)\epsilon_1$$

from the observed actions. Note that the right and left-hand side above are normally distributed. Using that the noise on average is zero, the observer can perfectly explain the inexperienced agent's action with the correct $\beta\kappa - \Theta$ and an estimated variance of the signal of $\hat{\sigma}_\epsilon^2 = \sigma_\epsilon^2(1-l_1)^2/(1-l)^2$. Denote by $\hat{f}(\epsilon)$ the density of the normal distribution with mean zero and variance $\hat{\sigma}_\epsilon^2$.

We next show that the observer infers $\hat{\tilde{\beta}} - \hat{\beta} = \tilde{\beta} - \beta$ from the agents' expectation of their own consumption (which by assumption of the proposition is known). The inexperienced agents expect a mean consumption of

$$\int (u')^{-1}[\tilde{\beta}\kappa - \Theta - (1-l_1)\epsilon] f(\epsilon) d\epsilon,$$

which, as $(1-l_1)\epsilon$ has the same distribution under f as $(1-l)\epsilon$ has under $\hat{f}$, is the same as

$$\int (u')^{-1}[\hat{\tilde{\beta}}\kappa - \hat{\Theta} - (1-l)\epsilon] \hat{f}(\epsilon) d\epsilon\,.$$

Hence, knowing the functional form of u and believing that the density of ϵ is $\hat{f}(\epsilon)$, the observer correctly infers $\kappa\tilde{\beta} - \Theta$ from the inexperienced agents' observed predicted mean consumption. Subtracting his previously obtained estimate of $\beta\kappa - \Theta$, the observer learns $(\tilde{\beta} - \beta)\kappa$ and since she knows κ, she correctly infers $\hat{\tilde{\beta}} - \hat{\beta} = \tilde{\beta} - \beta$.

Finally, we show that the observer infers $\hat{\tilde{\beta}} = \tilde{\beta}$ from the agents' marginal value of decreasing consumption without discounting (which is observable by the assumptions of the proposition). Intuitively, the agent's perceived benefit of marginally decreasing self 1's action only depends on the perceived conflict of interest between agent's selves, which is captured by $(1-\tilde{\beta})\kappa$. Thus, knowing κ, the observer correctly infers $\tilde{\beta}$ from the agent's stated benefit of a marginal reduction of the action. Formally, the marginal value of decreasing the subjective choice of $a_1 \equiv \tilde{\pi}_{\mu_1}(\Theta + \epsilon)$ to the long-run self is

$$-\int \left(\underbrace{u'(\tilde{\pi}_{\mu_1}(\Theta+\epsilon)) + \Theta + (1-l_1)\epsilon - \tilde{\beta}\kappa}_{=0} - (1-\tilde{\beta})\kappa \right) f(\epsilon) d\epsilon = (1-\tilde{\beta})\kappa,$$

while the observer believes that agent's belief about his action is given by $a_1 \equiv \tilde{\pi}_{\Theta}(\Theta + \epsilon)$ and the subjective marginal value of decreasing the action is

$$-\int \left(\underbrace{u'(\tilde{\pi}_{\hat{\Theta}}(\hat{\Theta}+\epsilon)) + \hat{\Theta} + (1-l)\epsilon - \hat{\tilde{\beta}}\kappa}_{=0} - (1-\hat{\tilde{\beta}})\kappa \right) \hat{f}(\epsilon) d\epsilon = (1-\hat{\tilde{\beta}})\kappa.$$

Given the observer's inference of $\tilde{\beta}-\beta$ above, she thus also correctly infers β. Furthermore, with these parameters and the estimated density $\hat{f}(\epsilon)$ of the error distribution, the observer can perfectly explain all observed choices of inexperienced agents.

(ii) We next consider the experienced agent, i.e. the limit of the observer's inferences when $t \to \infty$ and behavior converged to an SOE.

We first show that from observing the agent's actions the observer comes to believe that $\hat{\beta}\kappa - \hat{\Theta} = \tilde{\beta}\kappa - \tilde{\theta}$. By Observation 1, for the realized fundamental Θ, an experienced agent's long-run SOE belief $\tilde{\theta}$ for every ϵ solves

$$\pi_{\tilde{\theta}}(\Theta+\epsilon) = \tilde{\pi}_{\tilde{\theta}}(\tilde{\theta}+\epsilon).$$

Given the above SOE, the experienced agent's policy function is $\pi_{\tilde{\theta}}(\Theta+\epsilon)$, so that

$$\pi_{\tilde{\theta}}(\Theta+\epsilon) = \tilde{\pi}_{\tilde{\theta}}(\tilde{\theta}+\epsilon) = (u')^{-1}[\tilde{\beta}\kappa - \tilde{\theta} - (1-l)\epsilon)] \,. \tag{32}$$

Under the incorrect assumption that the agent knows the true fundamental (but allowing it to differ from the actually true realization), the observer believes the agent uses a policy function $\pi_{\hat{\Theta}}(\hat{\Theta}+\epsilon)$ for some to her unknown fundamental $\hat{\Theta}$, that is she thinks the agent's policy function satisfies

$$\pi_{\hat{\Theta}}(\hat{\Theta}+\epsilon) = (u')^{-1}[\hat{\beta}\kappa - \hat{\Theta} - (1-l)\epsilon)] \,. \tag{33}$$

Comparing the right-hand sides of (32) and (33) yields that to explain the average action, the observer must conclude that $\hat{\beta}\kappa - \hat{\Theta} = \tilde{\beta}\kappa - \tilde{\theta}$. Furthermore, to explain the variance of actions, she concludes that ϵ is distributed according to the true distribution $f(\epsilon)$. Given this, she can perfectly explain the distribution of the experienced agent's observed actions.

We next establish that the outside observer infers from the agent's reported expected subjective mean consumption that $\hat{\beta} - \hat{\hat{\beta}} = 0$. The experienced agent expects a mean consumption of

$$\int (u')^{-1}[\tilde{\beta}\kappa - \tilde{\theta} - (1-l)\epsilon] f(\epsilon) d\epsilon,$$

while the observer incorrectly thinks the agent expects a mean consumption of

$$\int (u')^{-1}[\hat{\hat{\beta}}\kappa - \hat{\Theta} - (1-l)\epsilon] f(\epsilon) d\epsilon\,.$$

Thus the observer, supposing that the experienced agent knows the true fundamental, misinfers that $\hat{\hat{\beta}}\kappa - \hat{\Theta} = \tilde{\beta}\kappa - \tilde{\theta}$. Combining this with her conclusion from above that $\hat{\beta}\kappa - \hat{\Theta} = \tilde{\beta}\kappa - \tilde{\theta}$, yields $\hat{\hat{\beta}} = \hat{\beta}$.

Finally, we establish that the outside observer learns $\hat{\tilde{\beta}} = \tilde{\beta}$. The experienced long-run self's perceived marginal value of decreasing the perceived choice of $a_t \equiv \tilde{\pi}_{\tilde{\theta}}(\tilde{\theta} + \epsilon)$ is

$$-\int \left\{ \underbrace{u'(\tilde{\pi}_{\tilde{\theta}}(\tilde{\theta}+\epsilon)) + (l\tilde{\theta} + (1-l)(\tilde{\theta}+\epsilon)) - \tilde{\beta}\kappa}_{=0} - (1-\tilde{\beta})\kappa \right\} f(\epsilon) d\epsilon = (1-\tilde{\beta})\kappa,$$

while the observer thinks $a_t \equiv \tilde{\pi}_{\hat{\Theta}}(\hat{\Theta} + \epsilon)$ and thus the perceived marginal value of decreasing it is

$$-\int \left\{ \underbrace{u'(a_t) + (l\hat{\Theta} + (1-l)(\hat{\Theta}+\epsilon)) - \hat{\hat{\beta}}\kappa}_{=0} - (1-\hat{\tilde{\beta}})\kappa \right\} f(\epsilon) d\epsilon = (1-\hat{\tilde{\beta}})\kappa\,.$$

Hence, the observer correctly learns $\hat{\tilde{\beta}} = \tilde{\beta}$ from the experienced agent's expected marginal value of decreasing consumption. Note that the correct inference of $\tilde{\beta}$ together with the misinference that $\hat{\beta} = \hat{\hat{\beta}}$ allows the observer to perfectly explain all observed choices of the experienced agent. □

Proof of Proposition 5. Recall that $\tilde{\theta}_t = \mathbb{E}_{\mu_t}[\theta]$ denotes the posterior mean (or log-likelihood maximizing) belief regarding the fundamental, and σ_t^2 the corresponding variance, prior to observing the signal s_t. Given our normal-normal structure, μ_t has precision $\sigma_t^{-2} = \sigma_1^{-2} + \sigma_\epsilon^{-2}(t-1)$, where σ_ϵ^{-2} is the precision of ϵ_t.

Part (i): For $t \geq 2$, define

$$\omega_t \equiv \frac{\frac{\sigma_t^2}{\sigma_\epsilon^2}}{1 - l\left(1 - \frac{\sigma_t^2}{\sigma_\epsilon^2}\right)},$$

and for future reference note that $\omega_t \in (0,1)$ since $\sigma_t^2 < \sigma_\epsilon^2$. We next argue that $\omega_t > (1-\omega_t)\omega_{t-1}$. This is equivalent to

$$\sigma_t^2\sigma_\epsilon^2(1-l) + \sigma_{t-1}^2\sigma_t^2 > \sigma_{t-1}^2\sigma_\epsilon^2(1-l).$$

This inequality is linear in l and holds for $l = 1$. Furthermore, because $\sigma_\epsilon^2 + \sigma_{t-1}^2 = \sigma_\epsilon^2\sigma_{t-1}^2\sigma_t^{-2}$ the left-hand side above equals the right-hand side when $l = 0$. Thus, the inequality holds for all $l \in (0,1)$.

Upon observing the signal s_t, normal updating implies that the agent's posterior mean belief becomes

$$\mathbb{E}_{\mu_t}[\Theta|s_t] = \tilde{\theta}_t + \frac{\sigma_{t+1}^2}{\sigma_\epsilon^2}(s_t - \tilde{\theta}_t).$$

Denote the chosen action as a function of s_t and $\tilde{\theta}_t$ by $a_t(s_t,\tilde{\theta}_t)$. Rewriting the first-order-condition for the maximization of v using that the expectation of ϕ_t given s_t and $\tilde{\theta}_t$ equals $l\mathbb{E}_{\mu_t}[\Theta|s_t]+(1-l)s_t$ yields

$$a_t(s_t,\tilde{\theta}_t) = l\left(1 - \frac{\sigma_{t+1}^2}{\sigma_\epsilon^2}\right)\tilde{\theta}_t + \left[1 - l\left(1 - \frac{\sigma_{t+1}^2}{\sigma_\epsilon^2}\right)\right]s_t - \beta\kappa. \tag{34}$$

Hence, the true signal s_t as a function of the agent's chosen action a_t can be inferred using

$$s_t = \frac{a_t + \beta\kappa - l\left(1 - \frac{\sigma_{t+1}^2}{\sigma_\epsilon^2}\right)\tilde{\theta}_t}{1 - l\left(1 - \frac{\sigma_{t+1}^2}{\sigma_\epsilon^2}\right)}.$$

But because the agent when looking back thinks she used $\tilde{\beta}$, she infers from a_t that she observed the signal

$$\tilde{s}_t = \frac{a_t + \tilde{\beta}\kappa - l\left(1 - \frac{\sigma_{t+1}^2}{\sigma_\epsilon^2}\right)\tilde{\theta}_t}{1 - l\left(1 - \frac{\sigma_{t+1}^2}{\sigma_\epsilon^2}\right)}. \tag{35}$$

Since when looking back at his action the agent updates his beliefs based on the perceived signal

$\tilde{s}_t$, normal-normal updating implies that

$$\begin{aligned}\tilde{\theta}_t &= \tilde{\theta}_{t-1} + \frac{\sigma_t^2}{\sigma_\epsilon^2}(\tilde{s}_{t-1} - \tilde{\theta}_{t-1}) = \tilde{\theta}_{t-1} + \frac{\sigma_t^2}{\sigma_\epsilon^2}\frac{1}{1 - l\left(1 - \frac{\sigma_t^2}{\sigma_\epsilon^2}\right)}\left(a_{t-1} + \tilde{\beta}\kappa - \tilde{\theta}_{t-1}\right) \\ &= (1-\omega_t)\tilde{\theta}_{t-1} + \omega_t(a_{t-1} + \tilde{\beta}\kappa)\,.\end{aligned} \tag{36}$$

Now using (36) to express $\tilde{\theta}_t$ as a function of past actions yields

$$\tilde{\theta}_t = \sum_{j=1}^{t-1} a_{t-j}\left(\omega_{t-j+1}\prod_{k=t-j+2}^{t}(1-\omega_k)\right) + \tilde{\beta}\kappa\left(\sum_{j=1}^{t-1}\omega_{t-j+1}\prod_{k=t-j+2}^{t}(1-\omega_k)\right) + \tilde{\theta}_1\prod_{j=2}^{t}(1-\omega_j)\,.$$

Using this and (34), one has for $\tau < t$ that

$$\frac{\partial a_t(s_t, a^{t-1})}{\partial a_\tau} = \frac{\partial a_t(s_t, \tilde{\theta}_t)}{\partial \tilde{\theta}_t}\frac{\partial \tilde{\theta}_t}{\partial a_\tau} = l\left(1 - \frac{\sigma_{t+1}^2}{\sigma_\epsilon^2}\right)\omega_{\tau+1}\left(\prod_{k=\tau+2}^{t}(1-\omega_k)\right) > 0.$$

Furthermore because $\omega_{\tau+2} > \omega_{\tau+1}(1-\omega_{\tau+2})$, it follows that

$$\frac{\partial a_t(s_t, a^{t-1})}{\partial a_{\tau+1}} > \frac{\partial a_t(s_t, a^{t-1})}{\partial a_\tau}.$$

Part II. Since the difference between the perceived and objective signal

$$\tilde{s}_t - s_t = \frac{(\tilde{\beta} - \beta)\kappa}{1 - l\left(1 - \frac{\sigma_{t+1}^2}{\sigma_\epsilon^2}\right)} > 0, \tag{37}$$

one has $\tilde{\theta}_t = \mathbb{E}_{\mu_t}[\Theta] > \mathbb{E}[\Theta|s^t]$. Furthermore, since the difference $\tilde{s}_t - s_t$ is increasing in κ so does the agent's overestimation of the fundamental. Since $a_t(s_t, \tilde{\theta}_t)$ is increasing in $\tilde{\theta}_t$, it follows that $a_t - a_t^* > 0$ and strictly increasing in κ.

Part III. Since from an ex ante perspective $\mathbb{E}_{\mu_1}[s_t] = \mathbb{E}_{\mu_1}[\Theta]$, (34) implies that

$$\mathbb{E}_{\mu_1}[a_t] = l\left(1 - \frac{\sigma_{t+1}^2}{\sigma_\epsilon^2}\right)\mathbb{E}_{\mu_1}[\tilde{\theta}_t] + \left[1 - l\left(1 - \frac{\sigma_{t+1}^2}{\sigma_\epsilon^2}\right)\right]\mathbb{E}_{\mu_1}[\Theta] - \beta\kappa. \tag{38}$$

Using (36) and (37) to express the mean posterior belief as function of the prior mean and the objective signals yields

$$\tilde{\theta}_t = \sum_{k=2}^{t} \prod_{\tau=k+1}^{t} \left(1 - \frac{\sigma_\tau^2}{\sigma_\epsilon^2}\right) \frac{\sigma_k^2}{\sigma_\epsilon^2} s_k + \prod_{k=2}^{t} \left(1 - \frac{\sigma_k^2}{\sigma_\epsilon^2}\right) \tilde{\theta}_1 + \left[\sum_{k=2}^{t} \prod_{\tau=k+1}^{t} \left(1 - \frac{\sigma_\tau^2}{\sigma_\epsilon^2}\right) \omega_k\right] (\tilde{\beta} - \beta)\kappa.$$

Taking the ex ante expectation using that $\mathbb{E}_{\mu_1}[s_k] = \mathbb{E}_{\mu_1}[\Theta]$ and $\tilde{\theta}_1 = \mathbb{E}_{\mu_1}[\Theta]$ gives

$$\mathbb{E}_{\mu_1}[\tilde{\theta}_t] = \mathbb{E}_{\mu_1}[\Theta] + \left[\sum_{k=2}^{t} \prod_{\tau=k+1}^{t} \left(1 - \frac{\sigma_\tau^2}{\sigma_\epsilon^2}\right) \omega_k\right] (\tilde{\beta} - \beta)\kappa.$$

We now argue that the term in square brackets, and hence $\mathbb{E}_{\mu_1}[\tilde{\theta}_t]$, is increasing in t. One has

$$\begin{aligned}
\left[\sum_{k=2}^{t} \prod_{\tau=k+1}^{t} \left(1 - \frac{\sigma_\tau^2}{\sigma_\epsilon^2}\right) \omega_k\right] - \left[\sum_{k=2}^{t-1} \prod_{\tau=k+1}^{t-1} \left(1 - \frac{\sigma_\tau^2}{\sigma_\epsilon^2}\right) \omega_k\right] &= \omega_t - \frac{\sigma_t^2}{\sigma_\epsilon^2} \sum_{k=2}^{t-1} \prod_{\tau=k+1}^{t-1} \left(1 - \frac{\sigma_\tau^2}{\sigma_\epsilon^2}\right) \omega_k \\
&\geq \omega_t - \frac{\sigma_t^2}{\sigma_\epsilon^2} \omega_2 \left(1 - \frac{\sigma_{t-1}^2}{\sigma_\epsilon^2}\right) \frac{1 - \left(1 - \frac{\sigma_{t-1}^2}{\sigma_\epsilon^2}\right)^4}{1 - \left(1 - \frac{\sigma_{t-1}^2}{\sigma_\epsilon^2}\right)} \\
&\geq \omega_t - \frac{\sigma_t^2}{\sigma_{t-1}^2} \omega_2 \left(1 - \frac{\sigma_{t-1}^2}{\sigma_\epsilon^2}\right) \\
&\geq \frac{\frac{\sigma_t^2}{\sigma_\epsilon^2}}{1 - l\left(1 - \frac{\sigma_t^2}{\sigma_\epsilon^2}\right)} - \frac{\frac{\sigma_t^2}{\sigma_\epsilon^2} \frac{\sigma_2^2}{\sigma_{t-1}^2} \left(1 - \frac{\sigma_{t-1}^2}{\sigma_\epsilon^2}\right)}{1 - l\left(1 - \frac{\sigma_t^2}{\sigma_\epsilon^2}\right)} \\
&= \omega_t \left[1 - \frac{\sigma_2^2}{\sigma_{t-1}^2} \left(1 - \frac{\sigma_{t-1}^2}{\sigma_\epsilon^2}\right)\right] \\
&= \omega_t \left[1 - \frac{\sigma_\epsilon^2 - \sigma_{t-1}}{\sigma_\epsilon^2 + \sigma_{t-1}}\right] \\
&> 0,
\end{aligned}$$

where the first inequality uses the facts that σ_t and ω_t are decreasing in t, the second follows from replacing the numerator of the final ratio by 1, the third from the fact that $\sigma_{t-1} < \sigma_t$, and the last equality uses that $\sigma_\epsilon^2 + \sigma_{t-1}^2 = \sigma_\epsilon^2 \sigma_{t-1}^2 \sigma_t^{-2}$. We conclude that $\mathbb{E}_{\mu_1}[\tilde{\theta}_t]$ is increasing in t, and also that $\mathbb{E}_{\mu_1}[\tilde{\theta}_t] > \mathbb{E}_{\mu_1}[\theta]$ for all $t > 1$. Since also $\sigma_{t+1}^2 < \sigma_t^2$, (38) implies that $\mathbb{E}_{\mu_1}[a_t]$ is strictly increasing in t. Furthermore, since $\mathbb{E}_{\mu_1}[\tilde{\theta}_t] - \mathbb{E}_{\mu_1}[\tilde{\theta}_{t-1}]$ is strictly increasing in κ, (38) implies that so is $\mathbb{E}_{\mu_1}[a_t] - \mathbb{E}_{\mu_1}[a_{t-1}]$.

Finally, by the law of iterated expectations (which applies to the subjective beliefs of the agent as he believes to be correctly specified), we get that the agent's ex ante belief $\tilde{\mathbb{E}}_{\mu_1}[a_t] = \mathbb{E}_{\mu_1}[\Theta] - \tilde{\beta}k$, which is constant. $\square$

Proof of Corollary 1. By Equation (34), the agent reacts to the new information according to $\partial a_t/\partial\kappa = -\beta$. But plugging the SOE belief in Equation (7) into the formula for consumption a_t in Equation (4), the agent's long-response is $\partial a_t/\partial\kappa = -\beta + l(\tilde{\beta} - \beta)/(1-l)$. $\square$

C.6 Verifying Convergence Conditions for Model of Section 4

We consider the setting from Section 4 in which the agent's decision utility and perceived decision utility is given by

$$\begin{aligned} v(a_t, s_t, \theta) &= (\phi_t - b)a_t - \frac{a_t^2}{2} \\ \tilde{v}(a_t, s_t, \theta) &= \phi_t a_t - \frac{a_t^2}{2} \\ \phi_t &= l\Theta + (1-l)s_t\,. \end{aligned} \tag{39}$$

Taking the first-order condition of (39) yields that the objectively and subjectively optimal policies $\pi_\mu(s), \tilde{\pi}_\mu(s)$ given a normally distributed belief $\mu = \mathcal{N}(\mathbb{E}_\mu[\Theta], \sigma_t^2)$ and signal s observed by the agent satisfy

$$\pi_\mu(s) = \bar{\phi}(s,\mu) - b \qquad \text{and} \qquad \tilde{\pi}_\mu(s) = \bar{\phi}(s,\mu)\,, \tag{40}$$

where

$$\bar{\phi}(s,\mu) = (l\mathbb{E}_\mu[\Theta|s] + (1-l)s) = l\frac{\sigma_t^{-2}\mathbb{E}_\mu[\Theta] + \sigma_\epsilon^{-2}s}{\sigma_t^{-2} + \sigma_\epsilon^{-2}} + (1-l)s\,.$$

Lemma 11. *If the agent's utility satisfies* (39) *then Assumption 1 is satisfied.*

Proof of Lemma 11. The result follows immediately from the fact that $\bar{\phi}(s,\mu)$ increases in s for every μ. $\square$

Lemma 12. *If the agent's utility satisfies* (39) *then Assumption 2 is satisfied.*

Proof. Since $\bar{\phi}(s,\mu)$ is continuous in s, so are $\pi_\mu(s)$ and $\tilde{\pi}_\mu(s)$. Thus, the function $s \mapsto \tilde{\pi}_{\mu_t}^{-1}(\pi_{\mu_t}(s))$ is also continuous.

By (40), we have that

$$b = \bar{\phi}(\pi_\mu^{-1}(a), \mu) - \bar{\phi}(\tilde{\pi}_\mu^{-1}(a), \mu) = \left(\pi_\mu^{-1}(a) - \tilde{\pi}_\mu^{-1}(a)\right) \left[l \frac{\sigma_\epsilon^{-2}}{\sigma_t^{-2} + \sigma_\epsilon^{-2}} + (1-l) \right].$$

Taking absolute values implies that the distance between the true and perceived signal if the action equals a is bounded

$$\left|\pi_\mu^{-1}(a) - \tilde{\pi}_\mu^{-1}(a)\right| = \frac{b}{l \frac{\sigma_\epsilon^{-2}}{\sigma_t^{-2}+\sigma_\epsilon^{-2}} + (1-l)} \leq \frac{b}{1-l}.$$

□

Lemma 13. *If the agent's utility satisfies* (39) *then Assumption 3 is satisfied.*

Proof. Recall that for every normally distributed belief ν with posterior variance σ^2

$$\bar{\phi}(s, \nu) = \begin{cases} l \frac{\sigma^{-2}\mathbb{E}_\nu[\Theta] + \sigma_\epsilon^{-2} s}{\sigma^{-2}+\sigma_\epsilon^{-2}} + (1-l)s & \text{if } \sigma^2 > 0 \\ l\mathbb{E}_\nu[\Theta] + (1-l)s & \text{else} \end{cases}.$$

As $l < 1$, the function $\bar{\phi}$ is invertible in it's first argument and we denote it's inverse by $\bar{\phi}^{-1}(\cdot, \nu)$. By (40) we have that for any belief ν

$$\begin{aligned} \tilde{\pi}_\nu^{-1}(\pi_\nu(s)) &= \tilde{\pi}_\nu^{-1}\left(\bar{\phi}(s,\nu) - b\right) \\ &= \bar{\phi}^{-1}(\bar{\phi}(s,\nu) - b, \nu) \\ &= \bar{\phi}^{-1}(\bar{\phi}(s,\nu), \nu) - \frac{\partial \bar{\phi}^{-1}(\bar{\phi}, \nu)}{\partial \bar{\phi}} b \\ &= s - \frac{\sigma_\epsilon^2 + \sigma^2}{(1-l)\sigma_\epsilon^2 + \sigma^2} b. \end{aligned} \tag{41}$$

Here the second to last equation follows as $\bar{\phi}$ and thus also $\bar{\phi}^{-1}$ are linear, and the final follows from the fact that ${}^{\partial\bar{\phi}^{-1}(\bar{\phi},\nu)}/_{\partial\bar{\phi}} = \left({}^{\partial\bar{\phi}(s,\nu)}/_{\partial s}\right)^{-1}$. As σ_t is the variance of the belief μ_t, we get that

$$\begin{aligned} \left|\tilde{\pi}_{\mu_t}^{-1}\left(\pi_{\mu_t}(s)\right) - \tilde{\pi}_{\tilde{\theta}_t}^{-1}\left(\pi_{\tilde{\theta}_t}(s)\right)\right| &= \left|\frac{\sigma_\epsilon^2 + \sigma_t^2}{(1-l)\sigma_\epsilon^2 + \sigma_t^2} - \frac{\sigma_\epsilon^2}{(1-l)\sigma_\epsilon^2}\right| b \\ &\leq \frac{l\sigma_t^2}{(1-l)^2\sigma_\epsilon^2 + (1-l)\sigma_t^2} b \leq \frac{l}{(1-l)^2\sigma_\epsilon^2} b\sigma_t^2 \\ &= \frac{l}{(1-l)^2\sigma_\epsilon^2} b \frac{1}{\sigma_1^{-2} + \sigma_\epsilon^{-2}(t-1)} \\ &\leq \frac{l}{(1-l)^2} b(t-1)^{-1} \leq \frac{3lb}{(1-l)^2}\, t^{-1}. \end{aligned}$$

□

C.7 Proof of Proposition 6

By Equation 41,

$$\tilde{\pi}_{\tilde{\theta}}^{-1}\left(\pi_{\tilde{\theta}}(\Theta+\epsilon)\right)=\Theta-\frac{b}{1-l}+\epsilon.$$

Since by (1), any SOE needs to solve the following first-order condition, the above fact yields

$$0=\frac{\partial}{\partial z}\int \log f\left(\tilde{\pi}_{\tilde{\theta}}^{-1}\left(\pi_{\tilde{\theta}}(\Theta+\epsilon)\right)-z\right)f(\epsilon)d\epsilon\Big|_{z=\tilde{\theta}}=-\int \frac{f'\left(\tilde{\pi}_{\tilde{\theta}}^{-1}\left(\pi_{\tilde{\theta}}(\Theta+\epsilon)\right)-z\right)}{f\left(\tilde{\pi}_{\tilde{\theta}}^{-1}\left(\pi_{\tilde{\theta}}(\Theta+\epsilon)\right)-z\right)}\Bigg|_{z=\tilde{\theta}}f(\epsilon)d\epsilon$$

$$=-\int \frac{\left(\tilde{\pi}_{\tilde{\theta}}^{-1}\left(\pi_{\tilde{\theta}}(\Theta+\epsilon)\right)-\tilde{\theta}\right)}{\sigma_\epsilon^2}f(\epsilon)d\epsilon=-\frac{1}{\sigma_\epsilon^2}\left[\Theta-\frac{b}{1-l}-\tilde{\theta}+\underbrace{\int \epsilon f(\epsilon)d\epsilon}_{=0}\right].$$

Thus there is a unique SOE $\tilde{\theta}$ given by

$$\tilde{\theta}=\Theta-\frac{b}{1-l}.$$

Furthermore, Lemmas 11, 12, and 13 verify that Assumptions 1, 2, and 3 hold. Thus, Proposition 8 implies that beliefs converge with probability one to the unique SOE above. □

C.8 Proof of Proposition 7

If the agent has point beliefs $\tilde{\theta}$, his policy function $\pi_{\tilde{\theta}}(\Theta+\epsilon)$ maximizes

$$\arg\max_{a_t}\ e^{\Theta+\epsilon_t}u(1-a_t)+\beta h(\tilde{\theta})u(a_t)\,.$$

Rewriting the corresponding first order condition, gives

$$\frac{u'(1-\pi_{\tilde{\theta}}(\Theta+\epsilon))}{u'(\pi_{\tilde{\theta}}(\Theta+\epsilon))}=\frac{\beta h(\tilde{\theta})}{e^{\Theta+\epsilon_t}}. \tag{42}$$

Similarly, his perceived policy function $\tilde{\pi}_{\tilde{\theta}}(\tilde{\theta}+\epsilon))$ satisfies

$$\frac{u'(1-\tilde{\pi}_{\tilde{\theta}}(\tilde{\theta}+\epsilon))}{u'(\tilde{\pi}_{\tilde{\theta}}(\tilde{\theta}+\epsilon))}=\frac{\tilde{\beta}h(\tilde{\theta})}{e^{\tilde{\theta}+\epsilon_t}}.$$

The right-hand side of the above equals that of (42) — and hence the policy functions $\pi_{\tilde{\theta}}(\Theta+\epsilon) = \tilde{\pi}_{\tilde{\theta}}(\tilde{\theta}+\epsilon)$ for any given ϵ — if and only if $\tilde{\theta} = \Theta + \ln(\tilde{\beta}/\beta)$. Thus, by Observation 1, $\tilde{\theta} = \Theta + \ln(\tilde{\beta}/\beta)$ is the unique SOE.

Equation (12) follows immediately from the fact that $\tilde{\pi}_{\Theta}(\Theta+\epsilon))$ satisfies

$$\frac{u'(1-\tilde{\pi}_{\Theta}(\Theta+\epsilon))}{u'(\tilde{\pi}_{\Theta}(\Theta+\epsilon))} = \frac{\tilde{\beta}h(\Theta)}{e^{\Theta+\epsilon_t}},$$

and (42). $\square$

D Appendix: Two Dimensions of Uncertainty

Consider a variant of our main model in which the agent is uncertain about the benefit Θ^b and the harm Θ^κ of consuming. The prior regarding each fundamental is normal, and the agent believes these fundamentals to be independently drawn. In every period, nature chooses signals $s_t^b = \Theta^b + e_t^b$ and $s_t^\kappa = \Theta^b + e_t^b$, where the error terms are drawn normally and independently. In odd periods the agent observes s_t^b, and in even periods he observes s_t^κ before choosing his consumption level a_t. He maximizes the expectation of $v(a_t, s_t^b, s_t^\kappa, \theta^b, \theta^\kappa) = u(a_t) + \phi_t^b a_t - \beta e^{\phi_t^\kappa} a_t$, where $\phi_t^b = l\Theta^b + (1-l)s_t^b$ and $\phi_t^\kappa = l\Theta^\kappa + (1-l)s_t^\kappa$. He believes that all other selves maximize the utility function in which $\tilde{\beta}$ replaces β above.

Because the agent only observes s_t^κ in even periods, he changes his beliefs regarding Θ^κ only in even periods. We first focus on updating in these periods. Let $\tilde{\theta}_t^b$ denote the mean benefit of consumption as perceived by the agent in period t. And analogously to our basic model, denote by μ_t^κ the agent's belief about Θ^κ at the beginning of period t. The agent chooses a_t to satisfy

$$u'(a_t) + \tilde{\theta}_t^b = \beta \mathbb{E}_{\mu_t^\kappa}[e^{\phi_t^\kappa} | s_t^\kappa]$$

Since in an even period the agent does not observe s_t^b, he later has correct beliefs regarding $\tilde{\theta}_t^b$. Furthermore, since μ_t^κ only depends on previous actions, he has the correct belief about μ_t^κ as well. Hence, he believes that in period t he chooses a_t according to

$$u'(a_t) + \tilde{\theta}_t^b = \tilde{\beta} \mathbb{E}_{\mu_t^\kappa}[e^{\phi_t^\kappa} | \tilde{s}_t^\kappa].$$

These observations imply that the signal $\tilde{s}_t^\kappa$ the agent believes he has observed solves $\tilde{\beta}\mathbb{E}_{\mu_t^\kappa}[e^{\phi_t^\kappa}|\tilde{s}_t^\kappa] = \beta\mathbb{E}_{\mu_t^\kappa}[e^{\phi_t^\kappa}|s_t^\kappa]$. Notice that this is independent of $\tilde{\theta}_t^b$. Consequently, for any t, μ_t^κ, and s_t^κ the agent extracts the same $\tilde{s}_t^\kappa$ as when $\tilde{\theta}_t^b = 0$, i.e., as in the model of Section 3.3. By Proposition 3, therefore, his beliefs regarding future harm converge with probability one to

$$\tilde{\theta}^\kappa = \Theta^\kappa - \frac{\ln\tilde{\beta} - \ln\beta}{1-l}.$$

Now note that the agent changes his beliefs about the benefit only in odd periods. With abuse of notation, we restrict attention to the agent's behavior in odd periods, with $\tau = {}^{t+1}/_2$ denoting the position of the period in the sequence of odd numbers. Let $\tilde{\kappa}_t = \tilde{\kappa}_{2\tau-1} = \mathbb{E}_{\mu_t^\kappa}[e^{\phi_t^\kappa}]$ be the expected marginal harm as perceived by the agent when choosing $a_{2\tau-1}$. Crucially, the agent has the correct belief about this the harm perceived in previous (odd) periods. This allows us to follow the first steps of Proposition 5 with minor modifications. In particular, Equation (34) becomes

$$u'(a_{2\tau-1}) = l\left(1 - \frac{\sigma_{\tau+1}^2}{\sigma_\epsilon^2}\right)\tilde{\theta}_{2\tau-1}^b + \left[1 - l\left(1 - \frac{\sigma_{\tau+1}^2}{\sigma_\epsilon^2}\right)\right]s_{2\tau-1}^b - \beta\tilde{\kappa}_{2\tau-1}.$$

Hence,

$$s_{2\tau-1}^b = \frac{u'(a_{2\tau-1}) + \beta\tilde{\kappa}_{2\tau-1} - l\left(1 - \frac{\sigma_{\tau+1}^2}{\sigma_\epsilon^2}\right)\tilde{\theta}_{2\tau-1}^b}{1 - l\left(1 - \frac{\sigma_{\tau+1}^2}{\sigma_\epsilon^2}\right)}$$

and

$$\tilde{s}_t^b = \frac{u'(a_{2\tau-1}) + \tilde{\beta}\tilde{\kappa}_{2\tau-1} - l\left(1 - \frac{\sigma_{\tau+1}^2}{\sigma_\epsilon^2}\right)\tilde{\theta}_{2\tau-1}^b}{1 - l\left(1 - \frac{\sigma_{\tau+1}^2}{\sigma_\epsilon^2}\right)},$$

and therefore

$$\tilde{s}_{2\tau-1}^b - s_{2\tau-1}^b = \frac{(\tilde{\beta} - \beta)\tilde{\kappa}_{2\tau-1}}{1 - l\left(1 - \frac{\sigma_{\tau+1}^2}{\sigma_\epsilon^2}\right)}. \tag{43}$$

Since $\tilde{\kappa}_t$ converges with probability one to a constant $\tilde{\kappa}$ and the denominator on the right-hand side converges to $1 - l > 0$, the right-hand side converges with probability one to $(\tilde{\beta} - \beta)\tilde{\kappa}/(1-l)$.

The law of large numbers implies that $(\sum_{t=\tau}^T s_\tau^b)/T$ converges with probability one to Θ^b. Hence,

$(\sum_{\tau=1}^{T} \tilde{s}_{\tau}^{b})/T$ converges with probability one to

$$\Theta^b + \frac{(\tilde{\beta} - \beta)\tilde{\kappa}}{1 - l}.$$

This is therefore the agent's limiting belief $\tilde{\theta}^b$ about the benefit of consumption.