# Mislearning from Censored Data: Gambler's Fallacy in Optimal-Stopping Problems

Kevin He*

First version: March 21, 2018
This version: September 4, 2018

**Abstract**

I explore learning dynamics for people who believe in the "gambler's fallacy" — anticipating too much regression to the mean for realizations of independent random events. Agents arrive in large generations and face the same stage game, an optimal-stopping problem. They are initially uncertain about the distributions generating draws in different periods of the stage game and must infer these fundamentals from the decision histories of their predecessors. Each agent's stopping strategy thus imposes a *censoring effect* on the datasets of her successors, as her history does not record the future draws she would have found had she persisted longer in the stage game. While innocuous for rational agents, this censoring effect interacts with the gambler's fallacy and creates a positive-feedback loop between distorted stopping strategies and pessimistic beliefs about the fundamentals. In general settings, stopping strategies of successive generations converge monotonically to a steady-state strategy that stops earlier than optimal. If agents jointly infer the means and variances of the distributions, they will exaggerate variances to an extent that depends on the severity of censoring in their datasets of histories. The positive-feedback loop continues to obtain provided the optimal-stopping problem is convex.

Latest version of this paper: https://scholar.harvard.edu/files/kevin/files/gambler.pdf
Online Appendix: https://scholar.harvard.edu/files/kevin/files/gambler_oa.pdf

*Department of Economics, Harvard University. Email: hesichao@gmail.com. I am indebted to Drew Fudenberg, Matthew Rabin, Tomasz Strzalecki, and Ben Golub for their guidance and support. I thank In-Koo Cho, Martin Cripps, Krishna Dasaratha, Jetlir Duraj, Ben Enke, Ignacio Esponda, Yizhou Jin, Michihiro Kandori, Jonathan Libgober, Matthew Lilley, Xiaosheng Mu, Roberto Serrano, Michael Thaler, Linh T. Tô, Maria Voronina, and various seminar participants for insightful comments.

# 1 Introduction

The gambler's fallacy is widespread. Many people believe a fair coin has a higher chance of landing on tails after landing on heads three times in a row, think a son is "due" to a woman who has given births to consecutive daughters, and in general expect too much mean reversion from sequential realizations of independent random events. Early experimental evidence of this statistical bias came from the abstract domain of producing or recognizing i.i.d. random sequences based on a given alphabet of digits, letters, or colors (see Bar-Hillel and Wagenaar (1991) for a review). Recent work by Chen, Moskowitz, and Shue (2016) has shown that this fallacy also affects experienced decision-makers in high-stakes settings, such as judges in asylum courts. I discuss the empirical literature on this bias in Section 1.3.

This paper undertakes the first study of the dynamics of endogenous learning for a society of agents suffering from this statistical bias. Consider successive generations of agents facing the same single-player optimal-stopping problem in turn: they receive a draw each period and must decide between stopping to receive a payoff based on the current draw, or continuing for a future payoff depending on the next draw. Examples include managers choosing between hiring the current job applicant or continuing their recruitment search in hopes of finding a better candidate, and entrepreneurs deciding between liquidating their early-stage startup at the current market value or continuing to improve their prototype in hopes of greater future returns. Agents are uncertain about the distributions generating these draws in different periods, so they must use histories of their predecessors (e.g. experience of managers from previous hiring seasons, or historical data about entrepreneurs from past years) to learn parameters of these distributions — the "fundamentals" of the environment.

This situation presents a novel learning obstacle: the stopping decisions of earlier agents impose an endogenous *censoring effect* on the dataset of the current generation, as the future draws that these predecessors would have generated had they persisted longer in their decision problems remain unknown. When a manager decides to fill her job vacancy with a candidate discovered early in the hiring cycle, she stops her recruitment efforts and future managers do not observe what alternative candidates the firm would have found with additional search in the same hiring cycle. If an entrepreneur decides to liquidate her early-stage startup, future entrepreneurs cannot learn how her innovation would have matured had she kept working on her project. While harmless if agents were rational, the censoring effect interacts with the gambler's fallacy bias, leading to misinference about the fundamentals.

I suppose agents are Bayesians except for the gambler's fallacy bias, isolating the learning implications of the particular bias under consideration[1]. More precisely, agents start with a prior belief over a family of subjective models about the joint distribution of draws in different

[1]As an extension, Appendix B shows the main results of the paper are robust to non-Bayesian agents using a natural method-of-moments procedure to estimate the fundamentals.

periods, with the models differing in the unconditional means of the draws in different periods (the *fundamentals*) but all specifying the same negative correlation between draws. Agents apply Bayes' rule to update their belief over this class of misspecified models after observing a dataset of histories generated by their predecessors, in an environment where draws across periods are objectively independent.

This is not a theory about misunderstanding missing data — agents know the censoring mechanism and perform a Bayesian estimation procedure taking censoring into account. In fact, it is precisely this understanding of censoring that leads them astray. To develop an intuition for how the censoring effect leads to misinference, consider the psychology of a biased manager following the unlucky draw of a below-average first-period candidate in the hiring problem. The manager expects an above-average second-period candidate, for he expects the second-period draw to reflect a combination of the second-period fundamental and a positive reversal to "balance out" the bad first-period draw. But in a dataset containing the decision histories from past hiring cycle, the second-period candidate is only observed following this kind of bad first-period candidate, as the past managers would have stopped searching when they discovered an outstanding early candidate. A manager with the gambler's fallacy bias therefore expects the unconditional mean of second-period candidate quality to fall below its sample mean in the censored dataset, for he thinks the sample mean reflects a combination of the second-period fundamental and a positive contribution from the expected reversal following bad early draws. In reality, the early and late candidate qualities are independent. Therefore, the bias manager's line of reasoning leads him to underestimate the second-period fundamental, where the extent of the underestimation grows with the severity of censoring.

I focus on how this misinference unfolds over the course of learning. The severity of the censoring effect evolves continuously across generations as agents' beliefs about the fundamentals drift and their stopping strategies adjust accordingly. As an example, Section 4 documents the welfare implication of a positive-feedback cycle between wrong beliefs and wrong behavior in the case of hiring managers searching for a candidate across two periods. Starting with correct beliefs about fundamentals in the first generation, managers' beliefs about second-period candidate quality become more pessimistic across generations, leading to more relaxed acceptance standards for first-period candidates[2]. This, in turn, imposes a more severe censoring effect on future generations, as future managers only observe the second-period candidate quality if the first-period candidate fails to meet the newly lowered standard. The increased censoring effect brings about even more pessimistic beliefs about the second-period fundamental for managers with the gambler's fallacy, and hence a fur-

[2]In this example and for most of the paper, I consider the case where agents find it possible that draws are generated from different distributions in different periods, so they may estimate different values of the fundamental for different periods. The dynamics discussed here remain unchanged even if agents dogmatically believe the distributions are the same in all periods.

ther lowering of the acceptance threshold. In the steady state, the biased managers uses an acceptance threshold that is not only below the objectively optimal one, but also below the threshold they would have used if they only observed the histories generated in the $t$-th generation for any $t = 0, 1, 2, \ldots$. This monotonic mislearning is driven by the interaction between the censoring effect and the gambler's fallacy bias, not by either assumption alone. In a counterfactual world where today's managers observe all candidates that would have been drawn in previous hiring seasons, regardless of the actual stopping decisions of the earlier managers, a society of biased managers would nevertheless correctly infer the fundamentals and play the objectively optimal stopping strategy in every generation. On the other hand, managers who do not suffer from the gambler's fallacy can always correctly infer the fundamentals, even from a censored dataset.

Section 5 considers general optimal-stopping problems, fully characterizing the interplay between biased beliefs about fundamentals and distorted stopping strategies across generations. The key phenomena from Section 4's example remain robust. I find that the censoring effect always enable positive feedback between belief and behavior, so that both beliefs and stopping thresholds converge monotonically across generations to their steady-state values. In the long run, biased agents use suboptimal stopping rules with strictly lower stopping thresholds than the objectively optimal threshold. That is, the early-stopping phenomenon taking place every generation in Section 4's example obtains generally for late enough generations.

Section 6 extends the model and considers agents who are uncertain about both the means and variances of the draw distributions. I show that this joint estimation leads to the same misinference about means as in the baseline model on each censored dataset, but exaggerates the variances in a way that depends on the censoring threshold. I derive two results that illustrate how this *fictitious variation* interacts with endogenous learning. First, provided the optimal-stopping problem is convex, the positive-feedback cycle of the baseline model continues to obtain. This is because a more severely censored dataset not only makes successors more pessimistic about the second-period fundamental due to the usual censoring effect, but also decreases their belief in fictitious variation. Due to convexity of the optimal-stopping problem, both forces discourage successors from continuing into the second period, leading to a lower stopping threshold and even heavier data censoring in the future. Second, a society with agents uncertain about variances can end up with a different long-run belief about the means of the distributions than another society that knows the correct variances, even though agents in both societies would make the same (mis)inference about the means given the same dataset of histories.

## 1.1 Related Theoretical Work

Heidhues, Koszegi, and Strack (2018) find a similar mislearning pattern in the context of learners playing a static stage game with a different bias: overconfidence about own ability. But these results stem from two different sources. In addition to the assumed psychological bias, Heidhues, Koszegi, and Strack (2018)'s results depend on sign restrictions of cross partial derivatives of observable output, which ensures that initial action adjustments accentuate rather than dampen mislearning[3]. By contrast I consider a dynamic stage game, as the gambler's fallacy is a behavioral bias concerning the serial correlation of data. Since the censoring effect relies on the dynamic structure of the decision problem, it has no analog in a static-game setting. Also, the monotonicity of the stopping threshold across generations holds generally in my model — it does not depend on assumptions about prior belief, the relative pool qualities of different periods, or whether the draws are interpreted as "values" that the agent gets if he chooses to stop, or as "costs" that the agent must pay if he chooses to stop, as shown in Online Appendix OA 3.2. The positive feedback result is endogenously driven by the optimal stopping rule censoring future draws only after favorable early draws, rather than by exogenous assumptions on environmental primitives.

Rabin (2002) and Rabin and Vayanos (2010) were the first to study the inferential mistakes implied by the gambler's fallacy. With the exception of an example in Rabin (2002), all of their investigations have focused on passive inference where learners observe an exogenous information process. By contrast, I examine an endogenous learning setting where the actions of predecessors censor the dataset of the current learners. This setting allows me to ask whether the feedback loop between learners' actions and biased beliefs will attenuate or exaggerate the distortions caused by the fallacy over the course of learning. Also, relative to this existing literature, the current paper provides a unique focus on the dynamics of mislearning under the gambler's fallacy, tracing out the trajectory of beliefs and behavior over generations.

Rabin (2002) Section 7 discusses an example of endogenous learning under a finite-urn model of the gambler's fallacy. However, the nature of his endogenous data is unrelated to the censoring effect central to the present paper.[4] In Appendix D, I modify the Rabin (2002)

[3] Heidhues, Koszegi, and Strack (2018) assume that the marginal product of effort is non-increasing in ability, which rules out cases where effort and ability are strong complements such as in Bénabou and Tirole (2002). If this sign restriction is reversed, then over-confidence does not lead to misguided learning but under-confidence does.

[4] In Rabin (2002)'s example, the biased agents (correctly) believe that the part of the data which is always observable is independent of the part of the data that is sometimes missing. However, what I term the "censoring effect" is about misinference resulting from agents wrongly believing in negative correlation between the early draw that is always observed and the late draw that may be censored depending on the realization of the early draw. In this sense, my central mechanism highlights a novel interaction between censoring and the gambler's fallacy bias that is not present in the previous literature.

example to induce the censoring effect. I find a misinference result in his finite-urn model of the gambler's fallacy similar to what I find in the continuous Gaussian model of this paper, showing the robustness of the my results to different models of the statistical bias.

When the biased agents' subjective models specify a different auto-correlation between draws in the decision problem than the objective auto-correlation, no estimates of the fundamentals exactly match the data. I assume that agents in each generation observe histories from infinitely many decision problems from the previous generation and end up with a doctrinaire belief in the fundamentals minimizing Kullback–Leibler divergence to the observed data. Across generations, play converges to a Berk-Nash equilibrium of Esponda and Pouzo (2016). But rather than focusing only on equilibrium analysis, I focus on inter-generational learning dynamics to illustrate how the censoring effect drives the society towards the suboptimal steady state step by step. In Section 8, I provide a justification of the learning dynamics I study as the limit of finite-population dynamics when population size tends to infinity.

Finally, Fudenberg, Romanyuk, and Strack (2017) studied a continuous-time model of active learning under misspecification, where a single long-lived learner's belief is supported on two subjective models but neither corresponds to the truth. In contrast, I look at misspecified endogenous learning in a large-generations learning model, where agents entertain a continuum of subjective models all reflecting the psychological bias of the gambler's fallacy. The central mechanism of this paper, the censoring effect, does not appear in Fudenberg, Romanyuk, and Strack (2017), where information is generated through a static decision problem each instant.

## 1.2 Roadmap

The rest of the paper is organized as follows. The remainder of Section 1 reviews the empirical literature on the gambler's fallacy. Section 2 introduces the model, including both the optimal-stopping problem that serves as the stage game and the large-generations learning model. Section 3 contains the keys preliminary results about inference from censored datasets and the agent's behavior in the stage game given her belief about the fundamentals. Section 4 applies these results in an illustrative example about managers searching for candidates to fill an open position, showing how the interaction between the censoring effect and the gambler's fallacy leads to monotonic mislearning across generations. Section 5 turns to the general learning dynamics for a class of optimal-stopping problems, showing that qualitatively similar results obtain. Specifically, the positive feedback cycle holds generally while the long-run stopping threshold is suboptimally low. Section 6 turns to agents who jointly estimate means and variances from censored datasets. Section 7 demonstrates the robustness of the results to a number of extensions: (1) agents observing histories from multiple predecessor generations,

(2) agents with a non-dogmatic prior belief about the correlation (but with the support of this belief bounded below 0), and (3) societies with a fraction of selection neglecters in each generation. Section 8 provides a finite-population foundation for the large-generations inference results. Appendix B considers agents estimating the joint distributions of the draws from a family of general, possibly non-Gaussian subjective models, under a natural method-of moments-procedure. Appendix C studies misinference about fundamentals from censored datasets in decision problems with $L$ periods.

### 1.3 Empirical Evidence on the Gambler's Fallacy

Bar-Hillel and Wagenaar (1991) review classical psychology studies on the gambler's fallacy. In "production tasks" where subjects are asked to produce i.i.d. random sequences using a given alphabet, they tend to generate sequences with too many alternations between symbols as they attempt to balance out symbol frequencies locally. In "judgment tasks" where subjects are asked to identify which sequence of binary symbols appears most like consecutive tosses of a fair coin, subjects find sequences with alternation probability 0.6 more random than those with alternation probability of 0.5. The gambler's fallacy persists in the lab even when subjects are given feedback about the randomness of the sequences they generate (Budescu, 1987), when they are playing the matching pennies game where the strategy of randomizing 50-50 between heads and tails is the minimax strategy (Rapoport and Budescu, 1992), or when given monetary incentives so that the bet on a fair coin continuing its streak pays strictly more than the bet on the streak reversing (Benjamin, Moore, and Rabin, 2017). Recently, Barron and Leider (2010) showed that experiencing a streak of binary outcomes one at a time exacerbates the gambler's fallacy, compared with simply being told the past sequence of outcomes all at once.

A number of other studies have identified the gambler's fallacy using field data on lotteries and casino games. Unlike in experiments, agents in field settings are typically not explicitly told the underlying probabilities of the randomization devices. In state lotteries, players tend to avoid betting on numbers that have very recently won. This under-betting behavior is strictly costly for the players when lotteries have a pari-mutuel payout structure (as in the studies of Terrell (1994) and Suetens, Galbo-Jørgensen, and Tyran (2016)), since it leads to a larger-than-average payout per winner in the event that the same number is drawn again in the following week. Using security video footage, Croson and Sundali (2005) show that roulette gamblers in casinos bet more on a color after a long streak of the opposite color. Narayanan and Manchanda (2012) use individual-level data tracked using casino loyalty cards to find that a larger recent win has a negative effect on the next bet that the gambler places, while a larger recent loss increases the size of the next bet. This result extends gambler's fallacy beyond the binary outcomes domain and suggests the same psychology also operates for continuous outcomes, with the severity of the recent bad outcome believed to foretell the degree mean reversal "due" in the near future. Finally, using data from the diverse areas of asylum granting, loan approvals, and baseball umpire calls, Chen, Moskowitz, and Shue (2016) show that even very experienced decision-makers show a tendency to alternate between two decisions across a sequence of randomly ordered decision problems. This can be explained by gambler's fallacy, as the fallacy leads to the belief that the objectively "correct" decision is negatively autocorrelated across the sequence of decision problems. The authors rule out a number of other explanations including contrast effect and quotas.

As Rabin (2002) and Rabin and Vayanos (2010) have argued, someone who dogmatically believes in the gambler's fallacy must attribute the lack of reversals in the data to the fundamental probabilities of the randomizing device, leading to overinference from small dataset. This overinference can be seen in the field data. Cumulative win/loss (as opposed to very recent win/loss) on a casino trip is positively correlated with the size of future bets (Narayanan and Manchanda, 2012). A player who believes in the gambler's fallacy rationalizes his persistent good luck on a particular day by thinking he must be in a "hot" state, where his fundamental probability of winning in each game is higher than usual. In a similar vein, a number that has been drawn more often in past 6 weeks, excluding the most recent past week, gets more bets in the Denmark lottery (Suetens, Galbo-Jørgensen, and Tyran, 2016). This kind of overinference result from small samples persists even in a market setting where participants have had several rounds of experience and feedback (Camerer, 1987). In line with these evidence, the model I consider involves agents who dogmatically believe in the gambler's fallacy and misinfer some parameter of the world as a consequence — though the misinference mechanism in my model is further complicated by the presence of endogenous data censoring.

# 2 Model

I introduce four aspects of the model in turn. Section 2.1 sets up the optimal-stopping problem that serves as the stage game of the learning environment. Section 2.2 discusses stopping strategies and histories in the stage game, highlighting how stopping strategies naturally "censor" game histories as the hypothetical values that would have been drawn after stopping remain unobserved. Section 2.3 explains that agents have a prior belief over a class of "subjective models" about how draws in different periods are jointly generated, with all subjective models in the prior's support exhibiting the gambler's fallacy bias. Section 2.4 details the large-generation learning environment and how agents make inferences from their observations.

## 2.1 Optimal-Stopping Problem as a Stage Game

The stage game is a two-period optimal-stopping problem. In the first period, the agent draws a value $x_1 \in \mathbb{R}$ and decides whether to stop. If she decides to stop at $x_1$, her payoff is $u_1(x_1)$ and the stage game ends. Otherwise, she continues to the second period where she draws another value $x_2 \in \mathbb{R}$. The stage game then ends with the agent getting payoffs $u_2(x_1, x_2)$. The values $(x_1, x_2)$ are the realizations of a pair of random variables $(X_1, X_2)$. The agent holds some belief about the joint distribution of $(X_1, X_2)$ and makes her stopping decision in the first period as to maximize her expected utility given her belief.

The payoff functions $u_1 : \mathbb{R} \to \mathbb{R}$ and $u_2 : \mathbb{R}^2 \to \mathbb{R}$ are continuous and satisfy some regularity conditions to be introduced in Assumption 1. Through appropriate choices of $u_1$ and $u_2$, this general setup can accommodate a range of economic situations. I illustrate with two examples.

**Example 1** (Search)**.** Many industries have a regular hiring season each year. Suppose the agent is the manager of a firm in such an industry, who must fill a vacancy in her company. In the early phase of the hiring season, her human resource team interviews a number of applicants and identifies the best candidate, who would bring net benefit $x_1$ to the firm if hired. The manager must decide between hiring this candidate immediately or interviewing more candidates. If the manager decides to wait, her team can continue searching for a potentially better candidate in the late phase of the hiring season — one who would bring net benefit $x_2$ to the organization. But waiting carries the risk that the early candidate accepts an offer at a different firm in the interim. Suppose there is $q \in [0, 1)$ probability that the early candidate will remain available for hiring in the late hiring phase. Then we may let $u_1(x_1) = x_1$ and $u_2(x_1, x_2) = q \cdot \max(x_1, x_2) + (1 - q)x_2$ to model this situation. That is, there is $q$ probability the manager gets payoff equal to the highest of the two candidates' qualities, and $1 - q$ probability that she only has the option to hire the second candidate. ♦

**Example 2** (Startup)**.** An entrepreneur pays effort cost $\kappa_1 > 0$ in period 1 to develop a startup valued at $v_1(x_1)$, where $v_1 : \mathbb{R} \to \mathbb{R}$ is strictly increasing with $\lim_{x_1 \to \infty} v_1(x_1) = \infty$ and $x_1$ captures the idiosyncratic luck or difficulty she encounters. She chooses whether to sell the early startup for $v_1$, or whether to work more and mature her startup further. If she chooses to work more, she pays a further effort cost $\kappa_2 > 0$ and the market value of her startup changes to $\alpha v_1(x_1) + v_2(x_2)$, where $\alpha \in (0, 1)$ and $v_2 : \mathbb{R} \to \mathbb{R}$ is another strictly increasing function with $\lim_{x_2 \to \infty} v_2(x_2) = \infty$. To interpret, $x_2$ represents idiosyncratic factors affecting the growth of her project, while the baseline value of her prototype is discounted to $\alpha v_1(x_1)$ due to changing market demand or the entrance of competitors. In this example, $u_1(x_1) = v_1(x_1) - \kappa_1$, $u_2(x_1, x_2) = \alpha v_1(x_1) + v_2(x_2) - \kappa_1 - \kappa_2$. ♦

I now present a set of conditions on the payoff functions that will be maintained throughout.

**Assumption 1.** *The payoff functions satisfy:*

*(a) For* $x_1^{'} > x_1^{''}$ *and* $x_2^{'} > x_2^{''}$, $u_1(x_1^{'}) > u_1(x_1^{''})$ *and* $u_2(x_1^{'}, x_2^{'}) > u_2(x_1^{'}, x_2^{''})$.

*(b) For* $x_1^{'} > x_1^{''}$ *and any* $\bar{x}_2$, $u_1(x_1^{'}) - u_1(x_1^{''}) > u_2(x_1^{'}, \bar{x}_2) - u_2(x_1^{''}, \bar{x}_2)$.

*(c) There exists* $L > 0$ *so that* $u_1(L) - u_2(L, -L) \geq 0$, *while* $u_1(-L) - u_2(-L, L) \leq 0$.

Assumption 1(a) says $u_1, u_2$ are strictly increasing in the draws in their respective periods. Assumption 1(b) says a higher realization of the early draw is more helpful for first-period payoff than for second-period payoff. Under Assumption 1(a), Assumption 1(b) is satisfied whenever $u_2$ is not a function of $x_1$, or more generally when $u_2(x_1, x_2) = z_{2,1}(x_1) + z_{2,2}(x_2)$ is separable across the draws of the two periods with $z_{2,1}^{'}(x_1) < u_1^{'}(x_1)$ at all $x_1 \in \mathbb{R}$. Assumption 1(c) says if the agent knew that $x_1$ were sufficiently positive and $x_2$ were sufficiently negative, then she would prefer to stop in period 1. Conversely, if she knew that $x_1$ were very negative while $x_2$ were very positive, then she would prefer to continue to period 2. The examples above satisfy these conditions.

*Claim* 1. Examples 1 and 2 satisfy Assumption 1.

Omitted proofs from the main text can be found in Appendix A.

## 2.2 Stopping Strategies and Endogenous Censoring of Histories

I now turn to strategies and histories in the stage game.

**Definition 1.** A **strategy** is a function $s : \mathbb{R} \rightarrow$ {Stop, Continue} that maps the realization of the first-period draw $X_1 = x_1$ into a stopping decision.

Without loss I consider only pure strategies, because under any subjective belief about the joint distribution of $(X_1, X_2)$, the agent can maximize her expected utility using a pure stopping strategy.

**Definition 2.** The **history** of the stage game is an element $h \in H := \mathbb{R} \times (\mathbb{R} \cup \{\varnothing\})$. If an agent decides to stop after $X_1 = x_1$, her history is $(x_1, \varnothing)$. If the agent continues after $X_1 = x_1$ and draws $X_2 = x_2$ in the second period, her history is $(x_1, x_2)$.

The symbol $\varnothing$ is a *censoring indicator*, emphasizing that the hypothetical second-period draw is unobserved when an agent does not continue into the second period. This observation structure is natural in my examples. In Example 1, if a firm fills their vacancy in the early phase of the hiring cycle, they would stop their recruitment efforts and the counterfactual candidate that they would have found had they kept headhunting in the same hiring season remains unknown. In Example 2, when an entrepreneur decides to liquidate her early-stage startup, society never learns what her project could have grown into had she pursued it further.

To preview the learning environment that I will describe in Sections 2.3 and 2.4, agents in successive generations form beliefs about the joint of $(X_1, X_2)$ using their prior and their observations. Then, they choose a stopping strategy as to maximize their expected utility given their beliefs. The stage-game histories of agents in generation $t$ become observations for agents in generation $t + 1$ and this process repeats. A key feature of my model is that the censoring of histories is endogenous. How histories are censored depends on the stopping strategy of the predecessors, which in turn depends on their beliefs. In contrast, the existing literature on learning under the gambler's fallacy has focused on biased learners passively forming inference from observing an exogenous flow of information. For instance, Rabin and Vayanos (2010) interpret their model as an observer seeing the returns time series of a mutual fund run by a team of managers under an exogenous turnover process, then estimating the parameters of that process. The bulk of my results concern the implications of endogenous learning under the gambler's fallacy bias and the interaction between distorted stopping strategies and distorted beliefs.

## 2.3 Objective and Subjective Models of Draws

Objectively, $X_1, X_2$ are independently drawn from Gaussian distributions $X_1 \sim \mathcal{N}(\mu_1^\bullet, \sigma^2)$ and $X_2 \sim \mathcal{N}(\mu_2^\bullet, \sigma^2)$ where parameters $\mu_1^\bullet, \mu_2^\bullet \in \mathbb{R}$ are fixed and called *fundamentals*. In Example 1, $\mu_1^\bullet$ and $\mu_2^\bullet$ stand for the underlying qualities of the two applicant pools in the early and late phases of the hiring season. In Example 2, $\mu_1^\bullet$ relates to the selling price of an early startup and $\mu_2^\bullet$ is associated with expected improvement from maturing an early startup.

The agent's belief deviates from the objective model in two ways. First, the agent is uncertain about the fundamentals. In addition, the agent suffers from the gambler's fallacy, causing her to misperceive the joint distribution of $(X_1, X_2)$ conditional on the fundamentals. She believes that if the first draw is higher than expected (based on her belief about the fundamental), then bad luck is "due" in the near future and the second draw is most likely below average. Conversely, an exceptionally bad early draw likely portends she will have above average luck in the next period. Formally, the agent's belief is supported on a class of subjective models about the joint distribution of $(X_1, X_2)$, indexed by her estimate of the fundamentals $(\mu_1, \mu_2)$.

**Definition 3.** For $\mu_1, \mu_2 \in \mathbb{R}$, $\sigma_1^2, \sigma_2^2 > 0$, and $\gamma' \leq 0$, the **subjective model** $\Xi(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2; \gamma')$ refer to the joint distribution for $(X_1, X_2)$,

$$\left\{ \begin{array}{c} X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2) \\ (X_2 | X_1 = x_1) \sim \mathcal{N}(\mu_2 + \gamma'(x_1 - \mu_1), \sigma_2^2) \end{array} \right\}$$

where $X_2 | X_1 = x_1$ is the conditional distribution of $X_2$ given $X_1 = x_1$.[5] The **objective model** is $\Xi^\bullet = \Xi(\mu_1^\bullet, \mu_2^\bullet, \sigma^2, \sigma^2; 0)$. The agent's **feasible subjective models** is the set $\{\Xi(\mu_1, \mu_2, \sigma^2, \sigma^2; \gamma) : \mu_1, \mu_2 \in \mathbb{R}\}$, and $\gamma < 0$ is his **bias parameter**.

The fundamentals $(\mu_1, \mu_2)$ in the subjective model $\Xi(\mu_1, \mu_2, \sigma^2, \sigma^2; \gamma)$ represent the unconditional means of the two distributions. Although $(X_1, X_2)$ are objectively independent, every feasible subjective model predicts a worse $X_2$ following a better $X_1$, due to $\gamma < 0$. The magnitude of $\gamma$ corresponds to the severity of his gambler's fallacy bias. I interpret the agent's wrong belief about the correlation in early and late draws as coming from the same psychology that leads people to mispredict the likelihood of a fair coin landing heads after having landed on heads multiple times in a row. In this model, an $X_1$ realization far above its mean is analogous to a sequence of heads in a row — a highly unbalanced outcome that must be followed with an $X_2$ far below its unconditional mean if the sample is to be overall "representative" of the population means.

Throughout this paper I write $\mathbb{E}_\Xi$ and $\mathbb{P}_\Xi$ for expectation and probability with respect to the subjective model $(X_1, X_2) \sim \Xi$. When $\mathbb{E}$ and $\mathbb{P}$ are used without subscripts, they refer to expectation and probability under the objective model $\Xi^\bullet$.

*Remark* 1. Alternatively, the subjective model $\Xi(\mu_1, \mu_2, \sigma^2, \sigma^2; \gamma)$ may be written as follows:

$$X_1 = \mu_1 + \epsilon_1$$
$$X_2 = \mu_2 + \epsilon_2$$

where $\epsilon_1 \sim \mathcal{N}(0, \sigma^2)$ and $\epsilon_2 | \epsilon_1 \sim \mathcal{N}(\gamma \epsilon_1, \sigma^2)$. The terms $\epsilon_1, \epsilon_2$ can be interpreted as the decision-maker's luck in the first and second periods, which determine the realizations of the draws $X_1, X_2$ relative to their unconditional means $\mu_1, \mu_2$. The subjective model stipulates reversal of luck across the two periods, as $(\epsilon_1, \epsilon_2)$ are negatively correlated. ♦

*Remark* 2. Since all feasible subjective correctly specify the variance of $X_1$ and the condi-

[5] Equivalently, $(X_1, X_2)$ have a joint Gaussian distribution with

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim \mathcal{N}\left( \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \gamma \sigma_1^2 \\ \gamma' \sigma_1^2 & (\gamma')^2 \sigma_1^2 + \sigma_2^2 \end{pmatrix} \right).$$

The correlation between $X_1$ and $X_2$ is $\frac{\gamma' \sigma_1}{\sqrt{(\gamma')^2 \sigma_1^2 + \sigma_2^2}}$.

tional variance of $X_2|X_1$, I will sometimes abbreviate the subjective model $\Xi(\mu_1, \mu_2, \sigma^2, \sigma^2; \gamma)$ as $\Xi(\mu_1, \mu_2; \gamma)$. Section 6 investigates agents who are uncertain about these variances and must jointly infer the means and variances from their observations. ◆

Each agent starts with the same full-support prior belief over the feasible subjective models. This belief is induced by a strictly positive prior density function $g : \mathbb{R}^2 \to \mathbb{R}_{++}$ about the fundamentals $(\mu_1, \mu_2)$. Before playing her own stage game, each agent observes data generated from other people's experience and updates $g$ in a Bayesian way to form a posterior belief over the feasible subjective models. Returning to the examples, this corresponds to observing the interviewing and hiring histories of other firms in her industry from last year's hiring season, or the improvements made by previous entrepreneurs who chose to keep working on their early-stage startups. The agent then uses her observations to make Bayesian inferences about parameters of the environment: the qualities of the early and late applicant pools, the profitability of maturing an early-stage startup, etc.

The agent can update her belief about the fundamentals, but all of her feasible subjective models specify the same $\gamma$, so she not revise this aspect of her mental model in light of data. The agent's dogmatic belief in $\gamma < 0$ is restrictive, but allows me to focus attention on the learning implications of gambler's fallacy. As Section 7.2 shows, results are unchanged if the agent also updates his belief about the coefficient $\gamma$, provided the support of his prior belief lies to the left of 0 and is bounded away from it. This assumption seems broadly in line with Chen, Moskowitz, and Shue (2016)'s analysis of field data, showing that even very experienced decision-makers continue to exhibit a non-negligible amount of the gambler's fallacy in high-stakes settings.

Another reason why agents may never question their misspecified prior is that the misspecification is "attentionally stable" in the sense of Gagnon-Bartsch, Rabin, and Schwartzstein (2018). Under the theory that the true model falls within the feasible subjective models, an agent finds it harmless to coarsen her dataset by only paying attention to certain "summary statistics". In large datasets, the statistics extracted by the limited-attention agent do not lead her to question the validity of her theory, even though a full-attention agent who retains the entire raw dataset could calculate other statistics that lead her to believe that her prior is misspecified. To summarize the results in Appendix E, for arbitrary full-support prior $g$ over the fundamentals, the Bayesian posterior density $g(\cdot|(h_n)_{n=1}^N)$ after observing a finite dataset of $N$ stage-game histories $(h_n)_{n=1}^N$ only depends on the dataset through two sufficient statistics: (i) the sample average of first-period draws; (ii) the sample average of "re-centered", uncensored second-period observations, where in the history $h_n = (x_1, x_2)$ the re-centered observation is defined as $x_2 - \gamma x_1$. An agent who compresses every dataset to just these two statistics finds this coarsening harmless for decision-making purposes and never needs to notice the true correlation between $X_1$ and $X_2$. In large samples, the inference procedure that only uses these two extracted statistics produces the same results as the full-attention Bayesian procedure I outline in Section 2.4, and furthermore the realized values of the statistics can always be rationalized by some subjective model in the support of the agent's prior.

Since the gambler's fallacy is a statistical bias about sequential realizations of random variables, the baseline two-periods model is the minimum model capturing the implications of this bias and its interaction with censoring. In Appendix C, I study an $L$-periods model of the gambler's fallacy based on Rabin and Vayanos (2010) and derive results about inference from censored data in stage games with a longer horizon.

## 2.4 Learning in Large Generations

This section details the learning environment and describes how agents make inferences from datasets of histories.

There is an infinite sequence of generations, $t \in \{0, 1, 2, ...\}$. Each generation consists of a continuum of agents $n \in [0, 1]$, with each agent only living for one generation. In the search problem of Example 1, for instance, successive generations refer to cohorts of hiring managers working in successive hiring cycles. The realizations of draws $X_1, X_2$ are independent across all stage games, including those from the same generation.

Before playing her own stage game, each agent in generation $t \geq 1$ observes an infinite dataset of histories $(h_n)_{n \in [0,1]}$. This dataset contains all the stage-game histories from generation $t-1$, where $h_n$ is the history of predecessor $n$ from that generation. The distribution of observed histories depends on the joint distribution of $(X_1, X_2)$ as well as the stopping strategy[6] used by predecessors. Agents are told the stopping strategy used by their predecessors[7] and use the dataset of histories to infer the joint distribution between $(X_1, X_2)$ from the class of feasible subjective models. Equivalently, agents infer fundamentals $\mu_1, \mu_2 \in \mathbb{R}$.

Before I can describe this inference procedure, I first introduce some notations for the distribution of histories in a dataset. For any measurable strategy $s : \mathbb{R} \to \{\text{Stop, Continue}\}$ and subjective model $\Xi$, let $\mathcal{H}(\Xi; s) \in \Delta(H)$ denote the distribution of histories when draws are distributed according to $(X_1, X_2) \sim \Xi$ and the agent continues into the second period if and only if $s(X_1) = \text{Continue}$. This is formalized below.

**Definition 4.** For measurable strategy $s$ and subjective model $\Xi = \Xi(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2; \gamma')$ where $\mu_1, \mu_2, \gamma' \in \mathbb{R}$, $\sigma_1^2, \sigma_2^2 > 0$, let $\mathcal{H}(\Xi; s)$ be the distribution on the space of histories,

[6] Since agents in each generation start with the same prior and observe the same dataset, they all hold the same beliefs about the fundamentals and play the same subjectively optimal stopping strategy given their beliefs.

[7] This stopping rule can also be exactly inferred from the infinite dataset.

$H = \mathbb{R} \times (\mathbb{R} \times \{\varnothing\})$, given by

$$\begin{aligned} \mathcal{H}(\Xi; s)[E_1 \times E_2] &:= \quad \mathbb{P}_\Xi\left[(E_1 \cap s^{-1}(\text{Continue})) \times E_2\right] \quad \text{for } E_1, E_2 \in \mathcal{B}(\mathbb{R}) \\ \mathcal{H}(\Xi; s)[E_1 \times \{\varnothing\}] &:= \quad \mathbb{P}_\Xi[(E_1 \cap s^{-1}(\text{Stop})) \times \mathbb{R}] \qquad \text{for } E_1 \in \mathcal{B}(\mathbb{R}), \end{aligned}$$

where $\mathbb{P}_\Xi$ is the probability measure on $\mathbb{R}^2$ given by $\Xi$, while $\mathcal{B}(\mathbb{R})$ is the collection of Borel subsets of $\mathbb{R}$.

The distribution $\mathcal{H}(\Xi^\bullet; s)$ refers to the objective distribution on histories when predecessors use the stopping strategy $s$. I will abbreviate it as $\mathcal{H}^\bullet(s)$.

Next, I define the Kullback-Leibler (KL) divergence from $\mathcal{H}(\Xi; s)$, the history distribution under model, to the objective history distribution, $\mathcal{H}^\bullet(s)$. For a given stopping strategy, the *pseudo-true fundamentals* are $\mu_1^*, \mu_2^* \in \mathbb{R}$ such that the feasible subjective model $\Xi(\mu_1^*, \mu_2^*; \gamma)$ minimizes this KL divergence.

**Definition 5.** (a) The **Kullback-Leibler (KL) divergence** from $\mathcal{H}^\bullet(s)$ to $\mathcal{H}(\Xi(\mu_1, \mu_2; \gamma); s))$, denoted by $D_{KL}(\mathcal{H}^\bullet(s) \mid\mid \mathcal{H}(\Xi(\mu_1, \mu_2; \gamma); s))$ ), is

$$\int_{x_1 \in s^{-1}(\text{Stop})} \phi(x_1; \mu_1^\bullet, \sigma^2) \cdot \ln\left(\frac{\phi(x_1; \mu_1^\bullet, \sigma^2)}{\phi(x_1; \mu_1, \sigma^2)}\right) dx_1$$
$$+ \int_{x_1 \in s^{-1}(\text{Continue})} \left\{\int_{-\infty}^{\infty} \phi(x_1; \mu_1^\bullet, \sigma^2) \cdot \phi(x_2; \mu_2^\bullet, \sigma^2) \cdot \ln\left[\frac{\phi(x_1; \mu_1^\bullet, \sigma^2) \cdot \phi(x_2; \mu_2^\bullet, \sigma^2)}{\phi(x_1; \mu_1, \sigma^2) \cdot \phi(x_2; \mu_2 + \gamma(x_1 - \mu_1), \sigma^2)}\right] dx_2\right\} dx_1$$

where $\phi(x; a, b^2)$ is Gaussian density with mean $a$ and variance $b^2$.

(b) Say $\mu_1^*, \mu_2^*$ are the **pseudo-true fundamentals** with respect to the strategy $s$ if

$$(\mu_1^*, \mu_2^*) \in \underset{\mu_1, \mu_2 \in \mathbb{R}}{\arg\min}\ D_{KL}(\mathcal{H}^\bullet(s) \mid\mid \mathcal{H}(\Xi(\mu_1, \mu_2; \gamma); s))\ ).$$

To interpret, the likelihood of the history $h = (x_1, x_2)$ with $s(x_1) = \text{Continue}$ is $\phi(x_1; \mu_1^\bullet; \sigma^2) \cdot \phi(x_2; \mu_2^\bullet, \sigma^2)$ under the objective model $\Xi^\bullet$, $\phi(x_1; \mu_1, \sigma^2) \cdot \phi(x_2; \mu_2 + \gamma(x_1 - \mu_1), \sigma^2)$ under the subjective model $\Xi(\mu_1, \mu_2; \gamma)$. The likelihood of the history $h = (x_1, \varnothing)$ with $s(x_1) = \text{Stop}$ is $\phi(x_1; \mu_1^\bullet; \sigma^2)$ under the objective model, $\phi(x_1; \mu_1, \sigma^2)$ under the subjective model. The likelihoods of all other histories are 0 under both models. So the KL divergence expression given in Definition 5 is the expected log-likelihood of the history under the objective model versus under the subjective model with fundamentals $(\mu_1, \mu_2)$, with the expectation taken under the objective model. In general, this quantity depends on the stopping strategy $s$, so I will occasionally denote the pseudo-true fundamentals as $\mu_1^*(s)$, $\mu_2^*(s)$ to emphasize this dependence.

When the $t$-th generation agents observe an infinite dataset of histories with the distribution $\mathcal{H}^\bullet(s)$, the agents update their prior $g$ to put dogmatic belief in the pseudo-true fundamentals $\mu_1^*(s), \mu_2^*(s)$, then play the subjectively optimal stopping strategy for the model $\Xi(\mu_1^*(s), \mu_2^*(s); \gamma)$. In Section 8, I establish that when a Bayesian agent with prior $g$ observes a finite dataset of $N$ histories $(h_n)_{n=1}^N$ drawn from the distribution $\mathcal{H}^\bullet(s)$ where $s$ is a cutoff-based stopping rule, then as $N \to \infty$ her posterior belief about the fundamentals $g(\cdot|(h_n)_{1=1}^N)$ almost surely converges in mean to the point-mass belief on the pseudo-true parameters $(\mu_1^*, \mu_2^*)$, and furthermore her posterior expected payoff from any cutoff-based stopping rule converges to its expected payoff under the model $\Xi(\mu_1^*(s), \mu_2^*(s); \gamma)$. In the large-generations learning model, belief dynamics between generations are given a deterministic transition between point mass beliefs, greatly simplifying the analysis.

One assumption behind this procedure is that agents do not reason through why previous agents made their stopping decisions. So, agents infer nothing about the fundamentals from the strategic choices of the previous generation. As the analysis of the learning dynamics will show, the strategies of different generations converge, so that asymptotically agents find the strategies of the previous generation approximately optimal given their own beliefs about the fundamentals. Section 7.1 considers a modified learning model where each generation observes the histories of all previous generations. In that setting we may assume common knowledge of rationality among the agents, as information sets are nested and generation $t$ observes all the information that generation $t' < t$ had, so there is nothing more to infer from the actions of generation $t'$. Here agents will still converge to the same steady state as when they observe only the immediate predecessor generation and do not assume rationality of others, though the rate of convergence to the steady state may be slower.

The key intuition behind my main endogenous learning results do not depend on full Bayesianism or on the Gaussian functional form. Appendix B considers gambler's fallacy agents who start with a class of possibly non-Gaussian subjective models of $(X_1, X_2)$ and infer the joint distribution of the draws by applying a natural method-of-moments procedure to the dataset. The positive feedback between distorted beliefs and distorted stopping behavior continues to hold.

# 3 Optimal Stopping Rules and Inference from Censored Datasets

In this section, I develop a number of preliminary results. Section 3.1 derives the subjectively optimal stopping rule for an agent who believes in the model $(X_1, X_2) \sim \Xi(\mu_1, \mu_2; \gamma)$. I show that this stopping rule involves a cutoff threshold that increases in belief about the second-period fundamental. Section 3.2 characterizes the Bayesian inference about fundamentals from large datasets of histories censored with a cutoff-based stopping strategy, when agents start with a full-support prior over the (misspecified) class of feasible models. Finally, Section 3.3 study the constrained inference when agents know that the fundamentals are time-invariant across periods.

## 3.1 Subjectively Optimal Stopping Rules

Consider an agent who believes in the subjective model $(X_1, X_2) \sim \Xi(\mu_1, \mu_2; \gamma)$, implied by a belief in the fundamentals $(\mu_1, \mu_2) \in \mathbb{R}^2$. The next lemma characterizes the subjectively optimal stopping strategy given this model. I show that this stopping rule must be a cutoff strategy: there exists some value $c \in \mathbb{R}$ depending on $(\mu_1, \mu_2)$ such that the agent strictly prefers stopping after any $x_1 > c$ and strictly prefers continuing after any $x_1 < c$.

**Lemma 1.** *Under the belief that $(X_1, X_2) \sim \Xi(\mu_1, \mu_2; \gamma)$, for any $\gamma \leq 0$, there exists a cutoff $C(\mu_1\mu_2)$, such that the agent strictly prefers stopping after any $x_1 > C(\mu_1\mu_2)$ and strictly prefers continuing after any $x_1 < C(\mu_1\mu_2)$.*

The next result establishes monotonicity of the cutoff in terms of belief about the second-period fundamental.

**Lemma 2.** *The indifference threshold $C(\mu_1, \mu_2)$ is strictly increasing in $\mu_2$.*

*Proof.* Let $\hat{\mu}_1, \hat{\mu}_2, \hat{\hat{\mu}}_2 \in \mathbb{R}$ with $\hat{\hat{\mu}}_2 > \hat{\mu}_2$. I show that $C(\hat{\mu}_1\hat{\mu}_2) < C(\hat{\mu}_1\hat{\hat{\mu}}_2)$.

By Lemma 1, the threshold $C(\hat{\mu}_1, \hat{\mu}_2)$ is characterized by the indifference condition,

$$u_1(C(\hat{\mu}_1, \hat{\mu}_2)) = \mathbb{E}_{\tilde{X}_2 \sim \mathcal{N}(\hat{\mu}_2 + \gamma(C(\hat{\mu}_1, \hat{\mu}_2) - \hat{\mu}_1), \sigma^2)}[u_2(C(\hat{\mu}_1, \hat{\mu}_2), \tilde{X}_2)]$$

But if agent were to instead believe $(\hat{\mu}_1\hat{\hat{\mu}}_2)$ where $\hat{\hat{\mu}}_2 > \hat{\mu}_2$, then the conditional distribution of $X_2$ given $X_1 = C(\hat{\mu}_1, \hat{\mu}_2)$ would be $\mathcal{N}(\hat{\hat{\mu}}_2 + \gamma(C(\hat{\mu}_1, \hat{\mu}_2) - \hat{\mu}_1), \sigma^2)$. We have

$$u_1(C(\hat{\mu}_1, \hat{\mu}_2)) < \mathbb{E}_{\tilde{X}_2 \sim \mathcal{N}(\hat{\hat{\mu}}_2 + \gamma(C(\hat{\mu}_1, \hat{\mu}_2) - \hat{\mu}_1), \sigma^2)}[u_2(C(\hat{\mu}_1, \hat{\mu}_2), \tilde{X}_2)]$$

by Assumption 1(a). This means $C(\hat{\mu}_1, \hat{\mu}_2) < C(\hat{\mu}_1, \hat{\hat{\mu}}_2)$ by Lemma 1, as only values of $X_1$ below $C(\hat{\mu}_1, \hat{\hat{\mu}}_2)$ lead to strict preference for continuing. $\square$

## 3.2 Inference about Fundamentals from Censored Datasets

I now turn to generation $t+1$'s inference when all agents in generation $t$ use a cutoff strategy. For $c \in (-\infty, \infty]$, denote the stopping strategy $s$ where $s(x) =$ Stop for all $x > c$ and $s(x) =$ Continue for all $x < c$ as $c \uparrow$,[8] evocative of the stopping region $[c, \infty)$. So, when all

[8] Whether $s(c) =$ Stop or $s(c) =$ Continue only changes history on a zero-probability event, so it does not affect inference.

agents use strategy $c\uparrow$, the next generation observes a dataset of histories with distribution $\mathcal{H}(\Xi^\bullet; c\uparrow)$. I will abbreviate this distribution as $\mathcal{H}^\bullet(c\uparrow)$.

I now find an explicit expression of the pseudo-true fundamentals $(\mu_1^*, \mu_2^*)$ as a function of $c \in \mathbb{R}$.

**Proposition 1.** *The pseudo-true fundamentals minimizing $D_{KL}(\mathcal{H}^\bullet(c\uparrow) \mid\mid \mathcal{H}(\Xi(\mu_1, \mu_2; \gamma); c\uparrow))$ are $\mu_1^*(c) = \mu_1^\bullet$ and*

$$\mu_2^*(c) = \mu_2^\bullet + \gamma\left(\mu_1^\bullet - \mathbb{E}\left[X_1 | X_1 \leq c\right]\right).$$

*So $\mu_2^*(c)$ is strictly increasing in $c$.*

*Proof.* Applying Definition 5, we see that $D_{KL}(\mathcal{H}^\bullet(c\uparrow) \mid\mid \mathcal{H}(\Xi(\mu_1, \mu_2; \gamma); c\uparrow))$, the KL divergence in the special case of a cutoff strategy $c\uparrow$, is

$$\int_c^\infty \phi(x_1; \mu_1^\bullet, \sigma^2) \cdot \ln\left(\frac{\phi(x_1; \mu_1^\bullet, \sigma^2)}{\phi(x_1; \mu_1, \sigma^2)}\right) dx_1$$
$$+ \int_{-\infty}^c \left\{\int_{-\infty}^\infty \phi(x_1; \mu_1^\bullet, \sigma^2) \cdot \phi(x_2; \mu_2^\bullet, \sigma^2) \cdot \ln\left[\frac{\phi(x_1; \mu_1^\bullet, \sigma^2) \cdot \phi(x_2; \mu_2^\bullet, \sigma^2)}{\phi(x_1; \mu_1, \sigma^2) \cdot \phi(x_2; \mu_2 + \gamma(x_1 - \mu_1), \sigma^2)}\right] dx_2\right\} dx_1.$$

Rewrite this as

$$\begin{aligned}
&\int_c^\infty \phi(x_1; \mu_1^\bullet, \sigma^2) \cdot \ln\left(\frac{\phi(x_1; \mu_1^\bullet, \sigma^2)}{\phi(x_1; \mu_1, \sigma^2)}\right) dx_1 \\
&+ \int_{-\infty}^c \phi(x_1; \mu_1^\bullet, \sigma^2) \cdot \int_{-\infty}^\infty \phi(x_2; \mu_2^\bullet, \sigma^2) \cdot \ln\left[\frac{\phi(x_1; \mu_1^\bullet, \sigma^2)}{\phi(x_1; \mu_1, \sigma^2)}\right] dx_2 dx_1 \\
&+ \int_{-\infty}^c \phi(x_1; \mu_1^\bullet, \sigma^2) \cdot \int_{-\infty}^\infty \phi(x_2; \mu_2^\bullet, \sigma^2) \cdot \ln\left[\frac{\phi(x_2; \mu_2^\bullet, \sigma^2)}{\phi(x_2; \mu_2 + \gamma(x - \mu_1), \sigma^2)}\right] dx_2 dx_1
\end{aligned}$$

which is:

$$\begin{aligned}
&\int_{-\infty}^\infty \phi(x_1; \mu_1^\bullet, \sigma^2) \cdot \ln\left(\frac{\phi(x_1; \mu_1^\bullet, \sigma^2)}{\phi(x_1; \mu_1, \sigma^2)}\right) dx_1 \\
&+ \int_{-\infty}^c \phi(x_1; \mu_1^\bullet, \sigma^2) \cdot \int_{-\infty}^\infty \phi(x_2; \mu_2^\bullet, \sigma^2) \ln\left[\frac{\phi(x_2; \mu_2^\bullet, \sigma^2)}{\phi(x_2; \mu_2 + \gamma(x_1 - \mu_1), \sigma^2)}\right] dx_2 dx_1
\end{aligned}$$

The KL divergence between $\mathcal{N}(\mu_{\text{true}}, \sigma^2_{\text{true}})$ and $\mathcal{N}(\mu_{\text{model}}, \sigma^2_{\text{model}})$ is

$$\ln\frac{\sigma_{\text{model}}}{\sigma_{\text{true}}} + \frac{\sigma^2_{\text{true}} + (\mu_{\text{true}} - \mu_{\text{model}})^2}{2\sigma^2_{\text{model}}} - \frac{1}{2},$$

so we may simplify the first term and the inner integral of the second term:

$$\frac{(\mu_1 - \mu_1^\bullet)^2}{2\sigma^2} + \int_{-\infty}^c \phi(x_1; \mu_1^\bullet, \sigma^2) \cdot \left[\frac{\sigma^2 + (\mu_2 + \gamma(x_1 - \mu_1) - \mu_2^\bullet)^2}{2\sigma^2} - \frac{1}{2}\right] dx_1.$$

Dropping constant terms not depending on $\mu_1$ and $\mu_2$ and multiplying by $\sigma^2$, we get a simplified expression of the objective,

$$\xi(\mu_1, \mu_2) := \frac{(\mu_1 - \mu_1^\bullet)^2}{2} + \int_{-\infty}^{c} \phi(x_1; \mu_1^\bullet, \sigma^2) \cdot \left[ \frac{(\mu_2 + \gamma(x_1 - \mu_1) - \mu_2^\bullet)^2}{2} \right] dx_1$$

We have the partial derivatives by differentiating under the integral sign,

$$\frac{\partial \xi}{\partial \mu_2} = \int_{-\infty}^{c} \phi(x_1; \mu_1^\bullet, \sigma^2) \cdot (\mu_2 + \gamma(x_1 - \mu_1) - \mu_2^\bullet) dx_1$$

$$\begin{aligned} \frac{\partial \xi}{\partial \mu_1} &= (\mu_1 - \mu_1^\bullet) - \gamma \int_{-\infty}^{c} \phi(x_1; \mu_1^\bullet, \sigma^2) \cdot (\mu_2 + \gamma(x_1 - \mu_1) - \mu_2^\bullet) dx_1 \\ &= (\mu_1 - \mu_1^\bullet) - \gamma \frac{\partial \xi}{\partial \mu_2} \end{aligned}$$

By the first order conditions, at the minimum $(\mu_1^*, \mu_2^*)$, we must have:

$$\frac{\partial \xi}{\partial \mu_2}(\mu_1^*, \mu_2^*) = \frac{\partial \xi}{\partial \mu_1}(\mu_1^*, \mu_2^*) = 0 \Rightarrow \mu_1^* = \mu_1^\bullet$$

So $\mu_2^*$ satisfies $\frac{\partial \xi}{\partial \mu_2}(\mu_1^\bullet, \mu_2^*) = 0$, which by straightforward algebra shows

$$\mu_2^*(c) = \mu_2^\bullet + \gamma \left( \mu_1^\bullet - \mathbb{E}\left[ X_1 | X_1 \le c \right] \right).$$

□

To interpret, the period $t$ agents correctly estimate the mean of the early draw, but misperceive the mean of the late draw in a way that depends on the degree of gambler's fallacy bias, the true mean of the early draw, and the cutoff used by the previous generation.

The censoring effect leads to a more pessimistic estimate of the second-period fundamental for lower values of $c$. To understand the intuition, consider that when $c$ decreases, the average $X_1$ conditional on $X_2$ being uncensored in the same game also decreases. Returning to the example, if firms in last year's hiring cycle decreased their threshold for hiring the early-phase candidates, then those firms that nevertheless engaged in a second round of search must have gotten especially disappointing early candidates, that is they must have drawn very low values of $X_1$.

While objectively $X_2$ is independent of $X_1$, the agents' gambler fallacy reasoning leads them to think that better $X_2$ should follow worse $X_1$. So, biased agents think of the sample mean of uncensored $X_2$ as reflecting a combination of the second-period fundamental and a reversal effect based on how realizations of $X_1$ accompanying these uncensored $X_2$ deviate from the first-period fundamental. Holding fixed the true distribution of observed $X_2$ in the dataset, which is objectively independent of the censoring threshold $c$, the agents' inference about $\mu_2$ decreases as the $X_1$'s in histories with uncensored $X_2$ decrease. This is because the change leads agents to attribute a greater fraction of the $X_2$ sample mean to the reversal effect, thus making a more pessimistic inference about the unconditional mean of $X_2$.

*Remark* 3. The misinference comes from the gambler's fallacy, not from misunderstanding missing data. The biased learners understand that $X_2$ is censored when $X_1 \geq c$, and their estimation procedure takes this into account. Indeed, it is precisely this understanding that leads them astray in their inference. If the gambler fallacy agents also suffer from selection neglect, in the sense that they treat the history of each uncensored game as a pair consisting of a sample from $\mathcal{N}(\mu_1^\bullet, \sigma^2)$ together with an unrelated sample from $\mathcal{N}(\mu_2^\bullet, \sigma^2)$, they would then end up with the correct inferences about both fundamentals. I believe my learning environment is unlikely to evoke selection neglect, a psychology most likely to be present when the observed dataset contains does not contain reminders about selection.[9] By contrast, censoring is highly explicit in the datasets of histories in my model: the always-observed first-period draw is the criterion for history censoring, and a censoring indicator replaces each unobserved second-period draw. In Section 7.3, I study an extension where there is a fraction of agents in each generation who suffer from selection neglect. I find that the presence of selection neglecters moderate the pessimism in inference, but do not eliminate it completely. ♦

*Remark* 4. This result shows that the pseudo-true fundamentals have a method-of-moments interpretation. Suppose that instead of finding parameters $\mu_1^*, \mu_2^*$ to minimize the KL divergence between $\mathcal{H}^\bullet(c\uparrow)$ and $\mathcal{H}(\Xi(\mu_1, \mu_2; \gamma); c\uparrow)$, agents' inference procedure involves finding

[9] In Enke (2017)'s experiment on selection neglect, players (one human subject and five computer players following a mechanical rule) are asked to guess a "state of the world" based on the average of 6 private signals. Players are sorted into one of two groups based on whether their own private signal is high or low, then observe the signals of others in their group. In the baseline treatment, there is no reminder of the excluded data on the decision screen where subjects are shown the signals of others in the same group and asked to enter a guess. This treatment finds selection neglect. Another treatment where subjects are given a simple hint stating: "Also think about the computer players whom you do not communicate with!" reduces the number of selection neglecters by 60%. So I believe the much clearer reminders of selection in my environment should reduce the frequency of selection neglect even further.

Jehiel (2018) studies misperceived investment returns under selection neglect. In his model, each predecessor has a potential project and observes a private signal about the project's quality. Predecessors with high signals implement their projects. Agents in the current generation observe the pool of implemented projects, then generate their own signals about the qualities of these observed projects. These signals are independent of the actual private signals that the predecessors used for implementation decisions. Current agents infer the conditional quality given each signal using the empirical mean quality among past implemented projects generating the same signal. This is another environment where the dataset contains no hints about the existence of excluded data (the unimplemented projects) or the selection criterion (the private signals of predecessors). In fact, if datasets in Jehiel (2018)'s setting record the complete experience of the predecessors in their decision problems, as is the case in my history datasets, then the misinference result no longer holds.

$\mu_1^M, \mu_2^M \in \mathbb{R}$ so that $\mathcal{H}(\Xi(\mu_1^M, \mu_2^M; \gamma); c\uparrow)$ matches the observed distribution of histories $\mathcal{H}^\bullet(c\uparrow)$ in terms of two moments: the observed mean of first-period draws, and the observed mean of second-period draws.

Since history always reveals $X_1$ in every game, $\mu_1^M(c) = \mu_1^\bullet$ matches the mean of observed first-period draws. Under the pair of parameters $(\mu_1^\bullet, \hat{\mu}_2)$, the mean $X_2$ among uncensored histories will be $\mathbb{E}[\hat{\mu}_2 + \gamma(X_1 - \mu_1^\bullet)|X_1 \leq c] = \hat{\mu}_2 + \gamma(\mathbb{E}[X_1|X_1 \leq c] - \mu_1^\bullet)$. In the objective distribution of histories $\mathcal{H}^\bullet(c\uparrow)$, the mean of observed $X_2$ is $\mu_2^\bullet$, so we have

$$\mu_2^M(c) = \mu_2^\bullet + \gamma\left(\mu_1^\bullet - \mathbb{E}\left[X_1|X_1 \leq c\right]\right),$$

which is the same as $\mu_2^*(c)$. ♦

*Remark* 5. The inference about $\mu_1^\bullet$ is exactly correct. Agents can rationalize the lack of expected reversals between the two periods in two ways: either the second period fundamental is low (so that the observed $X_2$ are in fact above second-period mean), or the first period fundamental is low (so that the rejected early draws are not much below first period mean and not much improved luck is "due" in the second period). One might imagine that slightly distorting belief about $\mu_1^\bullet$ downwards can help fit second-period data better at the expense of a small cost in fitting first-period data. The intuition here is that the first-period data is always observed while the second-period data is only sometimes observed, so this kind of distortion always leads to a worse overall fit. ♦

### 3.3 Inference about Fundamentals under the Constraint $\mu_1^\bullet = \mu_2^\bullet$

I now consider the natural special case where the true fundamentals are time-invariant, $\mu_1^\bullet = \mu_2^\bullet = \mu^\bullet \in \mathbb{R}$. If agents have a full-support prior belief over the feasible subjective models $\{\Xi(\mu_1, \mu_2; \gamma) : \mu_1, \mu_2 \in \mathbb{R}\}$ as before, then Proposition 1 continues to apply. But now suppose agents know the fundamentals are time-invariant and only have uncertainty over this common value. Formally, this means that agents' prior belief about the fundamentals is supported on the diagonal $\{(x, x) : x \in \mathbb{R}\}$, instead of having full-support on $\mathbb{R}^2$. This induces a prior belief supported on the *constrained feasible subjective models*, $\{\Xi(\mu, \mu; \gamma) : \mu \in \mathbb{R}\}$.

Let $\mu^*(c) \in \mathbb{R}$ stand for the fundamental that minimizes the KL divergence of the observation under this dogmatic belief, that is

$$\mu_{12}^*(c) := \underset{\mu \in \mathbb{R}}{\arg\min}\ D_{KL}(\mathcal{H}^\bullet(c\uparrow) \mid\mid \mathcal{H}(\Xi(\mu, \mu; \gamma); c\uparrow))$$

The next lemma characterizes $\mu_{12}^*(c)$.

**Lemma 3.** *$\mu^*_{12}(c) = \frac{1}{1+\mathbb{P}[X_1 \le c]\cdot(1-\gamma)^2}\mu^\circ_1(c) + \frac{\mathbb{P}[X_1 \le c]\cdot(1-\gamma)^2}{1+\mathbb{P}[X_1 \le c]\cdot(1-\gamma)^2}\mu^\circ_2(c)$, where $\mu^\circ_1(c) = \mu^\bullet$ and $\mu^\circ_2(c) = \mu^\bullet + \frac{\gamma}{1-\gamma}(\mu^\bullet - \mathbb{E}[X_1|X_1 \le c])$.*

The learner faces two kinds of data: observations of first-period draws and observations of second-period draws. Subjective models $\Xi(\mu^\circ_1(c), \mu^\circ_1(c); \gamma)$ and $\Xi(\mu^\circ_2(c), \mu^\circ_2(c); \gamma)$ minimize the KL divergence of these two kinds of data, respectively.[10]

The overall KL divergence minimizing estimator is a certain convex combination between these two points. Through the term $\mathbb{P}[X_1 \le c]$, the relative weight given to $\mu^\circ_2(c)$ increases as the cutoff $c$ increases, because the second-period data is observed more often if previous agents have used a more stringent cutoff in the first period.

We have $\mu^\circ_2(c) < \mu^\bullet$ while $\mu^\circ_1(c) = \mu^\bullet$, which shows that for any cutoff $c$ that the previous generation may have used, the next generation underestimates the fundamental.

Compared with Proposition 1's result about pseudo-true fundamentals without the equality constraint across periods, we have $\mu^\circ_2(c) > \mu^*_2(c)$ since $|\frac{\gamma}{1-\gamma}| < |\gamma|$, hence $\mu^*_{12}(c) > \mu^*_2(c)$. In the setting where agents start with a dogmatic belief that the fundamentals are identical in both periods, their beliefs about second-period fundamental end up less pessimistic relative to agents who can flexibly estimate different $\mu_1$ and $\mu_2$ for the two periods.

In general, $\mu^*_{12}(c)$ does not always increase in $c$. This is because decreasing the censoring threshold $c$ now has two competing effects. First, similar to the intuition of Proposition 1, a lower acceptance threshold $c$ leads the gambler's fallacy agents to expect greater reversal towards better-than-average draws in the second period, conditional on first draw falling below the threshold. Given the distribution of $X_2$ is in fact independent of $X_1$, a lower $c$ therefore leads to a more pessimistic second-period fundamental $\mu^\circ_2(c)$. But at the same time, a lower $c$ decreases the relative weight given to $\mu^\circ_2(c)$ rather than $\mu^\circ_1(c)$, since the second-period data is observed less frequently and so carries less weight in the overall divergence minimizing procedure.

Indeed, Figure 1 shows, the effect of cutoff $c$ on inference about the fundamental is in general non-monotonic.

[10] Note that $\mu^\circ_2(c)$ differs from the pseudo-true fundamental $\mu^*_2(c)$ from Proposition 1. The estimator $\mu^\circ_2(c)$ minimizes the KL divergence of second-period draws under the constraint that the same fundamental must be inferred for both periods, whereas $\mu^*_2(c)$ minimizes this divergence when first-period fundamental is fixed at its true value, $\mu^\bullet_1$.

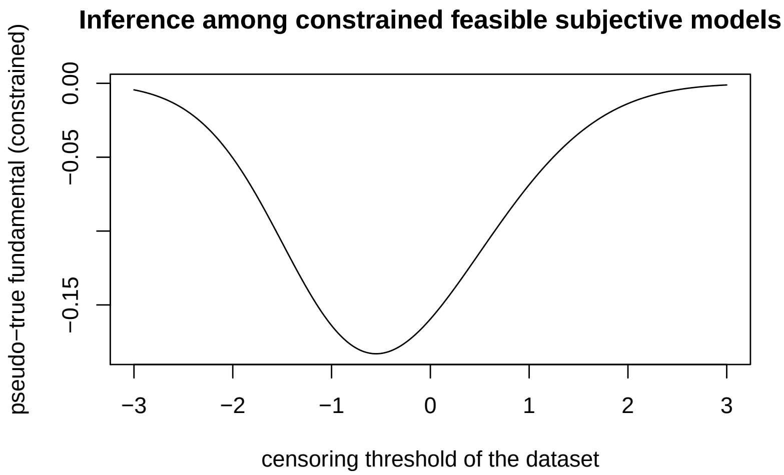


Figure 1: The pseudo-true fundamentals in constrained inference setting, with $\gamma = -0.5$ and $\mu^\bullet = 0$. The plot shows $\mu^*_{12}(c)$ for different values of stopping thresholds $c$.

# 4 Illustrative Example: Learning Dynamics in a Search Problem with $\mu_1^\bullet = \mu_2^\bullet$

To illustrate the main intuition of how the censoring effect interacts with the gambler's fallacy bias in a dynamic setting, I begin with a toy example involving a particularly simple optimal-stopping problem. I consider a simplified version of Example 1 where $q = 0$, that is a two-period search problem without recall. I suppose the two search periods have the same fundamental value $\mu_1^\bullet = \mu_2^\bullet = \mu^\bullet$, so the average candidate qualities in the early and late phases of the hiring season are equal.

In Section 4.1, I show that when agents (i.e. managers learning about the candidate qualities) start with any full-support prior beliefs about $\mu_1^\bullet$ and $\mu_2^\bullet$ and the 0th generation starts at the objectively optimal stopping strategy, a feedback loop emerges between distorted inferences and distorted acceptances thresholds. Agents' beliefs about the second-period fundamental and stopping rule monotonically drift away their objectively correct values, so that expected payoff is strictly decreasing across generations. In Section 4.2 I show that these learning dynamics are unchanged when agents' prior belief does not have full-support on $\mathbb{R}^2$, but reflects a (correct) dogmatic belief that the fundamentals are the same in the two periods. Finally, Section 4.3 shows the mislearning result relies crucially on the interaction between the censoring effect and the gambler's fallacy. Dropping either one of these elements from the example leads to the drastically different conclusion that agents have correct beliefs and play objectively optimal actions in every generation.

## 4.1 Feedback Loop and Monotonic Mislearning Across Generations

I first record the explicit expression for $C(\mu_1, \mu_2)$ in the search problem, which simply comes from rearranging the indifference condition.

**Lemma 4.** *In the two-period search problem given by the utility functions $u_1(x_1) = x_1$, $u_2(x_1, x_2) = x_2$, the cutoff is given by $C(\mu_1, \mu_2) = \frac{\mu_2 - \gamma\mu_1}{1-\gamma}$.*

In the setting of $\mu_1^\bullet = \mu_2^\bullet = \mu^\bullet$, let agents start with a full-support product prior belief $g : \mathbb{R}^2 \to \mathbb{R}_{++}$ about the fundamentals $(\mu_1, \mu_2)$. Suppose the 0th generation of agents start with the objectively optimal stopping strategy $c^\bullet \uparrow$ with $c^\bullet = \mu^\bullet$. Writing $\mu_{1,[t]}, \mu_{2,[t]}, c_{[t]}$ for the beliefs and stopping thresholds in generation $t$ for $t \geq 1$, I analyze the learning dynamics across generations.

By Proposition 1, the censoring effect leads agents in the first generation to infer that $\mu_{1,[1]} = \mu^\bullet$, $\mu_{2,[t]} < \mu_2^\bullet$. From Lemma 4, under the subjective model $\Xi(\mu_{1,[1]}, \mu_{2,[t]}; \gamma)$, generation 1 agents will revise their acceptance threshold to $c_{[t]} = \frac{\mu_{2,[t]} - \gamma\mu^\bullet}{1-\gamma} < \mu^\bullet$. This decrease in the cutoff rule between the 0th and 1st generation reflects the effect of distorted beliefs on behavior.

Importantly, the early stopping behavior of the 1st generation further distorts the beliefs of the next generation. By Proposition 1, the pseudo-true second-period fundamental is strictly increasing in the stopping threshold that generates the dataset of histories. This shows $\mu_2^*(c_{[1]}) < \mu_2^*(c_{[0]})$ since $c_{[1]} < c_{[0]}$. That is, as the second-generation agents face a more severe censoring effect, their beliefs end up even more pessimistic compared with the already distorted beliefs of the first-generation agents. Since the stopping threshold of Lemma 4 is strictly increasing in belief about the second-period fundamental, we conclude $c_{[2]} < c_{[1]}$. Figure 2 plots the beliefs and cutoff thresholds for generations 0 through 4 when $\gamma = -0.5$, $\mu^\bullet = 0$.

This feedback cycle between distorted stopping rule and distorted beliefs continues into all future generations, which I summarize in the next Proposition.

**Proposition 2.** *When the stage game is search without recall and $\mu_1^\bullet = \mu_2^\bullet = \mu^\bullet$, suppose the 0th generation starts with the objectively optimal stopping strategy $c^\bullet \uparrow$. The sequence of cutoff thresholds across different generations $(c_{[t]})_{t\geq 0}$ is strictly decreasing in $t$. So expected welfare is also strictly decreasing in $t$.*

Even though the 0th generation agents start at the objectively optimal play, the dynamics of learning across generations traces out a downward spiral that moves further and further away from it. The positive feedback between inference and behavior ensures that the mistake of the $t$-th generation is not corrected by agents in generation $t + 1$, but leads to further mislearning about the fundamental and further distortion away from the optimal behavior.

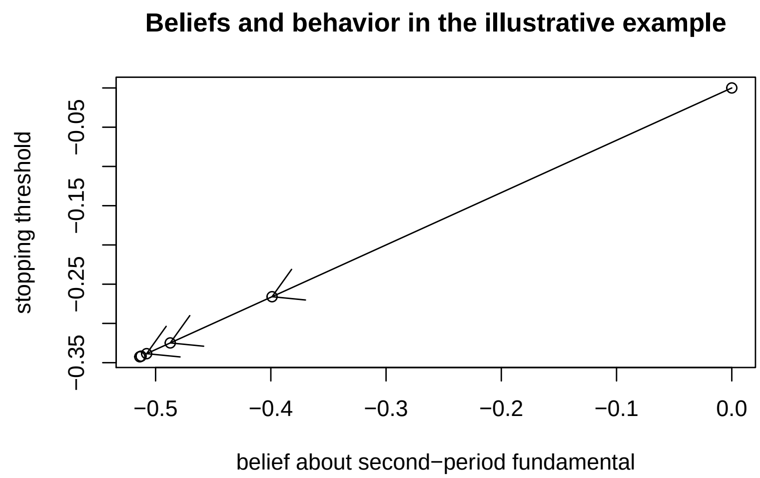


Figure 2: Beliefs $\mu_{2,[t]}$ and stopping thresholds $c_{[t]}$ of generations $1 \leq t \leq 5$, in search without recall with $\mu_1^\bullet = \mu_2^\bullet = 0$ and $\gamma = -0.5$.

## 4.2 When Agents Know That $\mu_1^\bullet = \mu_2^\bullet$

In Section 4.1, even though objectively $\mu_1^\bullet = \mu_2^\bullet = \mu^\bullet$, the agents are allowed to flexibly infer different values for these two fundamentals. Indeed, we have seen that they correctly infer $\mu_1^* = \mu^\bullet$ each period but under-infer the second-period fundamental. I now show the monotonic mislearning result obtained above not an artifact of this assumption.

Suppose agents know that $\mu_1^\bullet = \mu_2^\bullet$ and have a full-support belief over the constrained feasible subjective models, $\{\Xi(\mu, \mu; \gamma) : \mu \in \mathbb{R}\}$. As in Section 3.3, agents put full confidence in the pseudo-true fundamental

$$\mu_{12}^*(c) = \arg\min_{\mu \in \mathbb{R}} D_{KL}(\mathcal{H}^\bullet(c\uparrow) \;||\; \mathcal{H}(\Xi(\mu, \mu; \gamma); c\uparrow))$$

after observing the dataset of histories $\mathcal{H}^\bullet(c\uparrow)$.

A challenge in establishing the monotonic mislearning result in this setting is that Lemma 3 shows $\frac{d\mu_{12}^*}{dc}$ is not everywhere positive. However, it turns out $\frac{d\mu_{12}^*}{dc}$ is only negative for certain moderately negative values of $c$ that will not be visited for a society starting at $c_{[0]} = c^\bullet$. So under assumptions analogous to those in Proposition 2, we again get the harmful learning pattern that leads to worse welfare every generation.

**Proposition 3.** *When the stage game is search without recall, suppose agents know that the unknown fundamentals satisfy $\mu_1^\bullet = \mu_2^\bullet = \mu^\bullet$ and the 0th generation starts with the objectively optimal stopping strategy $c^\bullet \uparrow$. The sequence of cutoff thresholds across different generations $(c_{[t]})_{t\geq 0}$ is strictly decreasing in $t$. So expected welfare is also strictly decreasing in $t$.*

## 4.3 Turning Off the Censoring Effect or the Gambler's Fallacy

I now study some modifications of the baseline model to show that this harmful learning results emerges from the interaction of the gambler's fallacy and endogenous learning in the dynamic stage game. The results in this section imply that if the agents do not suffer from the gambler's fallacy, then they will learn to use the objectively optimal cutoff even in the presence of endogenous feedback (Proposition 4). I also consider a world where the generation $t+1$ agents observe what $X_2$ would have resulted in each of the previous generation's games, even those games where the early candidate was hired. Then even in the presence of the gambler's fallacy, the agents will learn to use the objectively optimal cutoff (Proposition 5). These results show that neither the censoring effect nor the gambler's fallacy is dispensable for the mislearning dynamics.

**Proposition 4.** *If $\gamma = 0$, then under the same assumptions as Proposition 2, $c_{[1]} = c^\bullet$ for all $t \geq 1$.*

*Proof.* The objective model for $(X_1, X_2)$ is within the class of feasible subjective models of the agents when $\gamma = 0$. Regardless of the cutoff $c_{[t-1]} \in \mathbb{R}$ used by the previous generation, only the objective model matches the censored distribution of $(X_1, X_2)$ and sets the KL divergence to 0. So, $\mu_{2,[t]} = \mu_2^\bullet$ for all $t \geq 1$. This means $c_{[t]} = c^\bullet$ for all $t \geq 1$. □

Now I introduce a new observability assumption where agents in generation $t+1$ observe the draws that would have been realized in each period of each game in generation $t$, regardless of the actual stopping choices of the generation $t$ agents.

**Definition 6.** The **full history** of predecessor $n$ consists of the pair of values $h_n^f = (x_1, x_2)$, where $x_1$ is the realization of $X_1$ in $n$'s decision problem and $x_2$ is the value of $X_2$ that would have been realized had $n$ continued into the second period.

Suppose generation $t+1$ observe an infinite dataset $(h_n^f)_{n \in [0,1]}$ of full histories corresponding to the hypothetical draws associated with all games in generation $t$. This full-observability environment turns off the censoring effect, as the distribution of generation $t+1$'s dataset is not affected by the stopping rule that generation $t$ agents use. I now show KL divergence in the new full-observability environment is minimized by the true fundamentals, even though agents have a misspecified class of feasible subjective models.

**Proposition 5.** *Under any full-support prior about the fundamentals, $c_{[t]} = c^\bullet$ for all $t \geq 1$ under full observations.*

*Proof.* Under full observations, regardless of the cutoff used by the previous generation, the inferred fundamentals $\mu_1^*$ and $\mu_2^*$ minimize the full-observations KL divergence,

$$\int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} \phi(x_1; \mu_1^\bullet, \sigma^2) \cdot \phi(x_2; \mu_2^\bullet, \sigma^2) \cdot \ln \left[ \frac{\phi(x_1; \mu_1^\bullet, \sigma^2) \cdot \phi(x_2; \mu_2^\bullet, \sigma^2)}{\phi(x_1; \mu_1, \sigma^2) \cdot \phi(x_2; \mu_2 + \gamma(x_1 - \mu_1), \sigma^2)} \right] dx_2 \right\} dx_1.$$

Note this integral is independent of $c$.

Performing rearrangements similar to the proof of Proposition 1, we find that $\mu_1^*$ and $\mu_2^*$ must minimize the objective function:

$$\xi(\mu_1, \mu_2) := \frac{(\mu_1 - \mu_1^\bullet)^2}{2} + \int_{-\infty}^{\infty} \phi(x_1; \mu_1^\bullet, \sigma^2) \cdot \left[ \frac{(\mu_2 + \gamma(x_1 - \mu_1) - \mu_2^\bullet)^2}{2} \right] dx_1$$

The unique pair solving the first-order conditions is $\mu_1^* = \mu_2^* = \mu^\bullet$. This shows for all $t \geq 1$, we have $\mu_{1,[t]} = \mu_{2,[t]} = \mu_2^\bullet$. So by using Lemma 4, we conclude $c_{[t]} = \mu^\bullet$ for every $t \geq 1$. $\square$

# 5 General Learning Dynamics

In this section, I investigate general learning dynamics for any stopping problem satisfying Assumption 1. Proposition 6 shows the positive feedback dynamics discussed in Section 4's example always obtains under these minimal assumptions. I then define the steady state of the learning system and provide additional assumptions to ensure its existence and uniqueness. When these additional assumptions are satisfied, I give a complete characterization of the generational learning dynamics (Corollary 1), showing that from any prior $g$ society always converges to the same steady state monotonically. Finally, dropping the additional restrictions ensuring a unique steady state, I show in Proposition 8 that all steady-state stopping rules of the learning dynamics are strictly lower than the objectively optimal one. This parallels Section 4's search example where agents end up using suboptimally low stopping thresholds. The overall picture is that the results and intuitions of Section 4's example remain qualitatively robust in the general environment.

I begin by showing the positive feedback phenomenon at the heart of Section 4's example holds generally. Changes in beliefs across successive generations are amplified, not dampened, by the corresponding changes in behavior leading to changes in the severity of the censoring effect.

**Proposition 6.** *Consider any optimal-stopping problem satisfying Assumption 1 and initialize the 0th generation at any cutoff stopping strategy*[11] $c_{[0]} \uparrow$. *Then beliefs about second period fundamental* $(\mu_{2,[t]}^*)_{t\geq 1}$ *and the stopping thresholds* $(c_{[t]})_{t\geq 1}$ *form monotonic sequences.*

*Proof.* Suppose $\mu_{2,[2]}^* \geq \mu_{2,[1]}^*$. Under Assumption 1, Lemma 2 applies, so $C$ is strictly increasing in its second argument. This shows $c_{[2]} = C(\mu_1^\bullet, \mu_{2,[2]}^*) \geq C(\mu_1^\bullet, \mu_{2,[1]}^*) = c_{[1]}$. But

[11] Throughout I will assume that 0th generation agents start with a cutoff-based stopping strategy. Lemma 1 does not directly apply to them, since they have a non-degenerate prior belief over the feasible subjective models and not a dogmatic belief in one model as agents in later generations. Online Appendix OA 2 provides sufficient conditions on the prior $g$ to ensure the subjectively optimal stopping strategy in the 0th generation involves a stopping threshold.

by Proposition 1, $\mu_2^*(c)$ increases in $c$, so $\mu_{2,[3]}^* = \mu_2^*(c_{[2]}) \geq \mu_2^*(c_{[1]}) = \mu_{2,[2]}^*$. Continuing this argument shows that $(\mu_{2,[t]}^*)_{t\geq 1}$ is a monotonically increasing sequence. Since $C$ is strictly increasing in its second argument, $(c_{[t]})_{t\geq 1}$ must also form a monotonically increasing sequence.

Conversely if $\mu_{2,[2]}^* < \mu_{2,[1]}^*$, then the analogous arguments show that $(\mu_{2,[t]}^*)_{t\geq 1}$ and $(c_{[t]})_{t\geq 1}$ are monotonically decreasing sequences. $\square$

The key driving force behind the result is the pair of monotonicity results: the indifference threshold $C$ is strictly increasing in its second argument and the pseudo-true fundamental strictly increases in the censoring threshold. The first comes from conditions in Assumption 1, ensuring that the cutoff choice of each generation increases with their belief about the second period fundamental. The second is a consequence of the interaction between the gambler's fallacy and the censoring effect — higher thresholds used by the previous generation lead to less censored datasets for the present generation, hence more optimistic beliefs about the second-period fundamental given the gambler's fallacy reasoning.

Now I turn to the long-run implications of the generational learning model. I first define the steady state, which depends on the stage game and the extent of the gambler's fallacy, but not on the prior. I then prove its existence and uniqueness under some additional assumptions.

**Definition 7.** A **steady state** consists of fundamentals $\mu_1^\infty, \mu_2^\infty \in \mathbb{R}$ and a stopping strategy $c^\infty \uparrow$ such that:

1. the cutoff is rational for the subjective model $\Xi(\mu_1^\infty, \mu_2^\infty; \gamma)$, that is $c^\infty = C(\mu_1^\infty, \mu_2^\infty)$.
2. beliefs correspond to the pseudo-true fundamentals given the dataset $\mathcal{H}^\bullet(c^\infty \uparrow)$, that is $\mu_1^\infty = \mu_1^*(c^\infty)$ and $\mu_2^\infty = \mu_2^*(c^\infty)$.

Note that the steady-state belief is endogenously determined by the steady-state stopping strategy. Every steady state is an instance of Esponda and Pouzo (2016)'s Berk-Nash equilibrium for an agent whose prior is supported on the feasible subjective models.

I now present a restriction on the stage-game payoff functions to ensure the existence and uniqueness of a steady state.

**Assumption 2.** *For every $x_1, x_2 \in \mathbb{R}$ and $w > 0$,*

$$u_1(x_1) - u_1(x_1 - w) > u_2(x_1, x_2) - u_2(x_1 - w, x_2 - (1+\gamma)w)$$

Essentially, this is a stronger version of Assumption 1's Condition 2, which already implies that

$$u_1(x_1) - u_1(x_1 - w) \geq u_2(x_1, x_2) - u_2(x_1 - w, x_2)$$

The new assumption requires that the inequality still holds even after the second argument $u_2(x_1 - w, x_2)$ is lowered by $(1+\gamma)w$, which makes the right-hand side larger.

For instance, Assumption 2 is satisfied for Example 1 for any $q \in [0,1)$, $\gamma < 0$.[12]

In Example 2, Assumption 2 is satisfied when the depreciation of the prototype value is significant for entrepreneurs who choose to improve their early-stage startup, or when it is easier to invent a high market-value prototype than it is to improve an existing prototype. More precisely, suppose $v_1(x_1) = b_1 x_1$, $v_2(x_2) = b_2 x_2$ for $b_1, b_2 > 0$, then Assumption 2 holds whenever $b_1 \geq \frac{1+\gamma}{1-\alpha} b_2$. This condition is easier to satisfy if $\alpha \in (0,1)$ is small or when $b_1$ is large relative to $b_2$.

**Proposition 7.** *When $-1 < \gamma < 0$, a unique steady state exists under Assumptions 1 and 2.*

To prove the existence and uniqueness of steady state, I consider the following belief iteration map on the second-period fundamental,

$$\Upsilon(\mu_2) := \mu_2^\bullet + \gamma\left(\mu_1^\bullet - \mathbb{E}\left[X_1 | X_1 \leq C(\mu_1^\bullet, \mu_2)\right]\right).$$

For $t \geq 1$, if the current generation of agents believe in fundamental values $(\mu_{1,[t]}^*, \mu_{2,[t]}^*) = (\mu_1^\bullet, \mu_2)$, they would choose the cutoff $C(\mu_1^\bullet, \mu_2)$ and, by Proposition 1, the next generation of agents would come to believe in $\mu_{1,[t+1]}^* = \mu_1^\bullet$, $\mu_{2,[t+1]}^* = \mu_2^\bullet + \gamma\left(\mu_1^\bullet - \mathbb{E}\left[X_1 | X_1 \leq C(\mu_1^\bullet, \mu_2)\right]\right)$. In short, the dynamics of beliefs about second-period fundamentals across successive generations are given by iterates of $\Upsilon$.

Every fixed point $\hat{\mu}_2$ of $\Upsilon$, $\Upsilon(\hat{\mu}_2) = \hat{\mu}_2$, is part of a steady state $\mu_1^\infty = \mu_1^\bullet$, $\mu_2^\infty = \hat{\mu}_2$, $c^\infty = C(\mu_1^\bullet, \hat{\mu}_2)$. Conversely, in any steady state $(\mu_1^\infty, \mu_2^\infty, c^\infty)$ we have to have $\mu_1^\infty = \mu_1^\bullet$, $c^\infty = C(\mu_1^\infty, \mu_2^\infty) = C(\mu_1^\bullet, \mu_2^\infty)$, and $\mu_2^\infty = \mu_2^*(c^\infty) = \mu_2^\bullet + \gamma\left(\mu_1^\bullet - \mathbb{E}\left[X_1 | X_1 \leq C(\mu_1^\bullet, \mu_2^\infty)\right]\right)$. So we see $\mu_2^\infty$ associated with any steady state must be a fixed point of $\Upsilon$.

The main idea of the proof of Proposition 7 involves showing that $\Upsilon$ is a contraction mapping. While there is a positive feedback loop between the censoring effect and gambler's fallacy, it turns out overall gain around the loop is positive but less than 1 provided the

[12] To see this, observe that

$$u_1(x_1) - u_1(x_1 - w) = (x_1) - (x_1 - w) = w$$

while

$$\begin{aligned}
&u_2(x_1, x_2) - u_2(x_1 - w, x_2 - (1+\gamma)w) \\
=&[q \cdot \max(x_1, x_2) + (1-q) \cdot x_2] - [q \cdot \max(x_1 - w, x_2 - (1+\gamma)w) + (1-q) \cdot (x_2 - (1+\gamma)w)] \\
\leq& q \cdot (\max(x_1, x_2) - \max(x_1 - w, x_2 - w)) + (1-q) \cdot (1+\gamma)w \\
=& qw + (1-q)(1+\gamma)w < qw + (1-q)w = w.
\end{aligned}$$

additional restrictions in Assumption 2 hold. So, this feedback does not cause beliefs and actions to run off to infinity across generations.

As a corollary of Proposition 6, I now give a complete characterization of the generations-based learning dynamics. When stopping problem satisfies the additional Assumption 2, agents have bias $-1 < \gamma < 0$, and when the 0th generation starts with any stopping strategy $c_{[0]} \uparrow$, both beliefs and behavior converge monotonically to the unique steady-state values. Thus the 0th generation stopping threshold $c_{[0]}$ affects the direction of convergence and the short-run behavior, but not the long-run behavior of the society.

**Corollary 1.** *Suppose Assumptions 1 and 2 hold, $-1 < \gamma < 0$, and 0th generation starts with a cutoff stopping rule $c_{[0]} \uparrow$. If $\mu^*_{2,[1]}$ is larger than the unique steady-state belief $\mu_2^\infty$, then beliefs and stopping thresholds decrease monotonically across generations, $\mu^*_{2,[t]} \searrow \mu_2^\infty$, $c_{[t]} \searrow c^\infty$. If $\mu^*_{2,[1]}$ is smaller than $\mu_2^\infty$, beliefs and stopping thresholds increase monotonically across generations, $\mu^*_{2,[t]} \nearrow \mu_2^\infty$, $c_{[t]} \nearrow c^\infty$. The rate of convergence in beliefs is at least exponential in $t$, with $|\mu^*_{2,[t]} - \mu_2^\infty| \leq \frac{|\gamma|^t}{1-|\gamma|}|\mu^*_{2,[2]} - \mu^*_{2,[1]}|$.*

*Proof.* Under these restrictions, by Proposition 7 there is a unique steady-state belief about the second-period fundamental, $\mu_2^\infty$. Since the sequence of beliefs across generations $(\mu^*_{2,[t]})_{t\geq 1}$ are the $\Upsilon$-iterates of $\mu^*_{2,[1]}$, that is $\mu^*_{2,[t+1]} = \Upsilon^{(t)}(\mu^*_{2,[1]})$ for all $t \geq 1$, the contraction mapping property of $\Upsilon$ established in the proof of Proposition 7 shows that $\lim_{t\to\infty} \mu^*_{2,[t]} = \mu_2^\infty$. The monotonicity of this convergence was established in Proposition 6. Also, Lemma A.1 in the Appendix used in the proof of Proposition 7 shows that $\left|C(\mu_1^\bullet, \mu_2^{'}) - C(\mu_1^\bullet, \mu_2^{''})\right| \leq |\mu_2^{'} - \mu_2^{''}|$ for all $\mu_1^{'}, \mu_2^{''} \in \mathbb{R}$. This in particular implies $C$ is a continuous function of its second argument, so we may exchange the limit:

$$\lim_{t\to\infty} c_{[t]} = \lim_{t\to\infty} C(\mu_1^\bullet, \mu^*_{2,[t]}) = C(\mu_1^\bullet, \lim_{t\to\infty} \mu^*_{2,[t]}) = C(\mu_1^\bullet, \mu_2^\infty) = c^\infty.$$

Finally, the proof of Proposition 7 showed that $\Upsilon$ has a modulus of $|\gamma|$, so the rate of convergence in belief comes from a well-known property about contraction mappings with modulus $|\gamma|$. □

Finally, I compare the steady-state stopping threshold with the objectively optimal one. Recall that the objectively optimal stopping strategy in the stage game has a cutoff form, as Lemma 1 also applies to the objective model $\Xi^\bullet$ specifying independent $(X_1, X_2)$. Let this objectively optimal cutoff be $c^\bullet \in \mathbb{R}$. I show that $c^\bullet > c^\infty$ for every steady-state cutoff $c^\infty$. (This result does not depend on uniqueness of the steady state and applies to all steady states if there are multiple, allowing me to drop the additional restrictions in Assumption 2.) That is, the early-stopping phenomenon of the illustrative example is robust to general stopping problems in the sense that it continues to obtain in the long-run.

**Proposition 8.** *Every steady-state stopping threshold $c^{\infty}$ is strictly lower than the objectively optimal stopping threshold, $c^{\bullet}$.*

*Proof.* Suppose stopping strategy $c^{\infty} \uparrow$ is subjectively optimal under the steady-state beliefs $(\mu_1^{\bullet}, \mu_2^{\infty})$. Indifference at $c^{\infty}$ under the subjective model $\Xi(\mu_1^{\bullet}, \mu_2^{\infty}; \gamma)$ implies that

$$u_1(c^{\infty}) = \mathbb{E}_{\tilde{X}_2 \sim \mathcal{N}(\mu_2^{\infty} + \gamma(c^{\infty} - \mu_1^{\bullet}), \sigma^2)}[u_2(c^{\infty}, \tilde{X}_2)].$$

By the definition of steady state, $\mu_2^{\infty} = \mu_2^*(c^{\infty}) = \mu_2^{\bullet} + \gamma\left(\mu_1^{\bullet} - \mathbb{E}\left[X_1 | X_1 \leq c^{\infty}\right]\right)$. This means

$$\mu_2^{\infty} - \gamma(\mu_1^{\bullet} - c^{\infty}) < \mu_2^{\infty} - \gamma(\mu_1^{\bullet} - \mathbb{E}\left[X_1 | X_1 \leq c^{\infty}\right]) = \mu_2^{\bullet}$$

since $c^{\infty} > \mathbb{E}\left[X_1 | X_1 \leq c^{\infty}\right]$.

Therefore, $\mathcal{N}(\mu_2^{\infty} + \gamma(c^{\infty} - \mu_1^{\bullet}), \sigma^2)$ is first-order stochastically dominated by $\mathcal{N}(\mu_2^{\bullet}, \sigma^2)$. Since $u_2$ is strictly increasing in its second argument by Assumption 1(a), we therefore have

$$u_1(c^{\infty}) < \mathbb{E}_{\tilde{X}_2 \sim \mathcal{N}(\mu_2^{\bullet}, \sigma^2)}[u_2(c^{\infty}, \tilde{X}_2)].$$

The LHS is the objective payoff of stopping at $c^{\infty}$ while the RHS is the objective expected payoff of continuing at $c^{\infty}$. By the structure of the optimal stopping rule under the objective model $\Xi^{\bullet}$, we must have $c^{\infty} < c^{\bullet}$. □

In the early generations, the comparison between the stopping behavior of biased agents and the optimal behavior is ambiguous. Given the correct beliefs about the fundamentals, a biased agent has a stronger incentive to continue after a below-average first-period draw than a rational agent. If the objectively optimal stopping rule involves a stopping threshold below the first-period fundamental, then this force pushes gambler's fallacy agents to use a higher stopping threshold than optimal. The content of this result is that in the long-run, society's pessimism about second-period fundamental always dominates and leads to the unambiguous prediction of stopping too early.

The intuition of this result is the clearest if we return to the illustrative example from Section 4. Under subjective model $\Xi$, agents choose a stopping threshold $c$ to optimize the wrong expected utility function $U_{\Xi}(c) := \mathbb{P}[X_1 \geq c] \cdot \mathbb{E}[X_1 | X_1 \geq c] \cdot + \mathbb{P}[X_1 < c] \cdot \mathbb{E}_{\Xi}[X_2 | X_1 < c]$. The expression $U_{\Xi}$ correctly describes the probability of hiring the early-phase candidate and the expected quality of the early candidate conditional on being hired, but is misspecified as to how the stopping threshold choice affects the quality of the late-phase candidate conditional on not hiring the first one. Whereas in reality $\mathbb{E}[X_2 | X_1 < c] = \mu_2^{\bullet}$ does not depend on $c$, the biased agents believe $\mathbb{E}_{\Xi}[X_2 | X_1 < c]$ decreases with $c$. Outside of the steady state, the agent's expected continuation payoff given that the early-phase candidate falls below some threshold $c$ is unrestricted by theory, so we can say little about how the biased agents behave relative to their rational counterpart. But in the steady state, we must have $\mathbb{E}_{\Xi(\mu_1^\bullet,\mu_2^\infty;\gamma)}[X_2|X_1 < c^\infty] = \mu_2^\bullet$, combining the steady-state condition $\mu_2^*(c^\infty) = \mu_2^\infty$ with the method-of-moments interpretation of the pseudo-true $\mu_2^\infty$. So the biased agents correctly know the expected second-period payoff if they use the steady-state cutoff $c^\infty$, but wrongly believe that increasing this cutoff will lead to a worse second-period payoff, under-estimating the expected benefits of a choosier stopping threshold. If the biased agents are indifferent at $X_1 = c^\infty$, rational agents must strictly prefer to increase the threshold.

The above discussion shows a connection between the structure of the steady state in this learning problem and Esponda (2008)'s behavioral equilibrium. In Esponda's world, buyers in a bilateral trade situation offer a price $p$, which sellers with different quality goods accept or reject. The buyer correctly knows the expected quality of the trade conditional on a seller accepting the price $p$, but holds wrong beliefs about the quality consequences of a deviation. Esponda (2008) had no explicit mechanism for how these beliefs are formed, but in my learning problem these deviation beliefs are pinned down by the gambler's fallacy.

# 6 Fictitious Variation and Censored Datasets

So far, I have considered agents who hold dogmatic and correct beliefs about the variance of $X_1$ and the conditional variance of $X_2|(X_1 = x_1)$. In this section, I turn to agents who are uncertain about the variances of the draws and jointly estimate variance and fundamentals using the histories of their predecessors.

Objectively, $X_1, X_2$ are independent Gaussian random variables each with a variance of $(\sigma^\bullet)^2 > 0$, so the true joint distribution of $(X_1, X_2)$ is $\Xi^\bullet = \Xi(\mu_1^\bullet, \mu_2^\bullet, (\sigma^\bullet)^2, (\sigma^\bullet)^2; 0)$. Suppose agents have a full-support belief over the class of models

$$\left\{\Xi(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2; \gamma) : \mu_1, \mu_2 \in \mathbb{R}, \sigma_1^2, \sigma_2^2 \geq 0\right\}.$$

Following Definition 5, $D_{KL}(\mathcal{H}^\bullet(c\uparrow)||\mathcal{H}(\Xi(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2; \gamma); c\uparrow))$ is given by

$$\int_c^\infty \phi(x_1; \mu_1^\bullet, (\sigma^\bullet)^2) \cdot \ln\left(\frac{\phi(x_1; \mu_1^\bullet, (\sigma^\bullet)^2)}{\phi(x_1; \mu_1, \sigma_1^2)}\right) dx_1 \tag{1}$$
$$+ \int_{-\infty}^c \left\{\int_{-\infty}^\infty \phi(x_1; \mu_1^\bullet, (\sigma^\bullet)^2) \cdot \phi(x_2; \mu_2^\bullet, (\sigma^\bullet)^2) \cdot \ln\left[\frac{\phi(x_1; \mu_1^\bullet, (\sigma^\bullet)^2) \cdot \phi(x_2; \mu_2^\bullet(\sigma^\bullet)^2)}{\phi(x_1; \mu_1, \sigma_2^2) \cdot \phi(x_2; \mu_2 + \gamma(x_1 - \mu_1), \sigma_2^2)}\right] dx_2\right\} dx_1.$$

The next proposition characterizes the pseudo-true parameters $\mu_1^*, \mu_2^*, (\sigma_1^*)^2, (\sigma_2^*)^2$ that minimize the above expression.

**Proposition 9.** *The pseudo-true parameters minimizing* $D_{KL}(\mathcal{H}^\bullet(c\uparrow)||\mathcal{H}(\Xi(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2;\gamma);c\uparrow))$ *are* $\mu_1^*=\mu_1^\bullet$, $\mu_2^*=\mu_2^\bullet+\gamma\left(\mu_1^\bullet-\mathbb{E}\left[X_1|X_1\leq c\right]\right)$, $(\sigma_1^*)^2=(\sigma^\bullet)^2$, $(\sigma_2^*)^2=(\sigma^\bullet)^2+\gamma^2\,Var[X_1|X_1\leq c]$.

Given any stopping rule $c\uparrow$, the agents' inferences about the fundamentals remain the same as in the case when they know the variances. Agents correctly estimate the first-period variance, $(\sigma_1^*)^2=(\sigma^\bullet)^2$, but their estimate of the second-period variance is too high. The magnitude of this distortion increases in the severity of the gambler's fallacy but decreases with the severity of the censoring, as $\mathrm{Var}[X_1|X_1\leq c]$ is smaller for lower $c$ when $X_1$ is Gaussian.

The intuition for misinferring the second-period conditional variance is the following. Whereas the objective conditional distribution of $X_2|(X_1=x_1)$ is independent of $x_1$, the agent has a different subjective model for this conditional distribution for each $x_1$. The agent's best-fitting belief about the second-period fundamental $\mu_2^*<\mu_2^\bullet$ ensures her subjective model about $X_2|X_1=x_1$ fits the data well following "typical" realizations of $x_1$ under the restriction $X_1\leq c$. However, following unusually high $X_1$ the agent is surprised by high values of $X_2$, while following unusually low $X_1$ she is surprised by low values of $X_1$. To better account for these surprising observations of $X_2$, the agent increases estimated conditional variance of $X_2|(X_1=x_1)$. The degree of overestimation increases in $\mathrm{Var}[X_1|X_1\leq c]$, for the frequency of these surprising observations depends on how much $X_1$ under the restriction $X_1\leq c$ tends to deviate from its typical value, $\mathbb{E}[X_1|X_1\leq c]$. And of course, the degree of overestimation increases in severity of the gambler's fallacy bias, which increases the size of these surprises.

An equivalent formulation of this result helps interpret the distorted $(\sigma_2^*)^2$. As in Remark 1, we may write the subjective model $\Xi(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2;\gamma)$, $\sigma_2^2=\sigma_1^2+\sigma_\eta^2$, $\sigma_\eta^2\geq 0$ as

$$X_1=\mu_1+\epsilon_1$$
$$X_2=\mu_2+\zeta+\epsilon_2$$

where $\epsilon_1\sim\mathcal{N}(0,\sigma_1^2)$, $\epsilon_2|\epsilon_1\sim\epsilon_2|\epsilon_1\sim\mathcal{N}(-\gamma\epsilon_1,\sigma_1^2)$, and $\zeta\sim\mathcal{N}(0,\sigma_\zeta^2)$, with $\zeta$ independent of $\epsilon_1,\epsilon_2$. In the context where $X_1$ and $X_2$ represent the quality realizations of two candidates from the early and late applicant pools, $\zeta$ is a vacancy-specific shift in the average quality of the second-period applicant. A positive $\sigma_\zeta^2$ means there are some vacancies for which the late applicants are an especially poor fit and some others for which they are especially suitable. Proposition 9 says that in an environment where all jobs are objectively homogeneous with respect to the fit of the late applicants, managers who find it possible that jobs are heterogeneous in this dimension will indeed estimate a positive amount of this heterogeneity, $\sigma_\zeta^2>0$, from the censored histories of their predecessors. This added heterogeneity allows agents to better rationalize histories $(X_1,X_2)$ where both candidates have unusually high/low qualities as vacancies that happen to be an especially good/bad fit for second-period applicants (i.e. the realization of $\zeta$ is far from 0.)

This phenomenon relates to findings in Rabin (2002) and Rabin and Vayanos (2010), who refer to this exaggeration of variance under the gambler's fallacy as *fictitious variation*. The key innovation of Proposition 9 is to show, in an endogenous learning setting, how the degree of fictitious variation depends on the severity of the censoring. To highlight this point, I now derive two results focusing on the interplay between fictitious variation and the endogenous censoring.

The first result says that when the second-period payoff $u_2(x_1, x_2)$ is a linear or convex function of $x_2$, the positive feedback cycle from Section 5 continues to obtain — cutoffs, beliefs about fundamentals, and beliefs about variances form monotonic sequences across generations. This includes the case of search with recall in Example 1 for any recall probability $0 \leq q < 1$.

**Definition 8.** The optimal-stopping problem is **convex** if for every $x_1 \in \mathbb{R}$, $x_2 \mapsto u_2(x_1, x_2)$ is convex with strict convexity for $x_2$ in a positive-measure set. The optimal-stopping problem is **concave** if for every $x_1 \in \mathbb{R}$, $x_2 \mapsto u_2(x_1, x_2)$ is concave with strict concavity for $x_2$ in a positive-measure set.

**Proposition 10.** *Suppose the optimal-stopping problem is convex. Suppose agents start with a full-support prior over* $\{\Xi(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2; \gamma) : \mu_1, \mu_2 \in \mathbb{R}, \sigma_1^2, \sigma_2^2 \geq 0\}$ *and let generation 0 use any cutoff stopping strategy* $c_{[0]} \uparrow$ *with* $c_{[0]} \in \mathbb{R}$. *For* $t \geq 1$, *denote the beliefs of generation* $t$ *as* $(\mu^*_{1,[t]}, \mu^*_{2,[t]}, (\sigma^*_{1,[t]})^2, (\sigma^*_{2,[t]})^2)$ *and their stopping strategy as* $c_{[t]} \uparrow$. *Then* $\mu^*_{1,[t]} = \mu^\bullet_1$, $(\sigma^*_{1,[t]})^2 = (\sigma^\bullet)^2$ *for all* $t$, *while* $(\mu^*_{2,[t]})_{t\geq 1}$, $(\sigma^*_{2,[t]})^2{}_{t\geq 1}$, *and* $(c_{[t]})_{t\geq 1}$ *are monotonic sequences.*

The intuition is straightforward. Suppose generation $t$ uses a more relaxed hiring threshold $c_{[t]} < c_{[t-1]}$ than generation $t-1$, resulting in a more severely censored dataset. By the usual censoring effect with known variances, generation $t+1$ becomes more pessimistic about second-period fundamental than generation $t$. In addition, by Proposition 9 we know that generation $t+1$ suffers less from fictitious variation than generation $t$. This implies generation $t+1$ agents would perceive less continuation value than generation $t$ agents even if they held the same beliefs about the fundamentals, for a larger variance in $X_2|(X_1 = x_1)$ improves the expected payoff when continuing after $X_1 = x_1$. Combining these two forces, we deduce $c_{[t+1]} < c_{[t]}$.

The second result compares the learning dynamics of two societies facing the same optimal-stopping problem. One society knows the correct variances of $X_1$ and $X_2|(X_1 = x_1)$. The other society is uncertain about the variances and infers them from data. Proposition 11 shows that in generation 1, the two societies hold the same beliefs about the means of the distributions, $\mu^\bullet_1$ and $\mu^\bullet_2$. But in all later generations $t \geq 2$, the society that must infer variances also end up with a more pessimistic/optimistic belief about the second-period fundamental compared with the society that knows the variances, provided the optimal-stopping problem is convex/concave. This divergence depends crucially on the endogenous-learning setting, for Proposition 9 implies that the two societies make the same inferences about the fundamentals when given the same dataset. But, since agents inferring variances end up believing in fictitious variation, they perceive a different continuation value than their peers in the same generation from the society that knows the variances. This causes the variance-inferring agents to use a different cutoff threshold, which affects the dataset that their successors observe. In short, allowing uncertainty on one dimension (variance) ends up affecting society's long-run inference in another dimension (mean).

Formally, consider two societies of agents, A and B. Agents in society A start with a full-support prior over $\{\Xi(\mu_1, \mu_2, (\sigma^\bullet)^2, (\sigma^\bullet)^2; \gamma) : \mu_1, \mu_2 \in \mathbb{R}\}$. Agents in society B start with a full-support prior over $\{\Xi(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2; \gamma) : \mu_1, \mu_2 \in \mathbb{R}, \sigma_1^2, \sigma_2^2 \geq 0\}$. Fix the same generation 0 cutoff stopping strategy $c_{[0]} \uparrow$ with $c_{[0]} \in \mathbb{R}$ for both societies. For $t \geq 1$, denote the beliefs of generation $t$ in society $k \in \{A, B\}$ as $(\mu^*_{1,[k,t]}, \mu^*_{2,[k,t]}, (\sigma^*_{1,[k,t]})^2, (\sigma^*_{2,[k,t]})^2)$ and their stopping strategy as $c_{[k,t]} \uparrow$.

**Proposition 11.** *In the first generation, $\mu^*_{1,[A,1]} = \mu^*_{1,[B,1]}$ and $\mu^*_{2,[A,1]} = \mu^*_{2,[B,1]}$. If the optimal-stopping problem is convex, then $\mu^*_{2,[B,t]} > \mu^*_{2,[A,t]}$ and $c_{[B,t]} > c_{[A,t]}$ for every $t \geq 2$. If the optimal-stopping problem is concave,[13] then $\mu^*_{2,[B,t]} < \mu^*_{2,[A,t]}$ and $c_{[B,t]} < c_{[A,t]}$ for every $t \geq 2$.*

# 7 Extensions of the Baseline Model

In this section I show that results of the baseline model are robust to a number of extensions. The Online Appendix OA 3 contains additional extensions.

## 7.1 Observing All Past Generations

As discussed in Section 2.4, an underlying assumption of the large-generations learning model is that generation $t$ agents only draw inference from the histories of generation $t-1$ agents, not from their actions. This amounts to not assuming the rationality of other players. When the stopping strategies associated with different generations converge, all late enough generations would find the strategy used by the immediate predecessor generation approximately optimal, given their own beliefs about the fundamentals. However, early generations may change their stopping thresholds by a non-negligible amount if they were to assume their predecessors

[13] For instance, the start-up problem in Example 2 is a concave optimal-stopping problem with $v_2(x_2) = x - e^{-x}$.

are rational and have the same payoff functions as they do, invert their predecessors' actions into their implied beliefs, then use this additional information in inferring the fundamentals.

I consider an extension where agents observe all stage-game histories from all previous generations (APG). In this extension, learning dynamics do not depend on whether agents know that others are rational and learn from their actions. Since generation $t_2$ observes all the histories that generation $t_1 < t_2$ saw, information sets are nested — generation $t_2$ agents can glean no additional information about the fundamentals from the strategy of generation $t_1$. In addition, generation $t_2$ would agree that generation $t_1$ played a subjectively optimal stopping strategy given their information.

To formally define the APG observation structure, write $h_{\tau,n}$ for the history of agent $n$ in generation $\tau$. Each agent in generation $t \geq 1$ observes an infinite dataset of histories $((h_{\tau,n})_{n\in[0,1]})_{\tau=0}^{t-1}$. If for each $0 \leq \tau \leq t-1$, generation $\tau$ agents the stopping strategy $c_\tau \uparrow$, then the distribution of each sub-dataset of histories $(h_{\tau,n})_{n\in[0,1]}$ is given by $\mathcal{H}^\bullet(c_\tau)$. This leads to the KL divergence objective to be minimized.

**Definition 9.** Under the APG observation structure, the **pseudo-true fundamentals** with respect to stopping rules $(c_\tau)_{\tau=0}^{t-1}$ are minimizers of

$$\sum_{\tau=0}^{t-1}\left(\begin{array}{c}\int_{c_\tau}^{\infty}\phi(x_1;\mu_1^\bullet,\sigma^2)\cdot\ln\left(\frac{\phi(x_1;\mu_1^\bullet,\sigma^2)}{\phi(x_1;\mu_1,\sigma^2)}\right)dx+\\ \int_{-\infty}^{c_k}\left\{\int_{-\infty}^{\infty}\phi(x_1;\mu_1^\bullet,\sigma^2)\cdot\phi(x_2;\mu_2^\bullet,\sigma^2)\cdot\ln\left[\frac{\phi(x_1;\mu_1^\bullet,\sigma^2)\cdot\phi(x_2;\mu_2^\bullet,\sigma^2)}{\phi(x_1;\mu_1,\sigma^2)\cdot\phi(x_2;\hat{\mu}_2+\gamma(x_1-\mu_1),\sigma^2)}\right]dx_2\right\}dx_1\end{array}\right) \tag{2}$$

across $\mu_1, \mu_2 \in \mathbb{R}$. Denote these pseudo-true fundamentals as $\mu_1^{\mathrm{A}}(c_0, ..., c_{t-1})$ and $\mu_2^{\mathrm{A}}(c_0, ..., c_{t-1})$.

The next lemma expresses the APG pseudo-true parameters in terms of the pseudo-true parameters in the baseline model. It shows that after observing data generated by cutoffs $c_0, ..., c_{t-1}$ in generations $0, \ldots\ t-1$, generation $t$'s posterior belief about the fundamentals can be computed by taking a weighted average of the $t$ different beliefs that the histories from each of the past $t$ generations would have generated on their own.

**Lemma 5.** *The pseudo-true parameters are given by* $\mu_1^A(c_0, ..., c_{t-1}) = \mu_1^\bullet$ *and*

$$\mu_2^A(c_0, ..., c_{t-1}) = \frac{1}{t \cdot \sum_{\tau=0}^{t-1}\mathbb{P}[X_1 \leq c_\tau]}\sum_{\tau=0}^{t-1}\mathbb{P}[X_1 \leq c_\tau]\cdot\mu_2^*(c_\tau),$$

*where* $\mu_2^*(c_\tau)$ *is the pseudo-true second-period fundamental associated with the dataset* $\mathcal{H}^\bullet(c_\tau \uparrow$ *).*

Intuitively speaking, generation $t$'s inference has to accommodate data from $t$ previous generations, which are censored using $t$ potentially different cutoffs. While the biased generation $t$ agents would draw the same inference about first-period fundamental using data from any of these past generation, the data of different past generations lead to different beliefs about second-period fundamental. The relative influence of generation $\tau$ data on the overall inference depends on the cutoff $c_\tau$ generating it, since this cutoff affects how many uncensored pairs $(X_1, X_2)$ are observed in this generation relative to other generations.

Next, I characterize the learning dynamics in the APG observations environment. I show that the positive feedback between distorted actions and distorted beliefs remains robust in this setting for general stopping problems. Furthermore, under the additional restrictions in Assumption 2 guaranteeing a unique steady state for the baseline model, APG learning converges to the same steady state. That is, whether agents observe only the histories of the immediate predecessor generation or all past generations has no effect on the long-run learning outcome. The intuition is that as beliefs converges across generations, the stopping thresholds converge as well. For agents in late enough generations, most of the histories in their dataset are censored according to stopping thresholds very similar to the limit threshold. So, the pseudo-true fundamental based on their APG dataset is similar to the pseudo-true fundamental of a one-generation dataset based on the limit cutoff, allowing us to compare inference under these two different observation structures.

While in the long-run both APG and the baseline models behave the same way, they can differ in their short-run welfare. For example, in settings where learning leads generations further and further astray from the objectively optimal strategy, the APG environment slows down this harmful learning, as the less-censored datasets from the early generations partially moderate the pessimistic inference about the second-period fundamental.

**Proposition 12.** *Suppose 0th generation starts with any cutoff stopping strategy* $c_{[0]} \uparrow$ *. Then in the APG observations environment, beliefs about second period fundamental* $(\mu^A_{2,[t]})_{t\geq 1}$ *and the stopping thresholds* $(c^A_{[t]})_{t\geq 1}$ *for generations* $t \geq 1$ *form monotonic sequences.*

*If in addition Assumption 2 holds and* $-1 < \gamma < 0$*, then* $(\mu^A_{2,[t]})_{t\geq 1}$ *and* $(c^A_{[t]})_{t\geq 1}$ *monotonically converge to* $\mu_2^\infty$ *and* $c^\infty$*, the belief and stopping threshold associated with the unique steady state of the baseline model.*

### 7.2 Uncertainty about $\gamma$

So far I have considered agents with a dogmatic belief in some $\gamma < 0$. Now I turn to the generalization where agents jointly estimate $\mu_1$, $\mu_2$, and $\gamma$ from data. While their prior beliefs about $\mu_1$ and $\mu_2$ have full support on $\mathbb{R}$, I assume their prior about $\gamma$ is supported on some finite interval $[\underline{\gamma}, \bar{\gamma}]$. The next result generalizes Proposition 1. It shows that when the draws $(X_1, X_2)$ have an objective distribution $\Xi(\mu_1^\bullet, \mu_2^\bullet; \gamma^\bullet)$ with $\gamma^\bullet \notin [\underline{\gamma}, \bar{\gamma}]$, the KL-divergence minimizing inference involves $\gamma^*$ equal to $\hat{\gamma} \in \{\underline{\gamma}, \bar{\gamma}\}$, boundary point of the support of $\gamma$ that is the closest to $\gamma^\bullet$. Given the estimated pseudo-true parameter $\hat{\gamma}$, the estimates of the first- and second-period fundamentals take similar forms to those in Proposition 1.

**Proposition 13.** *Suppose $\gamma^\bullet \notin [\underline{\gamma}, \bar{\gamma}]$. Let $\tilde{\gamma} = \bar{\gamma}$ if $\gamma^\bullet > \bar{\gamma}$, otherwise $\tilde{\gamma} = \underline{\gamma}$ when $\gamma^\bullet < \underline{\gamma}$. The pseudo-true parameters minimizing KL divergence*

$$\min_{\mu_1,\mu_2 \in \mathbb{R}, \gamma \in [\underline{\gamma}, \bar{\gamma}]} D_{KL}(\mathcal{H}(\Xi(\mu_1^\bullet, \mu_2^\bullet; \gamma^\bullet); c\uparrow) \mid\mid \mathcal{H}(\Xi(\mu_1, \mu_2; \gamma); c\uparrow))$$

*is given by $\mu_1^*(c) = \mu_1^\bullet$, $\mu_2^*(c) = \mu_2^\bullet + (\tilde{\gamma} - \gamma^\bullet) \cdot \left(\mu_1^\bullet - \mathbb{E}_{\Xi(\mu_1^\bullet, \mu_2^\bullet; \gamma^\bullet)}[X_1 | X_1 \le c]\right)$, $\gamma^*(c) = \tilde{\gamma}$.*

Intuitively, we may expect the closest distance (in the KL divergence sense) from the set of subjective models $\{\Xi(\mu_1, \mu_2; \hat{\gamma}) : \mu_1, \mu_2 \in \mathbb{R}\}$ to the objective distribution $\Xi(\mu_1^\bullet, \mu_2^\bullet; \gamma^\bullet)$ to decrease in $|\hat{\gamma} - \gamma^\bullet|$. Proposition 13 confirms this intuition, showing that the pseudo-true model from the set $\{\Xi(\mu_1, \mu_2; \gamma) : \mu_1, \mu_2 \in \mathbb{R}, \gamma \in [\underline{\gamma}, \bar{\gamma}]\}$ lies in the subset $\{\Xi(\mu_1, \mu_2; \gamma) : \mu_1, \mu_2 \in \mathbb{R}, \gamma = \tilde{\gamma}\}$, where $\tilde{\gamma}$ is the closest point (in the Euclidean sense) to $\gamma^\bullet$ in the interval $[\underline{\gamma}, \bar{\gamma}]$.

As an immediate corollary, when $X_1$ and $X_2$ are objectively uncorrelated ($\gamma^\bullet = 0$) and the agent's belief about $\gamma$ is supported on $[\underline{\gamma}, \bar{\gamma}]$ with by $\bar{\gamma} < 0$, the learning dynamics are exactly the same as in baseline the model where agent dogmatically believes in the value $\gamma = \bar{\gamma}$.

**Corollary 2.** *Suppose $\gamma^\bullet = 0$ and agents have a prior belief about the correlation that is supported on $[\underline{\gamma}, \bar{\gamma}]$ with $\bar{\gamma} < 0$. For any initial stopping strategy $c_{[0]}\uparrow$ in the 0th generation, all generations $t \ge 1$ believe with certainty that $\gamma = \bar{\gamma}$. The dynamics of of beliefs and behavior for generations $t \ge 1$ are identical to those in the baseline model when agents have a dogmatic belief in $\gamma = \bar{\gamma}$ and 0th generation starts with the same strategy $c_{[0]}\uparrow$.*

### 7.3 Population with Heterogeneity in Selection Neglect

In this section, I study an extension of the baseline model where a fraction $0 \le \alpha < 1$ of agents in each generation has full selection neglect, while the remainder are baseline agents with the gambler's fallacy. This mixture specification is inspired by Enke (2017)'s experimental results, who finds heterogeneity in subjects' degree of selection neglect with the full-selection-neglect subjects and no-selection-neglect subjects together accounting for a majority of the population. On the other hand, Enke (2017) does not find a significant mass of subjects at any "intermediate" level of selection neglect.

To model agents with full selection neglect, I assume that when faced with a dataset of histories $(h_{1,n}, h_{2,n})_{n \in [0,1]}$, they treat $(h_{1,n})_{n \in [0,1]}$ as a sample from the unconditional distribution of $X_1$, and $(h_{2,n})_{n: h_{2,n} \neq \varnothing}$ as an independent sample from the unconditional distribution of $X_2$. Relative to the base line agents, they make the error of of the selection process behind which $h_{2,n}$ appear in the dataset: they are not censored at random, but only censored when $h_{1,n}$ exceeds the acceptance threshold used by the predecessors. In this environment, the gambler's fallacy and selection neglect exactly cancel each other out, since in large datasets the mean of $h_{1,n}$ is $\mu_1^\bullet$ and the mean of uncensored $h_{2,n}$ is $\mu_2^\bullet$. This shows that from the dataset $\mathcal{H}^\bullet(c)$ for any $c \in \mathbb{R}$, the selection neglecters correctly infer the fundamentals and choose the stopping strategy with cutoff[14] $C(\mu_1^\bullet, \mu_2^\bullet)$.

Now consider a baseline agent with the gambler's fallacy, facing a dataset of histories generated by a heterogeneous population of predecessors. A fraction $\alpha$ of the histories are generated by selection neglecters using the stopping strategy $C(\mu_1^\bullet, \mu_2^\bullet) \uparrow$. The remaining $1-\alpha$ fraction are generated by baseline predecessors using the stopping strategy $c \uparrow$ . The next proposition characterizes the pseudo-true fundamentals maximizing the weighted-average KL-divergence objective,

$$\alpha D_{KL}(\mathcal{H}^\bullet(C(\mu_1^\bullet, \mu_2^\bullet))||\mathcal{H}(\Xi(\mu_1, \mu_2; \gamma); C(\mu_1^\bullet, \mu_2^\bullet))) + (1-\alpha) D_{KL}(\mathcal{H}^\bullet(c)||\mathcal{H}(\Xi(\mu_1, \mu_2; \gamma); c)). \tag{3}$$

The proof is similar to that of Proposition 5, except replacing multiple previous generations with two sub-populations within the immediate predecessor generation, and weighing these sub-populations differently due to their relative sizes.

**Proposition 14.** *The pseudo-true fundamentals minimizing Equation (3) when baseline predecessors use the stopping strategy $c \uparrow$ is $\mu_1^{SN} = \mu_1^\bullet$,*

$$\begin{aligned} \mu_2^{SN}(c) = & \frac{\alpha \mathbb{P}[X_1 \le C(\mu_1^\bullet, \mu_2^\bullet)]}{\alpha \mathbb{P}[X_1 \le C(\mu_1^\bullet, \mu_2^\bullet)] + (1-\alpha)\mathbb{P}[X_1 \le c]} \cdot \mu_2^*(C(\mu_1^\bullet, \mu_2^\bullet)) \\ & + \frac{(1-\alpha)\mathbb{P}[X_1 \le c]}{\alpha \mathbb{P}[X_1 \le C(\mu_1^\bullet, \mu_2^\bullet)] + (1-\alpha)\mathbb{P}[X_1 \le c]} \cdot \mu_2^*(c). \end{aligned}$$

That is, with a mixture of selection-neglecter and baseline predecessors, baseline agents' inference about the second-period fundamental is a convex combination between what they would infer from the histories of the selection neglecters alone and what they would infer from the histories of the baseline predecessors alone. The relative weights given to these two pseudo-true second-period fundamentals depend on the relative sizes of the two subpopulations, as well as on how frequently second-period draws are observed in each of the two sub-datasets.

Since both $\mu_2^*(C(\mu_1^\bullet, \mu_2^\bullet))$ and $\mu_2^*(c)$ are strictly below $\mu_2^\bullet$, we immediately conclude the same holds for $\mu_2^{SN}(c)$ for any $c \in \mathbb{R}$.

Next, I compare a baseline society with no selection neglecters with a second society containing a positive fraction of selection neglecters. I show that when two societies start with the same generation 0 behavior, the society with selection neglecters hold more optimistic

[14] This cutoff may nevertheless differ from the objectively optimal one, since the selection neglecters also suffer from the gambler's fallacy, so they believe in the joint distribution $\Xi(\mu_1^\bullet, \mu_2^\bullet; \gamma)$.

beliefs about the second-period fundamental and use a higher stopping threshold in every generation $t \geq 1$. So, the presence of a mixture of selection neglecters and basseline agents moderates the over-pessimism in inference without completely eliminating it.

**Corollary 3.** *Let $0 < \alpha < 1$. Consider two societies, 1 and 2, where society 1 has no selection neglecters and society 2 has an $\alpha$ fraction of selection neglecters in each generation. Suppose all agents in the 0th generation in both societies use the stopping rule $c_{[0]} \uparrow$. For $t \geq 1$, denote the baseline agents' beliefs and cutoff thresholds in society $k$ as $\mu^k_{1,[t]}, \mu^k_{2,[t]}, c^k_{[t]}$. Then for every $t \geq 2$, $\mu^2_{2,[t]} > \mu^1_{2,[t]}$ and $c^2_{[t]} > c^1_{[t]}$.*

# 8 A Finite-Population Foundation

In the analysis so far, I have assumed that each generation contains a continuum of agents. After observing an infinite dataset of histories with the distribution $\mathcal{H}^\bullet(c_{[t-1]})$, I assume that agents in each generation $t \geq 1$ place full confidence in pseudo-true fundamentals $(\mu_1^*(c_{[t-1]}), \mu_2^*(c_{[t-1]}))$ which solve the KL divergence minimization problem

$$\min_{\mu_1,\mu_2} D_{KL}(\mathcal{H}^\bullet(c_{[t-1]}) \mid\mid \mathcal{H}(\Xi(\mu_1, \mu_2; \gamma))\ ). \tag{4}$$

Agents in generation $t$ then use the optimal stopping strategy under the subjective model $(X_1, X_2) \sim \Xi(\mu_1^*(c_{[t-1]}), \mu_2^*(c_{[t-1]}))$.

The purpose of this section is to provide a finite-population foundation for this assumption. Suppose there are $N < \infty$ agents in each generation and agents in generation $t-1$ use the stopping strategy $c_{[t-1]} \uparrow$. I show that as $N$ grows large, the Bayesian posteriors of the generation $t$ agents about the fundamentals converge in mean the pseudo-true fundamentals solving (4). In addition, I also prove that for any $c' \in \mathbb{R}$, generation $t$'s posterior mean belief about the expected payoff to using stopping strategy $c' \uparrow$ converges to its expected payoff under $(\mu_1^*(c_{[t-1]}), \mu_2^*(c_{[t-1]}))$. In particular, if agents in generation $t$ contemplate between the stopping threshold $c^* = C(\mu_1^*(c_{[t-1]}), \mu_2^*(c_{[t-1]}))$ and finitely other alternatives, then as $N \to \infty$ they almost surely choosing $c^*$.

These results do not follow directly from the classical Berk (1966), because his result only establishes that for any open set containing the pseudo-true fundamentals, the mass that the posterior belief assigns to the open set almost surely converges to 1. Crucially, the prior distribution in my setting has full support on an unbounded domain, $(\mu_1, \mu_2) \in \mathbb{R}^2$. Indeed, one of the implications of my central inference result, Proposition 1, is that the pseudo-true parameter becomes unboundedly pessimistic as censoring threshold decreases. So, the weak mode of convergence in Berk (1966)'s conclusion leaves open the possibility that the posterior belief for increasing $N$ put decreasing mass on increasingly extreme values of $\mu_2$. If the magnitude of these extreme values grows more quickly in $N$ than the speed with which probability concentrates on the open set around the pseudo-true fundamentals, then payoff convergence could fail as $N \to \infty$. In practical terms, this implies the probability of generation $t$ agents playing the stopping strategy $c^* \uparrow$ as in my infinite-population model could be bounded away from 1 for all arbitrarily large $N$. Instead, the key Lemma 6 uses Bunke and Milhaud (1998) to derive the stronger convergence in mean that subsequently allows for convergence of payoffs and hence behavior as generations grow large.

To formalize the finite-population environment, suppose an agent has a full-support prior belief over the class of subjective models $\{\Xi(\mu_1, \mu_2, \sigma^2, \sigma^2; \gamma) : \mu_1, \mu_2 \in \mathbb{R}\}$ for some $\sigma^2 > 0$, $\gamma < 0$. Suppose the agent's prior over models is given by a density function with bounded magnitude, $g : \mathbb{R}^2 \to (0, B)$ for some $B < \infty$, over the fundamentals $(\mu_1, \mu_2)$. Let this agent observe $N$ pairs $(X_n, Y_n)_{n=1}^N$ generated in the following way: predecessor $n$ continues into the second period if and only if her first-period draw $X_{1,n}$ falls below some $c_{[t-1]} \in \mathbb{R}$. If $X_{1,n} < c_{[t-1]}$, then $X_n = X_{1,n}$ and $Y_n = X_{2,n}$ where $(X_{1,n}, X_{2,n})$ refers to pair of draws in $n$'s decision problem. If $X_{1,n} \geq c$, then $X_n = X_{1,n}$ but $Y_n \sim \mathcal{N}(0, 1)$ is a white noise term that is independent of the draws of any decision problem. The idea is that a censored draw is replaced by noise that is uninformative about the fundamentals. The generation $t$ agent knows the cutoff $c_{[t-1]}$ used to generate the dataset, so she knows which $Y_n$ terms are noise. The noise term is simply for notational convenience, allowing me to describe the joint distribution of $(X_n, Y_n)$ with a full support density.[15]

Since $N$ is finite, the agent will end up with a random, non-degenerate posterior density $\tilde{g}_N = g(\cdot|(X_n, Y_n)_{n=1}^N)$ about the fundamentals $(\mu_1, \mu_2)$, where the randomness comes from the randomness of the $N$ draws in the agent's finite sample. Lemma 6 shows that as $N \to \infty$, the random posterior $\tilde{g}_N$ converges to the pseudo-true parameters in mean.

**Lemma 6.** *Let $g : \mathbb{R}^2 \to (0, B)$ be a full-support prior density with bounded magnitude on the fundamentals. Fix a censoring threshold $c \in \mathbb{R}$ and write $\tilde{g}_N := g(\cdot|(X_n, Y_n)_{n=1}^N)$ for the random posterior after a censored sample of size $N$. Almost surely (with respect to the distribution of $(X_{1,n}, X_{2,n})_{n=1}^\infty$, the infinite sequence of $(X_1, X_2)$ pairs), we have*

$$\lim_{N \to \infty} \mathbb{E}_{(\mu_1, \mu_2) \sim \tilde{g}_N} \left( |\mu_1 - \mu_1^\bullet| + |\mu_2 - \mu_2^*(c)| \right) = 0.$$

Next, I turn to the convergence of expected payoffs for different cutoff strategies as sample size grows large. Let $g$ again satisfy the assumptions before and pairs $(X_n, Y_n)$ are still generated according to the censoring threshold $c \in \mathbb{R}$. For any $c' \in \mathbb{R}$ and $N \in \mathbb{N}$, let $U_N(c') := \mathbb{E}_{(\mu_1, \mu_2) \sim \tilde{g}_N} \left[ U(c'; \mu_1, \mu_2) \right]$ where $U(c; \mu_1, \mu_2)$ is the expected payoff of using

[15] The distribution of $Y_n$ can be replaced with any other distribution with a full-support density and all the results in this section will still go through.

the stopping strategy $c' \uparrow$ when $(X_1, X_2) \sim \Xi(\mu_1, \mu_2; \gamma)$. Note that $U_N(c')$ is a real-valued random variable representing agent's subjective expected payoff for the stopping strategy $c' \uparrow$, under the (random) non-degenerate posterior belief after observing a sample of size $N$. Proposition 15 shows that $U_N(c')$ converges almost surely to the subjective expected payoff of $c' \uparrow$ with a dogmatic belief in the pseudo-true fundamentals, provided the payoff functions $u_1, u_2$ of the optimal-stopping problem are Lipschitz continuous.

**Proposition 15.** *Suppose there are constants $K_1, K_2 > 0$ so that $|u_1(x_1') - u_1(x_1'')| < K_1 \cdot |x_1' - x_1''|$ and $|u_2(x_1', x_2') - u_2(x_1'', x_2'')| < K_2 \cdot (|x_1' - x_1''| + |x_2' - x_2''|)$ for all $x_1', x_1'', x_2', x_2'' \in \mathbb{R}$. Then for every $c' \in \mathbb{R}$, almost surely $U_N(c') \to U(c'; \mu_1^*(c), \mu_2^*(c))$.*

The Lipschitz continuity conditions are satisfied in the search problem (Example 1) for any $q \in [0, 1)$, as well as in the startup problem (Example 2) when initial value and later improvements are linear in draws, $v_1(x_1) = b_1 x_1 + a_1$ and $v_2(x_2) = b_2 x_2 + a_2$ for some $b_1, b_2 > 0$.

Suppose the agent contemplates between $c^*$ — the stopping threshold for generation $t$ agents after generation $t-1$ uses the stopping strategy $c_{[t-1]} \uparrow$ in the infinite-population model — and another threshold $c' \neq c^*$. It is an immediate corollary that the agent will choose $c^*$ with probability approaching 1 as sample size grows large.

**Corollary 4.** *Let $c^* = C(\mu_1^\bullet, \mu_2^*(c))$ and $c' \neq c^*$. Suppose an agent must choose between cutoff strategies $c^*$ or $c'$, based on expected payoff, after seeing the finite dataset $(X_n, Y_n)_{n=1}^N$ generated with the censoring threshold $X_n \geq c$. For every $\epsilon > 0$, there exists an $\underline{N} \in \mathbb{N}$ so that there is probability at least $1 - \epsilon$ the agent chooses $c^*$ whenever $N \geq \underline{N}$.*

# References


Bar-Hillel, M. and W. A. Wagenaar (1991): "The perception of randomness," *Advances in Applied Mathematics*, 12, 428–454.

Barron, G. and S. Leider (2010): "The role of experience in the gambler's fallacy," *Journal of Behavioral Decision Making*, 23, 117–129.

Bénabou, R. and J. Tirole (2002): "Self-confidence and personal motivation," *Quarterly Journal of Economics*, 117, 871–915.

Benjamin, D. J., D. A. Moore, and M. Rabin (2017): "Biased Beliefs About Random Samples: Evidence from Two Integrated Experiments," *Working Paper*.

Berk, R. H. (1966): "Limiting behavior of posterior distributions when the model is incorrect," *Annals of Mathematical Statistics*, 37, 51–58.

Bromiley, P. (2003): "Products and Convolutions of Gaussian Probability Density Functions," *Working Paper*.

Budescu, D. V. (1987): "A Markov model for generation of random binary sequences." *Journal of Experimental Psychology: Human Perception and Performance*, 13, 25.

Bunke, O. and X. Milhaud (1998): "Asymptotic behavior of Bayes estimates under possibly incorrect models," *Annals of Statistics*, 26, 617–644.

Bushong, B. and T. Gagnon-Bartsch (2018): "Learning with Misattribution of Reference Dependence," *Working Paper*.

Camerer, C. F. (1987): "Do biases in probability judgment matter in markets? Experimental evidence," *American Economic Review*, 77, 981–997.

Chen, D. L., T. J. Moskowitz, and K. Shue (2016): "Decision making under the gambler's fallacy: Evidence from asylum judges, loan officers, and baseball umpires," *Quarterly Journal of Economics*, 131, 1181–1242.

Croson, R. and J. Sundali (2005): "The gambler's fallacy and the hot hand: Empirical data from casinos," *Journal of Risk and Uncertainty*, 30, 195–209.

Enke, B. (2017): "What you see is all there is," *Working Paper*.

Esponda, I. (2008): "Behavioral equilibrium in economies with adverse selection," *American Economic Review*, 98, 1269–91.

Esponda, I. and D. Pouzo (2016): "Berk–Nash equilibrium: A framework for modeling agents with misspecified models," *Econometrica*, 84, 1093–1130.

Fudenberg, D., G. Romanyuk, and P. Strack (2017): "Active learning with a misspecified prior," *Theoretical Economics*, 12, 1155–1189.

Gagnon-Bartsch, T., M. Rabin, and J. Schwartzstein (2018): "Channeled Attention and Stable Errors," *Working Paper*.

Gumbel, E. J. (1960): "Bivariate exponential distributions," *Journal of the American Statistical Association*, 55, 698–707.

Heidhues, P., B. Koszegi, and P. Strack (2018): "Unrealistic expectations and misguided learning," *Econometrica*, 86, 1159–1214.

Jehiel, P. (2018): "Investment strategy and selection bias: An equilibrium perspective on overoptimism," *American Economic Review*, 108, 1582–97.

Narayanan, S. and P. Manchanda (2012): "An empirical analysis of individual level casino gambling behavior," *Quantitative Marketing and Economics*, 10, 27–62.

Pinelis, I. (2018): "Exact Bounds on the Inverse Mills Ratio and Its Derivatives," *Complex Analysis and Operator Theory*.

Rabin, M. (2002): "Inference by believers in the law of small numbers," *Quarterly Journal of Economics*, 117, 775–816.

Rabin, M. and D. Vayanos (2010): "The gambler's and hot-hand fallacies: Theory and applications," *Review of Economic Studies*, 77, 730–778.

Rapoport, A. and D. V. Budescu (1992): "Generation of random series in two-person strictly competitive games." *Journal of Experimental Psychology: General*, 121, 352.

Suetens, S., C. B. Galbo-Jørgensen, and J.-R. Tyran (2016): "Predicting lotto numbers: a natural experiment on the gambler's fallacy and the hot-hand fallacy," *Journal of the European Economic Association*, 14, 584–607.

Terrell, D. (1994): "A test of the gambler's fallacy: Evidence from pari-mutuel games," *Journal of Risk and Uncertainty*, 8, 309–317.

# Appendix

# A Omitted Proofs from the Main Text

## A.1 Proof of Claim 1

*Proof.* For Example 1, clearly $u_1$ and $u_2$ are strictly increasing functions of $x_1$ and $x_2$ respectively. We also have that $u_2(x_1^{'}, \bar{x}_2) - u_2(x_1^{''}, \bar{x}_2) \leq q(x_1^{'} - x_1^{''})$ for $x_1^{'} > x_1^{''}$ and any $\bar{x}_2$, while $u_1^{'}(x_1) = 1$. This shows Assumption 1(b) holds. Finally, if $x_1 > 0$ and $x_2 < 0$, then $u_2(x_1, x_2) = q \cdot x_1 + (1-q)x_2 \leq x_1 = u_1(x_1)$, and conversely $x_1 < 0, x_2 > 0$ imply $u_2(x_1, x_2) \geq u_1(x_1)$. This shows Assumption 1(c) holds.

For Example 2, Assumption 1(a) holds because $v_1, v_2$ are strictly increasing. Assumption 1(b) holds because for $x_1^{'} > x_1^{''}$ and any $\bar{x}_2$, $u_2(x_1^{'}, \bar{x}_2) - u_2(x_1^{''}, \bar{x}_2) = \alpha \cdot (u_1(x_1^{'}) - u_1(x_1^{''})) > 0$. Finally, $u_1(x_1) - u_2(x_1, x_2) = (1-\alpha)v_1(x_1) - v_2(x_2) + \kappa_2$. Since $v_1$ increases without bound as $x_1 \to \infty$, we can find a large enough $L_1 > 0$ so that $(1-\alpha)v_1(L_1) > v_2(0) - \kappa_2$. Then, for any $L \geq L_1$,

$$u_1(L) - u_2(L, -L) = (1-\alpha)v_1(L) - v_2(-L) + \kappa_2 \geq (1-\alpha)v_1(L_1) - v_2(0) + \kappa_2 > 0.$$

Also, since $v_2$ increases without bound as $x_2 \to \infty$, there is a large enough $L_2 > 0$ so that $(1-\alpha)v_1(0) - v_2(L_2) + \kappa_2 < 0$. Then for any $L \geq L_2$,

$$u_1(-L) - u_2(-L, L) = (1-\alpha)v_1(-L) - v_2(L) + \kappa_2 \leq (1-\alpha)v_1(0) - v_2(L_2) + \kappa_2 < 0.$$

Setting $L = \max(L_1, L_2)$ shows Example 2 satisfies Assumption 1(c). $\square$

## A.2 Proof of Lemma 1

*Proof.* I first show that if the agent is indifferent between stopping at some $\bar{x}_1$ and continuing, then he strictly prefers stopping at any $x_1^{'} > \bar{x}_1$. The indifference at $\bar{x}_1$ means that $u_1(\bar{x}_1) = \mathbb{E}[u_2(\bar{x}_1, \tilde{X}_2)]$ where $\tilde{X}_2 \sim \mathcal{N}(\mu_2 + \gamma(\bar{x}_1 - \mu_1), \sigma^2)$ is the conditional distribution $X_2|(X_1 = \bar{x}_1)$. The conditional distribution $X_2|(X_1 = x_1^{'})$ differs from $X_2|(X_1 = \bar{x}_1)$ by shifting the mean by $\gamma(x_1^{'} - \bar{x}_1)$. Since $u_2$ is strictly increasing in $x_2$ by Assumption 1(a), we get $\mathbb{E}[u_2(x_1^{'}, \tilde{X}_2 + \gamma(x_1^{'} - \bar{x}_1))] \leq \mathbb{E}[u_2(x_1^{'}, \tilde{X}_2)]$ seeing that $\gamma(x_1^{'} - \bar{x}_1) \leq 0$. Also, at any $x_2 \in \mathbb{R}$, by Assumption 1(b) we know that

$$u_1(x_1^{'}) - u_1(\bar{x}_1) > u_2(x_1^{'}, x_2) - u_2(\bar{x}_1, x_2).$$
$$\Rightarrow u_1(x_1^{'}) - u_2(x_1^{'}, x_2) > u_1(\bar{x}_1) - u_2(\bar{x}_1, x_2).$$

This then shows $u_1(x_1^{'}) - \mathbb{E}[u_2(x_1^{'}, \tilde{X}_2)] > u_1(\bar{x}_1) - \mathbb{E}[u_2(\bar{x}_1, \tilde{X}_2)]$. Combining these two facts with indifference at $\bar{x}_1$ gives us

$$u_1(x_1^{'}) - \mathbb{E}[u_2(x_1^{'}, \tilde{X}_2 + \gamma(x_1^{'} - \bar{x}_1))] \geq u_1(x_1^{'}) - \mathbb{E}[u_2(x_1^{'}, \tilde{X}_2)] > u_1(\bar{x}_1) - \mathbb{E}[u_2(\bar{x}_1, \tilde{X}_2)] = 0,$$

so stopping at $x_1^{'}$ is strictly preferable to continuing.

By an exactly symmetric argument, the agent also strictly prefers continuing at any $x_1 < \bar{x}_1$. So by continuity of $u_1, u_2$, the agent's optimal stopping strategy can only take 3 forms: either there is some threshold $c$ where he strictly prefers stopping for any $x_1 > c$ and strictly prefers continuing for any $x_1 < c$, or he strictly prefers to stop for all $x_1 \in \mathbb{R}$, or he strictly prefers to continue for all $x_1 \in \mathbb{R}$. I now use Assumption 1(c) to rule out these last two cases.

Under Assumption 1(c), there is some $L > 0$ so that $u_1(L) - u_2(L, -L) \geq 0$. Combining this with Assumption 1(a), $u_1(L) - u_2(L, -L-1) \geq h$ for some $h > 0$. Again by (a), $\lim_{d\to\infty} \left\{\mathbb{E}_{X_2 \sim \mathcal{N}(\mu_2 - d, \sigma^2)}[u_2(L, X_2)]\right\} \leq u_2(L, -L-1)$. That is to say, we can find some $\bar{d}$ so that $d \geq \bar{d}$ implies $\mathbb{E}_{X_2 \sim \mathcal{N}(\mu_2 - d, \sigma^2)}[u_2(L, X_2)] \leq u_2(L, -L-1) + h/2 \leq u_1(L) - h/2$. If $L \geq \mu_1 - \bar{d}/\gamma$, then $\mathbb{E}[X_2 | X_1 = L] \leq \mu_2 - \bar{d}$, which means agent strictly prefers stopping than continuing at $X_1 = L$. Otherwise, we just need to conclude that $u_1(\mu_1 - \bar{d}/\gamma) > \mathbb{E}_{X_2 \sim \mathcal{N}(\mu_2 - \bar{d}, \sigma^2)}[u_2(\mu_1 - \bar{d}/\gamma, X_2)]$. We do know that $L \leq \mu_1 - \bar{d}/\gamma$ and that $u_1(L) > \mathbb{E}_{X_2 \sim \mathcal{N}(\mu_2 - \bar{d}, \sigma^2)}[u_2(L, X_2)]$. So by Assumption 1(b), we get the desired inequality. This shows there is at least one value of $x_1$ such that $X_1 = x_1$ leads to strict preference for stopping.

Similar argument shows there is at least one value of $x_1$ such that $X_1 = x_1$ leads to strict preference for continuing. □

## A.3 Proof of Lemma 3

*Proof.* I start with the expression for the KL divergence from $\mathcal{H}^\bullet(c\uparrow)$ to $\mathcal{H}(\Xi(\mu, \mu; \gamma); c\uparrow)$. As in the proof of Proposition 1, this expression can be written as

$$\frac{(\mu - \mu^\bullet)^2}{2} + \int_{-\infty}^{c} \phi(x; \mu^\bullet, \sigma^2) \cdot \left[\frac{\sigma^2 + (\mu + \gamma(x_1 - \mu) - \mu^\bullet)^2}{2} - \frac{1}{2}\right] dx_1.$$

Dropping constant terms not depending on $\mu$, we get a simplified expression of the objective,

$$\xi(\mu) := \frac{(\mu - \mu^\bullet)^2}{2} + \int_{-\infty}^{c} \phi(x; \mu^\bullet, \sigma^2) \cdot \left[\frac{(\mu + \gamma(x_1 - \mu) - \mu^\bullet)^2}{2}\right] dx_1.$$

Taking the first-order condition, $\xi^{'}(\mu) = (\mu - \mu^\bullet) + (1-\gamma) \cdot \int_{-\infty}^{c} \phi(x_1; \mu^\bullet, \sigma^2) \cdot ((1-\gamma)\mu + \gamma x_1 - \mu^\bullet) dx_1$.

The term $\int_{-\infty}^{c} \phi(x_1; \mu^\bullet, \sigma^2) \cdot ((1-\gamma)\mu + \gamma x_1 - \mu^\bullet) dx_1$ may be rewritten as $\mathbb{P}[X_1 \leq c] \cdot$

$\mathbb{E}\left[(1-\gamma)\mu+\gamma X_1-\mu^\bullet|X_1\le c\right]$.

Setting the first-order condition to 0 and using straightforward algebra,

$$\mu_{12}^*(c)=\frac{1}{1+\mathbb{P}[X_1\le c]\cdot(1-\gamma)^2}\mu^\bullet+\frac{\mathbb{P}[X_1\le c]\cdot(1-\gamma)^2}{1+\mathbb{P}[X_1\le c]\cdot(1-\gamma)^2}\mu_2^\circ(c).$$

□

## A.4 Proof of Proposition 3

*Proof.* Write $\mu_{[t]}^*$ for the $t$'s generation's belief about the (common) fundamental value. By Lemma 4, under the subjective model $\Xi(\mu_{[t]}^*,\mu_{[t]}^*;\gamma)$, period $t$ managers will choose $c_{[t]}=\mu_{[t]}^*$.

By Lemma 3, $\mu_{[1]}^*<\mu^\bullet$, so $c_{[1]}<c_{[0]}$. Now assume we have $c_{[0]}>c_{[1]}>...>c_{[T]}$ for some $T\ge 1$. I will show that we also get $c_{[T+1]}<c_{[T]}$.

From the proof of Lemma 3, each $\mu_{[t]}^*$ for $t\ge 1$ minimizes the objective

$$\xi(\mu;c):=\frac{(\mu-\mu^\bullet)^2}{2}+\int_{-\infty}^{c}\phi(x_1;\mu^\bullet,\sigma^2)\cdot\left[\frac{(\mu+\gamma(x_1-\mu)-\mu^\bullet)^2}{2}\right]dx_1$$

with $c=c_{[t-1]}$. The objective $\xi$ has the derivatives:

$$\frac{\partial\xi}{\partial\mu}(\mu;c)=(\mu-\mu^\bullet)+(1-\gamma)\cdot\int_{-\infty}^{c}\phi(x_1;\mu^\bullet,\sigma^2)[(1-\gamma)\mu+\gamma x_1-\mu^\bullet]dx_1,$$

$$\frac{\partial^2\xi}{\partial\mu^2}(\mu;c)=1+(1-\gamma)^2\int_{-\infty}^{c}\phi(x_1;\mu^\bullet,\sigma^2)dx_1>0,$$

$$\frac{\partial^2\xi}{\partial c\partial\mu}(\mu;c)\propto(\mu-\mu^\bullet)-\gamma(\mu-c).$$

Since $\mu_{[T]}^*$ minimizes the objective $\xi(\cdot;c_{[T-1]})$, from the first-order condition we know $\frac{\partial\xi}{\partial\mu}(\mu_{[T]}^*;c_{[T-1]})=0$. To show that $\mu_{[T+1]}^*<\mu_{[T]}^*$, we need only establish that for any $\mu\ge\mu_{[T]}^*$, we have $\frac{\partial\xi}{\partial\mu}(\mu;c_{[T]})>0$, so by first-order condition the objective $\xi(\cdot;c_{[T]})$ cannot be minimized at any belief weakly more optimistic than the generation $T$ belief of $\mu_{[T]}^*$. Since we already have $\frac{\partial^2\xi}{\partial\mu^2}>0$ everywhere, it suffices to establish that $\frac{\partial^2\xi}{\partial c\partial\mu}(\mu_{[T]}^*;c)<0$ for all $c\in(c_{[T]},c_{[T-1]})$ (here $c_{[T]}<c_{[T-1]}$ by the inductive hypothesis). We have $\frac{\partial^2\xi}{\partial c\partial\mu}(\mu_{[T]}^*;c)\propto(\mu_{[T]}^*-\mu^\bullet)-\gamma(\mu_{[T]}^*-c)$, where $(\mu_{[T]}^*-\mu^\bullet)<0$ by Lemma 3 and also $\mu_{[T]}^*-c=c_{[T]}-c<0$ for $c$ in the range $(c_{[T]},c_{[T-1]})$. This shows the negativity of the cross partial derivative in the desired range, and establishes $\mu_{[T+1]}^*<\mu_{[T]}^*$ and $c_{[T+1]}<c_{[T]}$. Now by induction the sequence $(c_{[t]})_{t\ge1}$ is strictly decreasing. □

## A.5 Proof of Proposition 7

**Lemma A.1.** *Under under assumption 2, $|C(\mu_1^\bullet, \mu_2^{'}) - C(\mu_1^\bullet, \mu_2^{''})| \leq |\mu_2^{'} - \mu_2^{''}|$ for all $\mu_2^{'}, \mu_2^{''} \in \mathbb{R}$.*

*Proof.* Without loss of generality let $\mu_2^{'} > \mu_2^{''}$ with $w = \mu_2^{'} - \mu_2^{''} > 0$. The cutoff $C(\mu_1^\bullet, \mu_2^{'})$ satisfies the indifference condition,

$$u_1(C(\mu_1^\bullet, \mu_2^{'})) = \mathbb{E}_{\tilde{X}_2 \sim \mathcal{N}(\mu_2^{'} + \gamma(C(\mu_1^\bullet, \mu_2^{'}) - \mu_1^\bullet), \sigma^2)}[u_2(C(\mu_1^\bullet, \mu_2^{'}), \tilde{X}_2)]$$

I show that $C(\mu_1^\bullet, \mu_2^{''}) > C(\mu_1^\bullet, \mu_2^{'}) - w$. To do this, consider the difference in between payoff when stopping at $C(\mu_1^\bullet, \mu_2^{'}) - w$ and the expected payoff when continuing at $C(\mu_1^\bullet, \mu_2^{'}) - w$ for an agent who believes in fundamentals $(\mu_1^\bullet, \mu_2^{''})$:

$$\begin{aligned} &u_1(C(\mu_1^\bullet, \mu_2^{'}) - w) - \mathbb{E}_{\tilde{X}_2 \sim \mathcal{N}(\mu_2^{'} - w + \gamma(C(\mu_1^\bullet, \mu_2^{'}) - w - \mu_1^\bullet), \sigma^2)}[u_2(C(\mu_1^\bullet, \mu_2^{'}) - w, \tilde{X}_2)] \\ =&u_1(C(\mu_1^\bullet, \mu_2^{'}) - w) - \mathbb{E}_{\tilde{X}_2 \sim \mathcal{N}(\mu_2^{'} + \gamma(C(\mu_1^\bullet, \mu_2^{'}) - \mu_1^\bullet), \sigma^2)}[u_2(C(\mu_1^\bullet, \mu_2^{'}) - w, \tilde{X}_2 - (1+\gamma)w)] \end{aligned}$$

By Assumption 2, for every $x_2 \in \mathbb{R}$ we get

$$u_1(C(\mu_1^\bullet, \mu_2^{'}) - w) - [u_2(C(\mu_1^\bullet, \mu_2^{'}) - w, x_2 - (1+\gamma)w)] < u_1(C(\mu_1^\bullet, \mu_2^{'})) - [u_2(C(\mu_1^\bullet, \mu_2^{'}), x_2)]$$

which, combined with the indifference condition for $C(\mu_1^\bullet, \mu_2^{'})$, implies

$$u_1(C(\mu_1^\bullet, \mu_2^{'}) - w) - \mathbb{E}_{\tilde{X}_2 \sim \mathcal{N}(\mu_2^{'} - w + \gamma(C(\mu_1^\bullet, \mu_2^{'}) - w - \mu_1^\bullet), \sigma^2)}[u_2(C(\mu_1^\bullet, \mu_2^{'}) - w, \tilde{X}_2)] < 0.$$

That is, under belief $(\mu_1^\bullet, \mu_2^{''})$, the agent strictly prefers continuing at $C(\mu_1^\bullet, \mu_2^{'}) - w$. Since the optimal strategy at $(\mu_1^\bullet, \mu_2^{''})$ is given by a cutoff, above which the agent strictly prefers stopping, we then have $C(\mu_1^\bullet, \mu_2^{'}) - w < C(\mu_1^\bullet, \mu_2^{''})$. □

I now prove Proposition 7.

*Proof.* Consider the map $\Upsilon$ as discussed in the text,

$$\Upsilon(\mu_2) := \mu_2^\bullet + \gamma\left(\mu_1^\bullet - \mathbb{E}\left[X_1 | X_1 \leq C(\mu_1^\bullet, \mu_2)\right]\right).$$

I show $\Upsilon$ is a contraction mapping with modulus $|\gamma|$. We have

$$\Upsilon(\mu_2^{'}) - \Upsilon(\mu_2^{''}) = -\gamma \cdot \left(\mathbb{E}\left[X_1 | X_1 \leq C(\mu_1^\bullet, \mu_2^{'})\right] - \mathbb{E}\left[X_1 | X_1 \leq C(\mu_1^\bullet, \mu_2^{''})\right]\right).$$

By formula of the mean of a truncated Gaussian random variable, when $X_1 \sim \mathcal{N}(\mu_1^\bullet, \sigma^2)$ and

$c \in \mathbb{R}$, we get $\mathbb{E}[X_1|X_1 \le c] = \mu_1^\bullet - \left(\frac{\phi((c-\mu_1^\bullet)/\sigma)}{\Phi((c-\mu_1^\bullet)/\sigma)}\right)\sigma$. Therefore,

$$\begin{aligned}\Upsilon(\mu_2^{'}) - \Upsilon(\mu_2^{''}) &= -\gamma \cdot \left( \left( \mu_1^\bullet - \frac{\phi((C(\mu_1^\bullet, \mu_2^{'})/\sigma))}{\Phi((C(\mu_1^\bullet, \mu_2^{'})/\sigma))} \cdot \sigma \right) - \left( \mu_1^\bullet - \frac{\phi((C(\mu_1^\bullet, \mu_2^{''})/\sigma))}{\Phi((C(\mu_1^\bullet, \mu_2^{''})/\sigma))} \cdot \sigma \right) \right) \\ &= \gamma\sigma \cdot \left( \frac{\phi(C(\mu_1^\bullet, \mu_2^{'})/\sigma)}{\Phi(C(\mu_1^\bullet, \mu_2^{'})/\sigma)} - \frac{\phi(C(\mu_1^\bullet, \mu_2^{''})/\sigma)}{\Phi(C(\mu_1^\bullet, \mu_2^{''})/\sigma)} \right).\end{aligned}$$

The function $z \mapsto \frac{\phi(z)}{1-\Phi(z)}$ is the Gaussian inverse Mills ratio and it is well-known that its derivative is bounded by 1 in magnitude[16]. By symmetry this also applies to the function $z \mapsto \frac{\phi(z)}{\Phi(z)}$. This means $\left|\frac{\phi(z^{'})}{\Phi(z^{'})} - \frac{\phi(z^{''})}{\Phi(z^{''})}\right| \le |z^{'} - z^{''}|$. So we have

$$\left| \frac{\phi(C(\mu_1^\bullet, \mu_2^{'})/\sigma)}{\Phi(C(\mu_1^\bullet, \mu_2^{'})/\sigma)} - \frac{\phi(C(\mu_1^\bullet, \mu_2^{''})/\sigma)}{\Phi(C(\mu_1^\bullet, \mu_2^{''})/\sigma)} \right| \le \frac{1}{\sigma} \cdot \left| C(\mu_1^\bullet, \mu_2^{'}) - C(\mu_1^\bullet, \mu_2^{''}) \right| \le \frac{1}{\sigma} \cdot |\mu_2^{'} - \mu_2^{''}|$$

by Lemma A.1. This then showing $|\Upsilon(\mu_2^{'}) - \Upsilon(\mu_2^{''})| \le |\gamma| \cdot |\mu_2^{'} - \mu_2^{''}|$ for all $\mu_2^{'}, \mu_2^{''} \in \mathbb{R}$. So $\Upsilon$ is a contraction mapping with modulus $|\gamma| \in (0,1)$ and the proposition readily follows from properties of contraction mappings. □

## A.6 Proof of Proposition 9

*Proof.* Rewrite Equation (1) as

$$\begin{aligned}&\int_{-\infty}^{\infty} \phi(x_1; \mu_1^\bullet, (\sigma^\bullet)^2) \cdot \ln \left( \frac{\phi(x_1; \mu_1^\bullet, (\sigma^\bullet)^2)}{\phi(x_1; \mu_1, \sigma_1^2)} \right) dx_1 \\ &+ \int_{-\infty}^{c} \phi(x_1; \mu_1^\bullet, (\sigma^\bullet)^2) \cdot \int_{-\infty}^{\infty} \phi(x_2; \mu_2^\bullet, (\sigma^\bullet)^2) \ln \left[ \frac{\phi(x_2; \mu_2^\bullet, (\sigma^\bullet)^2)}{\phi(x_2; \mu_2 + \gamma(x_1 - \mu_1), \sigma_2^2)} \right] dx_2 dx_1.\end{aligned}$$

The KL divergence between $\mathcal{N}(\mu_{\text{true}}, \sigma_{\text{true}}^2)$ and $\mathcal{N}(\mu_{\text{model}}, \sigma_{\text{model}}^2)$ is $\ln \frac{\sigma_{\text{model}}}{\sigma_{\text{true}}} + \frac{\sigma_{\text{true}}^2 + (\mu_{\text{true}} - \mu_{\text{model}})^2}{2\sigma_{\text{model}}^2} - \frac{1}{2}$, so we may simplify the first term and the inner integral of the second term.

$$\begin{aligned}&\ln \frac{\sigma_1}{\sigma^\bullet} + \frac{(\mu_1 - \mu_1^\bullet)^2}{2\sigma_1^2} + \frac{(\sigma^\bullet)^2}{2\sigma_1^2} - \frac{1}{2} \\ &+ \int_{-\infty}^{c} \phi(x_1; \mu_1^\bullet, \sigma^\bullet) \cdot \left[ \ln \frac{\sigma_2}{\sigma^\bullet} + \frac{(\sigma^\bullet)^2 + (\mu_2 + \gamma(x_1 - \mu_1) - \mu_2^\bullet)^2}{2\sigma_2^2} - \frac{1}{2} \right] dx_1.\end{aligned}$$

[16]See for example Corollary 1.6 in Pinelis (2018)

Dropping terms not dependent on any of the four variables gives a simplified version of the objective,

$$\begin{aligned}\xi(\mu_1, \mu_2, \sigma_1, \sigma_2) := &\ln \frac{\sigma_1}{\sigma^\bullet} + \frac{(\mu_1 - \mu_1^\bullet)^2}{2\sigma_1^2} + \frac{(\sigma^\bullet)^2}{2\sigma_1^2} \\ &+ \int_{-\infty}^{c} \phi(x_1; \mu_1^\bullet, (\sigma^\bullet)^2) \cdot \left[\ln \frac{\sigma_2}{\sigma^\bullet} + \frac{(\sigma^\bullet)^2 + (\mu_2 + \gamma(x_1 - \mu_1) - \mu_2^\bullet)^2}{2\sigma_2^2}\right] dx_1.\end{aligned}$$

Differentiating under the integral sign,

$$\frac{\partial \xi}{\partial \mu_2} = \int_{-\infty}^{c} \phi(x_1; \mu_1^\bullet, (\sigma^\bullet)^2) \cdot \left[\frac{(\mu_2 + \gamma(x_1 - \mu_1) - \mu_2^\bullet)}{\sigma_2^2}\right] dx_1$$

$$\begin{aligned}\frac{\partial \xi}{\partial \mu_1} &= \frac{(\mu_1 - \mu_1^\bullet)}{\sigma_1^2} - \gamma \int_{-\infty}^{c} \phi(x_1; \mu_1^\bullet, (\sigma^\bullet)^2) \cdot \left[\frac{(\mu_2 + \gamma(x_1 - \mu_1) - \mu_2^\bullet)}{\sigma_2^2}\right] dx_1 \\ &= \frac{(\mu_1 - \mu_1^\bullet)}{\sigma_1^2} - \gamma \frac{\partial \xi}{\partial \mu_2}.\end{aligned}$$

At FOC $(\mu_1^*, \mu_2^*, \sigma_1^*, \sigma_2^*)$, we have $\frac{\partial \xi}{\partial \mu_2}(\mu_1^*, \mu_2^*, \sigma_1^*, \sigma_2^*) = 0$, hence $\mu_1^* = \mu_1^\bullet$. Similar arguments as before then establish $\mu_2^* = \mu_2^\bullet + \gamma\left(\mu_1^\bullet - \mathbb{E}\left[X_1 | X_1 \leq c\right]\right)$, where expectation is taken with respect to the true distribution of $X_1$ (with the true variance $(\sigma^\bullet)^2$). Then,

$$\frac{\partial \xi}{\partial \sigma_1}(\mu_1^*, \mu_2^*, \sigma_1^*, \sigma_2^*) = \frac{1}{(\sigma_1^*)} - \frac{(\sigma^\bullet)^2}{(\sigma_1^*)^3} = 0,$$

this gives $\sigma_1^* = \sigma^\bullet$ (since $\sigma_1^* \geq 0$).

Finally, from the FOC for $\sigma_2$,

$$\int_{-\infty}^{c} \phi(x_1; \mu_1^\bullet, (\sigma^\bullet)^2) \cdot \left[\frac{1}{\sigma_2^*} - \frac{(\sigma^\bullet)^2 + (\mu_2^* + \gamma(x_1 - \mu_1^*) - \mu_2^\bullet)^2}{(\sigma_2^*)^3}\right] dx_1 = 0.$$

Substituting in values of $\mu_1^*, \mu_2^*$ already solved for,

$$\begin{aligned}(\sigma_2^*)^2 &= (\sigma^\bullet)^2 + \mathbb{E}[(\mu_2^* + \gamma(X_1 - \mu_1^\bullet) - \mu_2^\bullet)^2 | X_1 \leq c] \\ &= (\sigma^\bullet)^2 + \mathbb{E}[(\mu_2^\bullet + \gamma\left(\mu_1^\bullet - \mathbb{E}\left[X_1 | X_1 \leq c\right]\right) + \gamma(X_1 - \mu_1^\bullet) - \mu_2^\bullet)^2 | X_1 \leq c] \\ &= (\sigma^\bullet)^2 + \gamma^2 \mathbb{E}\left[\left[(X_1 - \mu_1^\bullet) - (\mathbb{E}\left[X_1 | X_1 \leq c\right] - \mu_1^\bullet)\right]^2 | X_1 \leq c\right] \\ &= (\sigma^\bullet)^2 + \gamma^2 \mathrm{Var}[X_1 - \mu_1^\bullet | X_1 \leq c] \\ &= (\sigma^\bullet)^2 + \gamma^2 \mathrm{Var}[X_1 | X_1 \leq c]\end{aligned}$$

as desired. □

## A.7 Proof of Proposition 10

I start with a lemma that says, depending on the convexity of the decision problem, a stronger belief in fictitious variation either increases or decreases the subjectively optimal cutoff threshold.

**Lemma A.2.** *Suppose that under the subjective model $\Xi(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2; \gamma)$, the agent is indifferent between stopping at c and continuing. Suppose $\hat{\sigma}_2^2 > \sigma_2^2$. Then: (i) if $x_2 \mapsto u_2(c, x_2)$ is convex with strict convexity for $x_2$ in a positive-measure set, then under the subjective model $\Xi(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2; \gamma)$ the agent strictly prefers continuing at c; (ii) if $x_2 \mapsto u_2(c, x_2)$ is concave with strict concavity for $x_2$ in a positive-measure set, then under the subjective model $\Xi(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2; \gamma)$ the agent strictly prefers stopping at c.*

*Proof.* Indifference at $x_1 = c$ under the model $\Xi(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2; \gamma)$ implies that

$$u_1(c) = \mathbb{E}_{X_2 \sim \mathcal{N}(\mu_2 + \gamma(x_1 - \mu_1), \sigma_2^2)}[u_2(c, X_2)].$$

When hypothesis in (i) is satisfied,

$$\mathbb{E}_{X_2 \sim \mathcal{N}(\mu_2 + \gamma(x_1 - \mu_1), \sigma_2^2)}[u_2(c, X_2)] < \mathbb{E}_{X_2 \sim \mathcal{N}(\mu_2 + \gamma(x_1 - \mu_1), \hat{\sigma}_2^2)}[u_2(c, X_2)]$$

since $\hat{\sigma}_2^2 > \sigma_2^2$ implies that $\mathcal{N}(\mu_2 + \gamma(x_1 - \mu_1), \hat{\sigma}_2^2)$ is a strict mean-preserving spread of $\mathcal{N}(\mu_2 + \gamma(x_1 - \mu_1), \sigma_2^2)$. The RHS is the expected continuation payoff under model $\Xi(\mu_1, \mu_2, \sigma_1^2, \hat{\sigma}_2^2; \gamma)$, so the agent strictly prefers continuing when $X_1 = c$. The argument establishing (ii) is analogous. □

Now I give the proof of Proposition 10.

*Proof.* The result that $\mu_{1,[t]}^* = \mu_1^\bullet$, $(\sigma_{1,[t]}^*)^2 = (\sigma^\bullet)^2$ for all $t$ follows from Proposition 9.

Suppose $c_{[1]} \le c_{[0]}$. From Proposition 9, $\mu_{2,[2]}^* \le \mu_{2,[1]}^*$ and $(\sigma_{2,[2]}^*)^2 \le (\sigma_{2,[1]}^*)^2$. Let $c'_{[2]}$ be the indifference threshold under the model $\Xi(\mu_1^\bullet, \mu_{2,[2]}^*, (\sigma^\bullet)^2, (\sigma_{2,[1]}^*)^2)$. By Lemma 2, $c'_{[2]} \le c_{[1]}$. Also, from Lemma A.2, $c_{[2]} \le c'_{[2]}$ as generation 2 actually believes in the subjective model $\Xi(\mu_1^\bullet, \mu_{2,[2]}, (\sigma^\bullet)^2, (\sigma_{2,[2]}^*)^2)$ where $(\sigma_{2,[2]}^*)^2 \le (\sigma_{2,[1]}^*)^2$. This shows $c_{[2]} \le c_{[1]}$. Continuing this argument shows that $(c_{[t]})_{t \ge 1}$ forms a monotonically decreasing sequence. Since the pseudo-true parameters $\mu_2^*$ and $(\sigma_2^*)^2$ are monotonic functions of the censoring threshold $c$, we have established the proposition in the case where $c_{[1]} \le c_{[0]}$.

The argument for the case where $c_{[1]} \ge c_{[0]}$ is exactly analogous and therefore omitted. □

## A.8 Proof of Proposition 11

*Proof.* In the first generation, both societies A and B observe datasets censored according to the cutoff rule $c_{[0]} \uparrow$. So, by Proposition 9, two societies make the same inferences about the fundamentals.

Suppose the optimal-stopping problem is convex. Then due to fictitious variation in generation 1 and the convexity of $u_2$, it follows from Lemma A.2 that $c_{[B,1]} > c_{[A,1]}$. In the second generation, $\mu^*_{2,[B,2]} > \mu^*_{2,[A,2]}$ because the pseudo-true second-period fundamental increases in the censoring cutoff. Together again with the existence of fictitious variation, we conclude $c_{[B,2]} > c_{[A,2]}$. Continuing this argument establishes the proposition for the case where the optimal-stopping problem is convex. The case of concave optimal-stopping problems is analogous. □

## A.9 Proof of Lemma 5

*Proof.* By the same algebraic manipulations as in the proof of Proposition 1, we may rewrite the objective in Equation (2) as:

$$\frac{(\mu_1-\mu_1^\bullet)^2}{2\sigma^2}+\frac{1}{t}\sum_{\tau=0}^{t-1}\left\{\int_{-\infty}^{c_\tau}\phi(x_1;\mu_1^\bullet,\sigma^2)\cdot\left[\frac{\sigma^2+(\mu_2+\gamma(x_1-\mu_1)-\mu_2^\bullet)^2}{2\sigma^2}-\frac{1}{2}\right]dx_1\right\}.$$

Dropping terms not dependent on $\mu_1$, $\mu_2$ and multiplying through by $\sigma^2$, we get the simplified objective

$$\xi^{\mathrm{A}}(\mu_1,\mu_2):=\frac{(\mu_1-\mu_1^\bullet)^2}{2}+\frac{1}{t}\sum_{\tau=0}^{t-1}\left\{\int_{-\infty}^{c_k}\phi(x_1;\mu_1^\bullet,\sigma^2)\cdot\left[\frac{(\mu_2+\gamma(x_1-\mu_1)-\mu_2^\bullet)^2}{2\sigma^2}\right]dx_1\right\}$$

The same argument as in the proof of Proposition 1 gives $\mu_1=\mu_1^\bullet$ as the only value satisfying the first-order conditions, and following this the minimizing $\mu_2$ must satisfy $\frac{\partial\xi^{\mathrm{A}}}{\partial\mu_2}(\mu_1^\bullet,\mu_2)=0$. We now compute:

$$\frac{\partial\xi^{\mathrm{A}}}{\partial\mu_2}(\mu_1^\bullet,\mu_2)=\frac{1}{t}\sum_{\tau=0}^{t}\mathbb{P}[X_1\le c_\tau]\cdot\left(\mu_2-\mu_2^\bullet+\gamma\left(\mathbb{E}\left[X_1|X_1\le c_\tau\right]-\mu_1^\bullet\right)\right).$$

Since the derivative $\frac{\partial\xi^{\mathrm{A}}}{\partial\mu_2}$ is a linear function of $\mu_2$, when $\frac{\partial\xi^{\mathrm{A}}}{\partial\mu_2}(\mu_1^\bullet,\mu_2^*)=0$ we can rearrange to find

$$\begin{aligned}\mu_2^*&=\frac{1}{t\cdot\sum_{\tau=0}^{t-1}\mathbb{P}[X_1\le c_\tau]}\cdot\sum_{\tau=0}^{t-1}\mathbb{P}[X_1\le c_\tau]\left\{\mu_2^\bullet+\gamma\left(\mu_1^\bullet-\mathbb{E}\left[X_1|X_1\le c_\tau\right]\right)\right\}\\&=\frac{1}{t\cdot\sum_{\tau=0}^{t-1}\mathbb{P}[X_1\le c_\tau]}\sum_{\tau=0}^{t-1}\mathbb{P}[X_1\le c_\tau]\cdot\mu_2^*(c_\tau).\end{aligned}$$

This shows $\hat{\mu}_1^{\mathrm{A}}(c_0, ..., c_{t-1}) = \mu_1^\bullet$ and

$$\mu_2^{\mathrm{A}}(c_0, ..., c_{t-1}) = \frac{1}{t \cdot \sum_{\tau=0}^{t-1} \mathbb{P}[X_1 \le c_\tau]} \sum_{\tau=0}^{t-1} \mathbb{P}[X_1 \le c_\tau] \cdot \mu_2^*(c_\tau).$$

□

## A.10 Proof of Proposition 12

*Proof.* **Step 1**: If $c_{[1]}^{\mathrm{A}} > c_{[0]}^{\mathrm{A}}$, then $(\mu_2^{(t),\mathrm{A}})_{t \ge 1}$ and $(c_{[t]}^{\mathrm{A}})_{t \ge 0}$ are two increasing sequence, whereas $c_{[1]}^{\mathrm{A}} \le c_{[0]}^{\mathrm{A}}$ implies $(\mu_{2,[t]}^{\mathrm{A}})_{t \ge 1}$ and $(c_{[t]}^{\mathrm{A}})_{t \ge 0}$ are two decreasing sequences.

Suppose $c_{[1]}^{\mathrm{A}} > c_{[0]}^{\mathrm{A}}$. Note that by Lemma 5, $\mu_{2,[1]}^{\mathrm{A}} = \mu_2^*(c_{[0]}^{\mathrm{A}})$, whereas $\mu_{2,[2]}^{\mathrm{A}}$ is a weighted average between $\mu_2^*(c_{[0]}^{\mathrm{A}})$ and $\mu_2^*(c_{[1]}^{\mathrm{A}})$ where the latter is larger because $c_{[1]}^{\mathrm{A}} > c_{[0]}^{\mathrm{A}}$ and $\mu_2^*(c)$ is strictly increasing. This shows we have $\mu_{2,[2]}^{\mathrm{A}} > \mu_{2,[1]}^{\mathrm{A}}$ and hence $c_{[2]}^{\mathrm{A}} > c_{[1]}^{\mathrm{A}}$ as the cutoff is strictly increasing in its second argument by Lemma 2. Now assume the partial sequences $(c_{[\tau]}^{\mathrm{A}})_{\tau=0}^{T}$ and $(\mu_{2,[\tau]}^{\mathrm{A}})_{\tau=1}^{T}$ are both increasing. We show that $\mu_{2,[T+1]}^{\mathrm{A}} > \mu_{2,[T]}^{\mathrm{A}}$, which would also imply $c_{[T+1]}^{\mathrm{A}} > c_{[T]}^{\mathrm{A}}$. By comparing expressions for $\mu_{2,[T+1]}^{\mathrm{A}}$ and $\mu_{2,[t]}^{\mathrm{A}}$ given by Lemma 5,

$$\mu_{2,[T+1]}^{\mathrm{A}} = \delta \cdot \mu_2^*(c_{[T]}^{\mathrm{A}}) + (1-\delta) \cdot \mu_{2,[t]}^{\mathrm{A}}$$

where $\delta = \frac{\mathbb{P}[X_1 \le c_{[T]}^{\mathrm{A}}]}{(T+1) \cdot \sum_{\tau=0}^{T} \mathbb{P}[X_1 \le c^{(\tau),\mathrm{A}}]} > 0$ and $\mu_{2,[t]}^{\mathrm{A}}$ is itself a weighted average of the collection $\{\mu_2^*(c_{[\tau]}^{\mathrm{A}})\}_{\tau \le T-1}$ by Lemma 5. Now by the first part of the inductive hypothesis, $(c_{[\tau]}^{\mathrm{A}})_{\tau=0}^{T}$ is strictly increasing, meaning $\mu_2^*(c_{[T]}^{\mathrm{A}}) > \mu_2^*(c_{[\tau]}^{\mathrm{A}})$ for any $\tau < T$, which are the components making up $\mu_{2,[t]}^{\mathrm{A}}$. Since the weight $\delta$ on $\mu_2^*(c_{[T]}^{\mathrm{A}})$ in the expression of $\mu_{2,[T+1]}^{\mathrm{A}}$ is strictly positive, this shows $\mu_{2,[T+1]}^{\mathrm{A}} > \mu_{2,[t]}^{\mathrm{A}}$. So by induction, we have shown Step 1. (The other case of $c_{[1]}^{\mathrm{A}} < c_{[0]}^{\mathrm{A}}$ is symmetric.)

For the rest of this proof, suppose Assumption 2 holds and $-1 < \gamma < 0$.

**Step 2**: $(\mu_{2,[t]}^{\mathrm{A}})_{t \ge 1}$ is bounded and converges.

In the case that $c_{[1]}^{\mathrm{A}} \ge c_{[0]}^{\mathrm{A}}$ (so $\mu_{2,[2]}^{\mathrm{A}} \ge \mu_{2,[1]}^{\mathrm{A}}$), **Step 1** implies that $(\mu_{2,[t]}^{\mathrm{A}})_{t \ge 1}$ forms an increasing sequence. Since $\mu_2^*(\cdot)$ is bounded above by $\mu_2^\bullet$ by Proposition 1 and $\mu_{2,[t]}^{\mathrm{A}}$ for any $t \ge 1$ is a convex combinations of such terms, we also have $\mu_{2,[t]}^{\mathrm{A}} \le \mu_2^\bullet$ for every $t$. So in this case the sequence $(\mu_{2,[t]}^{\mathrm{A}})_{t \ge 1}$ is bounded between $\mu_{2,[1]}^{\mathrm{A}}$ and $\mu_2^\bullet$.

In the case that $c_{[1]}^{\mathrm{A}} \le c_{[0]}^{\mathrm{A}}$ (so $\mu_{2,[2]}^{\mathrm{A}} \le \mu_{2,[1]}^{\mathrm{A}}$), we notice that $c_{[0]}^{\mathrm{A}} = c^{(0)}$, $c_{[1]}^{\mathrm{A}} = c_{[1]}$, so by Corollary 1 the baseline model gives the learning dynamics $\mu_{2,[t]}^* \searrow \mu_2^\infty$, $c_{[t]} \searrow c^\infty$, where $(\mu_2^\infty, c^\infty)$ are associated with the unique steady state of the baseline model. So we have $\mu_{2,[1]}^{\mathrm{A}} = \mu_{2,[1]}^*$ while $\mu_{2,[2]}^{\mathrm{A}} \ge \mu_{2,[2]}^*$ since $\mu_{2,[2]}^{\mathrm{A}}$ is a convex combination between $\mu_2^*(c_{[0]})$ and

$\mu_2^*(c_{[1]}) = \mu_{2,[2]}^*$, with the latter being lower. This means $c_{[2]}^{\mathrm{A}} \geq c_{[2]}$. In the third generation,

$$\mu_2^{\mathrm{A}}(c_{[0]}^{\mathrm{A}}, c_{[1]}^{\mathrm{A}}, c_{[2]}^{\mathrm{A}}) \geq \mu_2^{\mathrm{A}}(c_{[0]}, c_{[1]}, c_{[2]}) \geq \mu_2^*(c_{[2]})$$

where the last inequality follows because $\mu_2^{\mathrm{A}}(c_{[0]}, c_{[1]}, c_{[2]})$ is a weighted average between $\mu_2^*(c_{[0]})$, $\mu_2^*(c_{[1]})$, and $\mu_2^*(c_{[2]})$, with the last one being the lowest since $c_{[t]}$ decreases in $t$. This shows $\mu_{2,[3]}^{\mathrm{A}} \geq \mu_{2,[3]}^*$ and $c_{[3]}^{\mathrm{A}} \geq c_{[3]}$. Iterating this argument shows that $\mu_{2,[t]}^{\mathrm{A}} \geq \mu_{2,[t]}^*$ for every $t$ in this case. Seeing as $(\mu_{2,[t]}^{\mathrm{A}})_{t\geq 1}$ forms a decreasing sequence by **Step 1**, it is bounded between $\mu_2^\infty$ and $\mu_{2,[1]}^{\mathrm{A}}$.

Since $(\mu_{2,[t]}^{\mathrm{A}})_{t\geq 1}$ is a bounded, monotonic sequence, it must converge. I denote this limit as $\mu_{2,[t]}^{\mathrm{A}} \to \mu_2^{\infty,\mathrm{A}}$. Also, Lemma A.1 shows that under Assumption 2, the indifference threshold $C$ is a continuous function of the second argument, so the sequence $c_{[T]}^{\mathrm{A}}$ must also converge. I denote this limit by $c_{[T]}^{\mathrm{A}} \to c^{\infty,\mathrm{A}}$.

**Step 3**: $\mu_2^{\infty,\mathrm{A}}$ is a fixed point of $\Upsilon$, so in particular $\mu_2^{\infty,\mathrm{A}} = \mu_2^\infty$ and $c^{\infty,\mathrm{A}} = c^\infty$ by uniqueness of the fixed point of $\Upsilon$.

The proof of Proposition 7 showed that under Assumption 2 and when $-1 < \gamma < 0$, $\Upsilon$ is a contraction mapping and hence must be continuous. Now let any $\epsilon > 0$ be given. I show there exists $\bar{t}$ so that $|\Upsilon(\mu_{2,[t]}^{\mathrm{A}}) - \mu_{2,[t]}^{\mathrm{A}}| < \epsilon$ for all $t > \bar{t}$. As this holds for all $\epsilon > 0$, continuity of $T$ then implies $\Upsilon(\mu_2^{\infty,\mathrm{A}}) - \mu_2^{\infty,\mathrm{A}} = 0$, that is $\mu_2^{\infty,\mathrm{A}}$ is a fixed point of $\Upsilon$.

We may write by Lemma 5,

$$\mu_{2,[T]}^{\mathrm{A}} = \frac{1}{t \cdot \sum_{\tau=0}^{t-1} \mathbb{P}[X_1 \leq c_{[\tau]}^{\mathrm{A}}]} \sum_{\tau=0}^{t-1} \mathbb{P}[X_1 \leq c_{[\tau]}^{\mathrm{A}}] \cdot \mu_2^*(c_{[\tau]}^{\mathrm{A}}).$$

The probabilities $\mathbb{P}[X_1 \leq c_{[\tau]}^{\mathrm{A}}]$ are bounded below since the beliefs $(\mu_{2,[t]}^{\mathrm{A}})_{t\geq 1}$ are bounded by **Step 2**. Also, since $\mu_2^*(\cdot)$ is continuous, there exists $\bar{t}_1$ so that $|\mu_2^*(c_{[T]}^{\mathrm{A}}) - \mu_2^*(c^{\infty,\mathrm{A}})| < \epsilon/2$ for all $t > \bar{t}_1$, that is to say $|\Upsilon(\mu_{2,[T]}^{\mathrm{A}}) - \mu_2^*(c^{\infty,\mathrm{A}})| < \epsilon/2$. When $t \to \infty$, the weight assigned to terms $\mu_2^*(c_{[\tau]}^{\mathrm{A}})$ with $\tau \geq \bar{t}_1$ in the expression for $\mu_{2,[T]}^{\mathrm{A}}$ grows to 1, which means $\limsup_{t\to\infty} |\mu_{2,[T]}^{\mathrm{A}} - \mu_2^*(c^{\infty,\mathrm{A}})| < \epsilon/2$. Combining these facts give $\limsup_{t\to\infty} |\Upsilon(\mu_{2,[t]}^{\mathrm{A}}) - \mu_{2,[t]}^{\mathrm{A}}| < \epsilon$ as desired. This establishes that $\mu_2^{\infty,\mathrm{A}}$ is a fixed point of $\Upsilon$, which combined with the uniqueness of $\Upsilon$'s fixed point implies it is equal to the unique steady-state belief about second-period fundamental in the baseline model. By continuity of $C$ under Assumption 2, $c^\infty = C(\mu_1^\bullet, \mu_2^{\infty,\mathrm{A}}) = C(\mu_1^\bullet, \mu_2^\infty) = c^\infty$. $\square$

### A.11 Proof of Proposition 13

*Proof.* In the true model, $X_2|(X_1 = x_1) \sim \mathcal{N}(\mu_2^\bullet + \gamma^\bullet(x_1 - \mu_1^\bullet), \sigma^2)$, while the agents' subjective model $\Xi(\mu_1, \mu_2; \gamma)$ has $X_2|(X_1 = x_1) \sim \mathcal{N}(\mu_2 + \gamma(x_1 - \mu_1), \sigma^2)$. So, we can write

$$D_{KL}(\mathcal{H}(\Xi(\mu_1^\bullet, \mu_2^\bullet; \gamma^\bullet); c\uparrow) \mid\mid \mathcal{H}(\Xi(\mu_1, \mu_2; \gamma); c\uparrow))$$

as the following:

$$\int_c^\infty \phi(x_1; \mu_1^\bullet, \sigma^2) \cdot \ln\left(\frac{\phi(x_1; \mu_1^\bullet, \sigma^2)}{\phi(x_1; \mu_1, \sigma^2)}\right) dx_1$$
$$+ \int_{-\infty}^c \left\{ \int_{-\infty}^\infty \begin{matrix} \phi(x_1; \mu_1^\bullet, \sigma^2) \cdot \phi(x_2; \mu_2^\bullet + \gamma^\bullet(x_1 - \mu_1^\bullet), \sigma^2) \cdot \\ \ln\left[\frac{\phi(x_1;\mu_1^\bullet,\sigma^2)\cdot\phi(x_2;\mu_2^\bullet+\gamma^\bullet(x_1-\mu_1^\bullet),\sigma^2)}{\phi(x_1;\mu_1,\sigma^2)\cdot\phi(x_2;\mu_2+\gamma(x_1-\mu_1),\sigma^2)}\right] \end{matrix} dx_2 \right\} dx_1.$$

Performing rearrangements similar to those in the proof of Proposition 1 and using the closed-form expression of KL divergence between two Gaussian distributions, the above can be rewritten as

$$\frac{(\mu_1 - \mu_1^\bullet)^2}{2\sigma^2} + \int_{-\infty}^c \phi(x_1; \mu_1^\bullet, \sigma^2) \cdot \frac{(\mu_2 + \gamma(x_1 - \mu_1) - \mu_2^\bullet - \gamma^\bullet(x_1 - \mu_1^\bullet))^2}{2\sigma^2} dx_1.$$

Multiplying through by$\sigma^2$ and dropping terms not depending on $\mu_1, \mu_2, \gamma$, we get a simplified objective with the same minimizers:

$$\xi(\mu_1, \mu_2, \gamma) = \frac{(\mu_1 - \mu_1^\bullet)^2}{2} + \int_{-\infty}^c \phi(x_1; \mu_1^\bullet, \sigma^2) \cdot \frac{1}{2} \cdot [\mu_2 + \gamma(x_1 - \mu_1) - \mu_2^\bullet - \gamma^\bullet(x_1 - \mu_1^\bullet)]^2 dx_1.$$

We have the partial derivatives by differentiating under the integral sign,

$$\frac{\partial \xi}{\partial \mu_2} = \int_{-\infty}^c \phi(x_1; \mu_1^\bullet, \sigma^2) \cdot [\mu_2 + \gamma(x_1 - \mu_1) - \mu_2^\bullet - \gamma^\bullet(x_1 - \mu_1^\bullet)] dx_1,$$

$$\begin{aligned} \frac{\partial \xi}{\partial \mu_1} &= (\mu_1 - \mu_1^\bullet) - \gamma \int_{-\infty}^c \phi(x_1; \mu_1^\bullet, \sigma^2) \cdot [\mu_2 + \gamma(x_1 - \mu_1) - \mu_2^\bullet - \gamma^\bullet(x_1 - \mu_1^\bullet)] dx_1 \\ &= (\mu_1 - \mu_1^\bullet) - \gamma \frac{\partial \xi}{\partial \mu_2}, \end{aligned}$$

$$\frac{\partial \xi}{\partial \gamma} = \int_{-\infty}^c \phi(x_1; \mu_1^\bullet, \sigma^2) \cdot [x_1 - \mu_1] \cdot [\mu_2 + \gamma(x_1 - \mu_1) - \mu_2^\bullet - \gamma^\bullet(x_1 - \mu_1^\bullet)] dx_1.$$

Suppose $(\mu_1^*, \mu_2^*, \gamma^*)$ is the minimum. By the first-order conditions for $\mu_1$ and $\mu_2$, we have:

$$\frac{\partial \xi}{\partial \mu_1}(\mu_1^*, \mu_2^*, \gamma^*) = \frac{\partial \xi}{\partial \mu_2}(\mu_1^*, \mu_2^*, \gamma^*) = 0 \Rightarrow \mu_1^* = \mu_1^\bullet.$$

Substituting this into the first-order condition for $\mu_2$,

$$\frac{\partial \xi}{\partial \mu_2}(\mu_1^\bullet, \mu_2^*, \gamma^*) = 0 \Rightarrow \mu_2^* = \mu_2^\bullet + (\gamma^* - \gamma^\bullet) \cdot (\mu_1^\bullet - \mathbb{E}[X_1 | X_1 \leq c]) .$$

It remains to show $\gamma^* = \tilde{\gamma}$. We have

$$\frac{\partial \xi}{\partial \gamma}(\mu_1^*, \mu_2^*, \gamma^*) = \mathbb{P}[X_1 \leq c] \cdot \mathbb{E}[(X_1 - \mu_1^*) \cdot (\mu_2^* + \gamma^*(X_1 - \mu_1^*) - \mu_2^\bullet - \gamma^\bullet(X_1 - \mu_1^\bullet)) | X_1 \leq c].$$

We rearrange the expectation term as:

$$\begin{aligned}
&\mathbb{E}[(X_1 - \mu_1^*) \cdot (\mu_2^* + \gamma^*(X_1 - \mu_1^*) - \mu_2^\bullet - \gamma^\bullet(X_1 - \mu_1^\bullet)) | X_1 \leq c] \\
=&\mathbb{E}[(X_1 - \mu_1^*) | X_1 \leq c] \cdot \mathbb{E}[(\mu_2^* + \gamma^*(X_1 - \mu_1^*) - \mu_2^\bullet - \gamma^\bullet(X_1 - \mu_1^\bullet)) | X_1 \leq c] \\
&+ \mathrm{Cov}(X_1 - \mu_1^*, \mu_2^* + \gamma^*(X_1 - \mu_1^*) - \mu_2^\bullet - \gamma^\bullet(X_1 - \mu_1^\bullet) | X_1 \leq c].
\end{aligned}$$

The first-order condition for $\mu_2$ implies $\mathbb{E}[(\mu_2^* + \gamma^*(X_1 - \mu_1^*) - \mu_2^\bullet - \gamma^\bullet(X_1 - \mu_1^\bullet)) | X_1 \leq c] = 0$ at the optimum $(\mu_1^*, \mu_2^*, \gamma^*)$. Also, we may drop terms without $X_1$ in the conditional covariance operator, and we get:

$$\frac{\partial \xi}{\partial \gamma}(\mu_1^*, \mu_2^*, \gamma^*) = \mathbb{P}[X_1 \leq c] \cdot (\gamma^* - \gamma^\bullet) \cdot \mathrm{Cov}(X_1, X_1 | X_1 \leq c).$$

We have $\mathbb{P}[X_1 \leq c] > 0$ and $\mathrm{Cov}(X_1, X_1 | X_1 \leq c) > 0$, hence we conclude

$$\frac{\partial \xi}{\partial \gamma}(\mu_1^*, \mu_2^*, \gamma^*) \begin{cases} > 0 & \text{for } \gamma^* > \gamma^\bullet \\ = 0 & \text{for } \gamma^* = \gamma^\bullet \\ < 0 & \text{for } \gamma^* < \gamma^\bullet \end{cases} .$$

In case that $\bar{\gamma} < \gamma^\bullet$, at the optimum we must have $\frac{\partial \xi}{\partial \gamma}(\mu_1^*, \mu_2^*, \gamma^*) < 0$. By Karush-Kuhn-Tucker condition, this means the optimum is $\gamma^* = \bar{\gamma}$. Conversely, when $\underline{\gamma} > \gamma^\bullet$, at the optimum we must have $\frac{\partial \xi}{\partial \gamma}(\mu_1^*, \mu_2^*, \gamma^*) > 0$. In that case, the optimum is $\gamma^* = \underline{\gamma}$. So in both cases, $\gamma^* = \tilde{\gamma}$ as desired. □

## A.12 Proof of Proposition 14

*Proof.* Let $w_1 = \alpha, w_2 = 1 - \alpha$, $c_1 = C(\mu_1^\bullet, \mu_2^\bullet)$, $c_2 = c$. By the same argument as in the proof of Proposition 5, we may rewrite the weighted KL divergence as

$$\frac{(\mu_1 - \mu_1^\bullet)^2}{2\sigma^2} + \sum_{k=1}^{2} w_k \left\{ \int_{-\infty}^{c_k} \phi(x_1; \mu_1^\bullet, \sigma^2) \cdot \left[ \frac{\sigma^2 + (\mu_2 + \gamma(x_1 - \mu_1) - \mu_2^\bullet)^2}{2\sigma^2} - \frac{1}{2} \right] dx_1 \right\} .$$

Dropping terms not dependent on $\mu_1$, $\mu_2$ and multiplying through by $\sigma^2$, we get the simplified objective

$$\xi^{SN}(\mu_1, \mu_2) := \frac{(\mu_1 - \mu_1^\bullet)^2}{2} + \sum_{k=1}^{2} w_k \left\{ \int_{-\infty}^{c_k} \phi(x_1; \mu_1^\bullet, \sigma^2) \cdot \left[ \frac{(\mu_2 + \gamma(x_1 - \mu_1) - \mu_2^\bullet)^2}{2\sigma^2} \right] dx_1 \right\}.$$

The same argument as in the proof of Proposition 5 shows that the first-order condition is only satisfied at $\mu_1^{SN} = \mu_1^\bullet$,

$$\mu_2^{SN} = \frac{1}{w_1 \mathbb{P}[X_1 \leq c_1] + w_2 \mathbb{P}[X_1 \leq c_2]} \sum_{k=1}^{2} w_k \mathbb{P}[X_1 \leq c_k] \left\{ \mu_2^\bullet + \gamma \left( \mu_1^\bullet - \mathbb{E}\left[ X_1 | X_1 \leq c_k \right] \right) \right\}.$$

This shows, in terms of expressions for pseudo-true fundamentals in the baseline model $\mu_2^*$,

$$\begin{aligned} \mu_2^{SN}(c) = & \frac{\alpha \mathbb{P}[X_1 \leq C(\mu_1^\bullet, \mu_2^\bullet)]}{\alpha \mathbb{P}[X_1 \leq C(\mu_1^\bullet, \mu_2^\bullet)] + (1-\alpha)\mathbb{P}[X_1 \leq c]} \cdot \mu_2^*(C(\mu_1^\bullet, \mu_2^\bullet)) \\ & + \frac{(1-\alpha)\mathbb{P}[X_1 \leq c]}{\alpha \mathbb{P}[X_1 \leq C(\mu_1^\bullet, \mu_2^\bullet)] + (1-\alpha)\mathbb{P}[X_1 \leq c]} \cdot \mu_2^*(c). \end{aligned}$$

□

## A.13 Proof of Corollary 3

*Proof.* From Proposition 14 (and Proposition 1 for the case of $t = 1$), $\mu_{1,[t]}^1 = \mu_{1,[t]}^1 = \mu_1^\bullet$ for every $t \geq 1$. Also, in the first generation, $\mu_{2,[1]}^1 = \mu_{2,[1]}^2$ and $c_{[1]}^1 = c_{[1]}^2$ since both societies face the same dataset $\mathcal{H}^\bullet(c_{[0]})$. Since $\mu_{2,[1]}^1 < \mu_2^\bullet$, we must have $c_{[1]}^1 = C(\mu_1^\bullet, \mu_{2,[1]}^1) < C(\mu_1^\bullet, \mu_2^\bullet)$ by Lemma 2. In the second generation, $\mu_{2,[2]}^1 = \mu_2^*(c_{[1]}^1)$ and $\mu_{2,[2]}^2$ is a convex combination between $\mu_2^*(c_{[1]}^2)$ and $\mu_2^*(C(\mu_1^\bullet, \mu_2^\bullet))$. As $\mu_2^*(c_{[1]}^1) = \mu_2^*(c_{[1]}^2) < \mu_2^*(C(\mu_1^\bullet, \mu_2^\bullet))$ due to Proposition 1, we conclude $\mu_{2,[2]}^2 > \mu_{2,[2]}^1$ and hence $c_{[2]}^2 > c_{[2]}^1$. But when $c_{[t]}^2 > c_{[t]}^1$ and $C(\mu_1^\bullet, \mu_2^\bullet) > c_{[t]}^1$ we have $\mu_2^*(c_{[t]}^1) < \mu_2^*(c_{[t]}^2)$, which shows in the next generation, $\mu_{2,[t+1]}^2$ is the convex combination of two terms both exceeding $\mu_2^*(c_{[t]}^1)$. This implies $\mu_{2,[t+1]}^2 > \mu_{2,[t+1]}^1$ and $c_{[t+1]}^2 > c_{[t+1]}^1$. By induction, the corollary holds for all $t \geq 2$. □

## A.14 Proof of Lemma 6

*Proof.* I check conditions A1 through A5 in Bunke and Milhaud (1998). The lemma follows from their Theorem 2 when these conditions are satisfied.

The parameter space is $\Theta = \mathbb{R}^2$. The data-generating density of observation $(x, y)$ is:

$$f^\bullet(x,y) = \begin{cases} \phi(x;\mu_1^\bullet,\sigma^2)\cdot\phi(y;\mu_2^\bullet,\sigma^2) & \text{if } x < c \\ \phi(x;\mu_1^\bullet,\sigma^2)\cdot\phi(y;0,1) & \text{if } x \geq c \end{cases}$$

where $\phi(\cdot;\mu,\sigma^2)$ is the Gaussian density with mean $\mu$ and variance $\sigma^2$. Under parameters $(\hat{\mu}_1, \hat{\mu}_2)$ (and with dogmatic belief in $\gamma < 0$), the same observation has density:

$$f_{\hat{\mu}_1,\hat{\mu}_2}(x,y) = \begin{cases} \phi(x;\hat{\mu}_1,\sigma^2)\cdot\phi(y;\hat{\mu}_2+\gamma\cdot(x-\hat{\mu}_1),\sigma^2) & \text{if } x < c \\ \phi(x;\hat{\mu}_1,\sigma^2)\cdot\phi(y;0,1) & \text{if } x \geq c. \end{cases}$$

**A1**. Parameter space is a closed, convex set in $\mathbb{R}^2$ with nonempty interior. The density $f_{\hat{\mu}_1,\hat{\mu}_2}(x,y)$ is bounded over $(\hat{\mu}_1,\hat{\mu}_2,x,y)$ and its carrier, $\{(x,y) : f_{\hat{\mu}_1,\hat{\mu}_2}(x,y) > 0\}$ is the same for all $\hat{\mu}_1,\hat{\mu}_2$.

Evidently $\mathbb{R}^2$ is closed in itself. The density $f_{\hat{\mu}_1,\hat{\mu}_2}(x,y)$ is bounded by the product of the modes of Gaussian densities with variance $\sigma^2$ and variance 1. The density $f_{\hat{\mu}_1,\hat{\mu}_2}(x,y)$ is strictly positive on $\mathbb{R}^2$ for any parameter values $\hat{\mu}_1,\hat{\mu}_2$.

**A2**. For all $\hat{\mu}_1,\hat{\mu}_2$, there is a sphere $S[(\hat{\mu}_1,\hat{\mu}_2),\eta]$ of center $(\hat{\mu}_1,\hat{\mu}_2)$ and radius $\eta > 0$ such that:

$$\mathbb{E}_{f^\bullet}\left[\sup_{(\mu_1',\mu_2')\in S[(\hat{\mu}_1,\hat{\mu}_2),\eta]}\left|\ln\frac{f^\bullet(X,Y)}{f_{\mu_1',\mu_2'}(X,Y)}\right|\right] < \infty.$$

Pick say $\eta = 1$. Consider the rectangle $R[(\hat{\mu}_1,\hat{\mu}_2),\eta]$ consisting of points $(\mu_1',\mu_2')$ such that $|\mu_1'-\hat{\mu}_1| < \eta$ and $|\mu_2'-\hat{\mu}_2| < \eta$. Since the the Gaussian distribution is single-peaked, for any $(x,y)\in\mathbb{R}^2$ the absolute value of the log likelihood ratio $\left|\ln\frac{f^\bullet(X,Y)}{f_{\mu_1',\mu_2'}(X,Y)}\right|$ on all of $R[(\hat{\mu}_1,\hat{\mu}_2),\eta]$ must be bounded by its value at the 4 corners. That is to say,

$$\begin{aligned}
&\sup_{(\mu_1',\mu_2')\in S[(\hat{\mu}_1,\hat{\mu}_2),\eta]}\left|\ln\frac{f^\bullet(X,Y)}{f_{\mu_1',\mu_2'}(X,Y)}\right| \\
\leq &\sup_{(\mu_1',\mu_2')\in R[(\hat{\mu}_1,\hat{\mu}_2),\eta]}\left|\ln\frac{f^\bullet(X,Y)}{f_{\mu_1',\mu_2'}(X,Y)}\right| \\
\leq &\left|\ln\frac{f^\bullet(X,Y)}{f_{\hat{\mu}_1-\eta,\hat{\mu}_2-\eta}(X,Y)}\right| + \left|\ln\frac{f^\bullet(X,Y)}{f_{\hat{\mu}_1-\eta,\hat{\mu}_2+\eta}(X,Y)}\right| + \left|\ln\frac{f^\bullet(X,Y)}{f_{\hat{\mu}_1+\eta,\hat{\mu}_2-\eta}(X,Y)}\right| + \left|\ln\frac{f^\bullet(X,Y)}{f_{\hat{\mu}_1+\eta,\hat{\mu}_2+\eta}(X,Y)}\right|.
\end{aligned}$$

It is easy to see that for any fixed parameter $\mathbb{E}_{f^\bullet}\left[\left|\ln\frac{f^\bullet(X,Y)}{f_{\mu_1',\mu_2'}(X,Y)}\right|\right]$ is finite, so the sum of these 4 terms gives a finite upper bound.

**A3**. For all fixed $(x_0,y_0)\in\mathbb{R}^2$, the map from parameters to density $(\mu_1,\mu_2)\mapsto f_{\mu_1,\mu_2}(x_0,y_0)$

has continuous derivatives with respect to parameters $(\mu_1, \mu_2) \mapsto \frac{\partial f}{\partial \mu_1}(x_0, y_0; \mu_1, \mu_2)$, $(\mu_1, \mu_2) \mapsto \frac{\partial f}{\partial \mu_2}(x, y; \mu_1, \mu_2)$. There exist positive constants $\kappa_0$ and $b_0$ with

$$\int\int \left\| [f_{\mu_1,\mu_2}(x,y)]^{-1} \cdot \begin{pmatrix} \frac{\partial f}{\partial \mu_1}(x, y; \mu_1, \mu_2) \\ \frac{\partial f}{\partial \mu_2}(x, y; \mu_1, \mu_2) \end{pmatrix} \right\|^{12} \cdot f_{\mu_1,\mu_2}(x,y) \cdot dydx < \kappa_0(1 + ||(\mu_1, \mu_2)||^{b_0})$$

satisfied for every $(\mu_1, \mu_2) \in \mathbb{R}^2$, where $|| \cdot ||$ is a norm on $\mathbb{R}^2$.

Let's choose the max norm, $||v|| = \max(|v_1|, |v_2|)$. For uncensored data $(x_0, y_0)$ with $x_0 < c$, we can compute

$$\frac{\partial f}{\partial \mu_1}(x_0, y_0; \mu_1, \mu_2) = f_{\mu_1,\mu_2}(x_0, y_0) \cdot \left[ \frac{(1+\gamma^2)}{\sigma^2} \cdot (x - \mu_1) - \frac{\gamma}{\sigma^2} \cdot (y - \mu_2) \right]$$

and

$$\frac{\partial f}{\partial \mu_2}(x_0, y_0; \mu_1, \mu_2) = f_{\mu_1,\mu_2}(x_0, y_0) \cdot \left[ -\frac{\gamma}{\sigma^2} \cdot (x - \mu_1) - \frac{1}{\sigma^2} \cdot (y - \mu_2) \right].$$

While for censored data $(x_0, y_0)$ where $x_0 > c$, the likelihood of the data is unchanged by parameter $\mu_2$ since it neither changes the distribution of the early draw quality nor the distribution of the white noise term, meaning $\frac{\partial f}{\partial \mu_2}(x_0, y_0; \mu_1, \mu_2) = 0$. Also, for the censored case

$$\frac{\partial f}{\partial \mu_1}(x_0, y_0; \mu_1, \mu_2) = f_{\mu_1,\mu_2}(x_0, y_0) \cdot \frac{1}{\sigma^2}(x - \mu_1).$$

This means the integral to be bounded is:

$$\int_{x=-\infty}^{x=c} \left[ \int_{-\infty}^{\infty} \left\| \begin{pmatrix} \frac{(1+\gamma^2)}{\sigma^2} \cdot (x - \mu_1) - \frac{\gamma}{\sigma^2} \cdot (y - \mu_2) \\ -\frac{\gamma}{\sigma^2} \cdot (x - \mu_1) - \frac{1}{\sigma^2} \cdot (y - \mu_2) \end{pmatrix} \right\|^{12} \cdot f_{\mu_1,\mu_2}(x,y) \cdot dy \right] dx$$
$$+ \int_{x=c}^{x=\infty} \left[ \int_{-\infty}^{\infty} (\frac{1}{\sigma^2}(x - \mu_1))^{12} \cdot f_{\mu_1,\mu_2}(x,y) \cdot dy \right] dx.$$

Since the inner integrals are non-negative, this expression is smaller than the version where the domains of the outer integrals are expanded and the densities $f_{\mu_1,\mu_2}(x,y)$ are simply replaced with the joint density on $\mathbb{R}^2$ of the subjective model for $\Xi(\mu_1, \mu_2)$, which I denote as $\tilde{f}_{\mu_1,\mu_2}(x,y)$.

$$\int_{-\infty}^{\infty} \left[ \int_{-\infty}^{\infty} \left\| \begin{pmatrix} \frac{(1+\gamma^2)}{\sigma^2} \cdot (x - \mu_1) - \frac{\gamma}{\sigma^2} \cdot (y - \mu_2) \\ -\frac{\gamma}{\sigma^2} \cdot (x - \mu_1) - \frac{1}{\sigma^2} \cdot (y - \mu_2) \end{pmatrix} \right\|^{12} \cdot \tilde{f}_{\mu_1,\mu_2}(x,y) \cdot dy \right] dx$$
$$+ \int_{-\infty}^{\infty} \left[ \int_{-\infty}^{\infty} (\frac{1}{\sigma^2}(x - \mu_1))^{12} \cdot \tilde{f}_{\mu_1,\mu_2}(x,y) \cdot dy \right] dx.$$

The second summand is a 12th moment of the joint normal random variable with distribution

$\Xi(\mu_1, \mu_2)$, so for all $\mu_1, \mu_2$ it is given by some 12th order polynomial $P_2(\mu_1, \mu_2)$. Similarly the first summand is also given by a 12th order polynomial $P_1(\mu_1, \mu_2)$. Therefore by choosing $b_0 = 12$ and choosing $\kappa_0$ appropriately according to the coefficients in $P_1$ and $P_2$, we achieved the desired bound.

**A4**. For some positive constants $b_1$ and $\kappa_1$, the affinity function

$$A(\mu_1, \mu_2) := \int\int [f_{\mu_1,\mu_2}(x, y) \cdot f^\bullet(x, y)]^{1/2} dydx$$

satisfies $A(\mu_1, \mu_2) < \kappa_1 \cdot ||(\mu_1, \mu_2)||^{-b_1}$ for all $\mu_1, \mu_2$.

We have $A(\mu_1, \mu_2) \leq [\int\int [f_{\mu_1,\mu_2}(x, y) \cdot f^\bullet(x, y)]dydx]^{1/2}$, so it's sufficient to find some $\kappa_1$ and $b_1$ that works to bound $\int\int [f_{\mu_1,\mu_2}(x, y) \cdot f^\bullet(x, y)]dydx$. We have:

$$\begin{aligned}
&\int\int [f_{\mu_1,\mu_2}(x, y) \cdot f^\bullet(x, y)]dydx \\
&= \int_{x=-\infty}^{c}\int_{-\infty}^{\infty} \phi(x; \mu_1, \sigma^2) \cdot \phi(x; \mu_1^\bullet, \sigma^2) \cdot \phi(y; \mu_2 + \gamma(x - \mu_1), \sigma^2) \cdot \phi(y; \mu_2^\bullet, \sigma^2)dydx \\
&\ + \int_{x=c}^{\infty}\int_{-\infty}^{\infty} \phi(x; \mu_1, \sigma^2) \cdot \phi(x; \mu_1^\bullet, \sigma^2) \cdot \phi(y; 0, 1) \cdot \phi(y; 0, 1)dydx \\
&\leq \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \phi(x; \mu_1, \sigma^2) \cdot \phi(x; \mu_1^\bullet, \sigma^2) \cdot \phi(y; \mu_2 + \gamma(x - \mu_1), \sigma^2) \cdot \phi(y; \mu_2^\bullet, \sigma^2)dydx \\
&\ + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \phi(x; \mu_1, \sigma^2) \cdot \phi(x; \mu_1^\bullet, \sigma^2) \cdot \phi(y; 0, 1) \cdot \phi(y; 0, 1)dydx.
\end{aligned}$$

I show how to find $\kappa_1$ and $b_1$ to bound the first summand in the last expression above. It is easy to similarly bound the second summand. By Bromiley (2003), the product of Gaussian densities $\phi(y; \mu_2+\gamma(x-\mu_1), \sigma^2)\cdot\phi(y; \mu_2^\bullet, \sigma^2)$ is itself a Gaussian density in $y$, $\tilde{\phi}(y)$, multiplied by a scaling factor equal to $(4\pi\sigma^2)^{-1/2} \cdot \exp\left(-\frac{\gamma^2}{4\sigma^2} \cdot [x - (\mu_1 + \frac{\mu_2^\bullet}{\gamma} - \frac{\mu_2}{\gamma})]^2\right)$. So we have

$$\begin{aligned}
&\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \phi(x; \mu_1, \sigma^2) \cdot \phi(x; \mu_1^\bullet, \sigma^2) \cdot \phi(y; \mu_2 + \gamma(x - \mu_1), \sigma^2) \cdot \phi(y; \mu_2^\bullet, \sigma^2)dydx \\
&= \int_{-\infty}^{\infty} \phi(x; \mu_1, \sigma^2) \cdot \phi(x; \mu_1^\bullet, \sigma^2) \cdot \left(4\pi\sigma^2\right)^{-1/2} \cdot \exp\left(-\frac{\gamma^2}{4\sigma^2} \cdot [x - (\mu_1 + \frac{\mu_2^\bullet}{\gamma} - \frac{\mu_2}{\gamma})]^2\right) \cdot \int_{-\infty}^{\infty} \cdot\tilde{\phi}(y)dydx \\
&= \left(4\pi\sigma^2\right)^{-1/2} \cdot \int_{-\infty}^{\infty} \phi(x; \mu_1, \sigma^2) \cdot \phi(x; \mu_1^\bullet, \sigma^2) \cdot \exp\left(-\frac{\gamma^2}{4\sigma^2} \cdot [x - (\mu_1 + \frac{\mu_2^\bullet}{\gamma} - \frac{\mu_2}{\gamma})]^2\right) \cdot dx.
\end{aligned}$$

Again applying Bromiley (2003), product of the two Gaussian densities $\phi(x; \mu_1, \sigma^2)\cdot\phi(x; \mu_1^\bullet, \sigma^2)$ is another Gaussian density with mean $\frac{\mu_1^\bullet+\mu_1}{2}$, variance $\frac{\sigma^2}{2}$, and multiplied to a scaling factor of $(4\pi\sigma^2)^{-1/2} \exp\left(-\frac{(\mu_1-\mu_1^\bullet)^2}{4\sigma^2}\right)$. So above expression is:

$$K_1 \cdot \exp\left(-\frac{(\mu_1 - \mu_1^\bullet)^2}{4\sigma^2}\right) \cdot \int_{-\infty}^{\infty} \phi(x; \frac{\mu_1^\bullet + \mu_1}{2}, \frac{\sigma^2}{2}) \cdot \exp\left(-\frac{\gamma^2}{4\sigma^2} \cdot [x - (\mu_1 + \frac{\mu_2^\bullet}{\gamma} - \frac{\mu_2}{\gamma})]^2\right) \cdot dx$$

where $K_1$ is a constant not dependent on $\mu_1, \mu_2$. Also, we may write

$$\exp\left(-\frac{\gamma^2}{4\sigma^2} \cdot [x - (\mu_1 + \frac{\mu_2^\bullet}{\gamma} - \frac{\mu_2}{\gamma})]^2\right) = K_2 \cdot \phi(x; (\mu_1 + \frac{\mu_2^\bullet}{\gamma} - \frac{\mu_2}{\gamma}), \sigma_B^2)$$

where $\sigma_B^2 = \frac{2\sigma^2}{\gamma^2}$ and $K_2 = (2\pi\sigma_B^2)^{1/2}$. Applying Bromiley (2003) one final time, the product $\phi(x; \frac{\mu_1^\bullet + \mu_1}{2}, \frac{\sigma^2}{2}) \cdot \phi(x; (\mu_1 + \frac{\mu_2^\bullet}{\gamma} - \frac{\mu_2}{\gamma}), \sigma_B^2)$ is a Gaussian density in $x$ scaled by $K_4 \exp(-K_3 \cdot (\frac{\mu_1^\bullet - \mu_1}{2} + \frac{\mu_2 - \mu_2^\bullet}{\gamma})^2)$ where $K_3, K_4 > 0$ are constants not dependent on $\mu_1, \mu_2$. So altogether, the second summand we are bounding is a constant multiple of $\exp\left(-\frac{(\mu_1 - \mu_1^\bullet)^2}{4\sigma^2}\right) \cdot \exp(-K_3 \cdot (\frac{\mu_1^\bullet - \mu_1}{2} + \frac{\mu_2 - \mu_2^\bullet}{\gamma})^2)$. For $|\mu_1| \geq |\mu_2|$, the max norm $||(\mu_1, \mu_2)|| = |\mu_1|$ and $\exp\left(-\frac{(\mu_1 - \mu_1^\bullet)^2}{4\sigma^2}\right)$ decreases exponentially fast in the norm. For $|\mu_1| < |\mu_2|$, and $\frac{|\mu_2|}{2} - \frac{|\mu_1^\bullet|}{2} - \frac{|\mu_2^\bullet|}{\gamma} > 0$,

$$\exp(-K_3 \cdot (\frac{\mu_1^\bullet - \mu_1}{2} + \frac{\mu_2 - \mu_2^\bullet}{\gamma})^2) \leq \exp(-K_3 \cdot (\frac{|\mu_2|}{2} - \frac{|\mu_1^\bullet|}{2} - \frac{|\mu_2^\bullet|}{\gamma})^2).$$

So for large enough $|\mu_2|$, $\exp(-K_3 \cdot (\frac{\mu_1^\bullet - \mu_1}{2} + \frac{\mu_2 - \mu_2^\bullet}{\gamma})^2)$ will decrease exponentially fast in the norm. These two facts imply that there is some $K > 0$ so that whenever $||(\mu_1, \mu_2)|| > K$,

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi(x; \mu_1, \sigma^2) \cdot \phi(x; \mu_1^\bullet, \sigma^2) \cdot \phi(y; \mu_2 + \gamma(x - \mu_1), \sigma^2) \cdot \phi(y; \mu_2^\bullet, \sigma^2) dy dx < ||(\mu_1, \mu_2)||^{-1}.$$

Now put $\kappa_1 = K^{-1}$ and we can ensure for any value of $||(\mu_1, \mu_2)||$ we will have

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi(x; \mu_1, \sigma^2) \cdot \phi(x; \mu_1^\bullet, \sigma^2) \cdot \phi(y; \mu_2 + \gamma(x - \mu_1), \sigma^2) \cdot \phi(y; \mu_2^\bullet, \sigma^2) dy dx < \kappa_1 \cdot ||(\mu_1, \mu_2)||^{-1}.$$

**A5**. There are positive constants $b_2, b_3$ so that for all $(\mu_1', \mu_2')$ and $r > 0$ it holds that $g(S[(\mu_1', \mu_2'), r]) \leq c r^{b_2}(1 + (||(\mu_1', \mu_2')|| + r)^{b_3})$. Moreover, $g$ assigns positive mass to every sphere with positive radius.

Since we have assumed that density $g$ is bounded by $B$, the prior mass assigned to the sphere $S[(\mu_1', \mu_2'), r]$ is bounded by $B^2$ times its Euclidean volume. So, take $b_2 = 2$ and $c = \pi B^2$ and the first statement is satisfied. Since we have assumed that $g$ is strictly positive everywhere, the second statement is satisfied. □

## A.15 Proof of Proposition 15

I start with a lemma that shows when the optimal-stopping problem's payoff functions $u_1, u_2$ are Lipschitz continuous, then $(\mu_1, \mu_2) \mapsto U(c'; \mu_1, \mu_2)$, the expected payoff of the stopping strategy $c' \uparrow$ under the subjective model $\Xi(\mu_1, \mu_2; \gamma)$, is locally Lipschitz continuous.

**Lemma A.3.** *Suppose there are constants $K_1, K_2 > 0$ so that $|u_1(x_1') - u_1(x_1'')| < K_1 \cdot |x_1' - x_1''|$ and $|u_2(x_1', x_2') - u_2(x_1'', x_2'')| < K_2 \cdot (|x_1' - x_1''| + |x_2' - x_2''|)$ for all $x_1', x_1'', x_2', x_2'' \in \mathbb{R}$. For each* *chosen center* $(\mu_1^\circ, \mu_2^\circ) \in \mathbb{R}^2$ *and each* $c^{'} \in \mathbb{R}$, *there corresponds a constant* $K > 0$ *so that for any* $\mu_1, \mu_2 \in \mathbb{R}$, $|U(c^{'}; \mu_1, \mu_2) - U(c^{'}; \mu_1^\circ, \mu_2^\circ)| < K \cdot (|\mu_1 - \mu_1^\circ| + |\mu_2 - \mu_2^\circ|)$.

*Proof.* Let $(\mu_1^\circ, \mu_2^\circ) \in \mathbb{R}^2$ and $c^{'} \in \mathbb{R}$ be given. For any $\mu_1, \mu_2 \in \mathbb{R}$, we have

$$\begin{aligned} U(c^{'}; \mu_1, \mu_2) = &\int_{c^{'}}^{\infty} u_1(x_1)\phi(x_1; \mu_1, \sigma^2)dx_1 \\ &+ \int_{-\infty}^{c^{'}} \left[\int_{-\infty}^{\infty} u_2(x_1, x_2)\phi(x_2; \mu_2 + \gamma(x_1 - \mu_1), \sigma^2)dx_2\right] \cdot \phi(x_1; \mu_1, \sigma^2)dx_1 \end{aligned}$$

where $\phi(x; \mu, \sigma^2)$ is the Gaussian density with mean $\mu$, variance $\sigma^2$, evaluated at $x$.

We first bound $|\int_{c^{'}}^{\infty} u_1(x_1)\phi(x_1; \mu_1, \sigma^2)dx_1 - \int_{c^{'}}^{\infty} u_1(x_1)\phi(x_1; \mu_1^\circ, \sigma^2)dx_1|$ by a multiple of $|\mu_1 - \mu_1^\circ|$. Suppose first $\mu_1 = \mu_1^\circ + \Delta$ for some $\Delta > 0$. We have

$$\int_{c^{'}}^{\infty} u_1(x_1)\phi(x_1; \mu_1, \sigma^2)dx_1 = \int_{c^{'}-\Delta}^{\infty} u_1(x_1 + \Delta)\phi(x_1; \mu_1^\circ, \sigma^2)dx_1.$$

By Lipschitz continuity of $u_1$, $|u_1(x_1) - u_1(x_1 + \Delta)| \leq K_1\Delta$ for all $x_1 \in \mathbb{R}$. Thus we conclude

$$|\int_{c^{'}}^{\infty} u_1(x_1)\phi(x_1; \mu_1, \sigma^2)dx_1 - \int_{c^{'}}^{\infty} u_1(x_1)\phi(x_1; \mu_1^\circ, \sigma^2)dx_1| \leq K_1\Delta + |\int_{c^{'}-\Delta}^{c^{'}} u_1(x_1)\phi(x_1; \mu_1^\circ, \sigma^2)dx_1|.$$

Again by Lipschitz continuity of $u_1$, for any $x_1 \in \mathbb{R}$, $|u_1(x_1)\phi(x_1; \mu_1, \sigma^2)| \leq (|u_1(c^{'})| + K_1|x_1 - c^{'}|) \cdot \phi(x_1; \mu_1^\circ, \sigma^2)$. Since the Gaussian density decreases to 0 exponentially fast as $x_1 \to \pm\infty$, the RHS is uniformly bounded for all $x_1$ by some constant, say $J_1 > 0$. This shows that $|\int_{c^{'}-\Delta}^{c^{'}} u_1(x_1)\phi(x_1; \mu_1^\circ, \sigma^2)dx_1| \leq \int_{c^{'}-\Delta}^{c^{'}} |u_1(x_1)\phi(x_1; \mu_1^\circ, \sigma^2)|dx_1 \leq J_1\Delta$. So altogether,

$$|\int_{c^{'}}^{\infty} u_1(x_1)\phi(x_1; \mu_1, \sigma^2)dx_1 - \int_{c^{'}}^{\infty} u_1(x_1)\phi(x_1; \mu_1^\circ, \sigma^2)dx_1| \leq (K_1 + J_1)\Delta.$$

If instead $\mu_1 = \mu_1^\circ - \Delta$, then a similar argument shows that

$$|\int_{c^{'}}^{\infty} u_1(x_1)\phi(x_1; \mu_1, \sigma^2)dx_1 - \int_{c^{'}}^{\infty} u_1(x_1)\phi(x_1; \mu_1^\circ, \sigma^2)dx_1| \leq K_1\Delta + |\int_{c^{'}}^{c^{'}+\Delta} u_1(x_1)\phi(x_1; \mu_1^\circ, \sigma^2)dx_1|,$$

and again we may bound the second term by $J_1\Delta$ as before.

We now turn to bounding the difference in the second summand making up $U(c^{'}; \mu_1, \mu_2)$. First consider the case where $\mu_2 = \mu_2^\circ$. For each $x_1, \mu_1 \in \mathbb{R}$, let $I(x_1; \mu_1) := \int_{-\infty}^{\infty} u_2(x_1, x_2)\phi(x_2; \mu_2^\circ + \gamma(x_1 - \mu_1), \sigma^2)dx_2$, the expected continuation utility after $X_1 = x_1$, in the subjective model $\Gamma(\mu_1, \mu_2^\circ)$. The second summand in $U(c^{'}; \mu_1, \mu_2)$ is given by $\int_{-\infty}^{c^{'}} I(x_1; \mu_1)\phi(x_1; \mu_1, \sigma^2)dx_1$.

For $x_1^{''} = x_1^{'} + d_1$, $\mu_1^{''} = \mu_1^{'} + d_2$, we have

$$
\begin{aligned}
I(x_1^{''}; \mu_1^{''}) &= \int_{-\infty}^{\infty} u_2(x_1^{''}, x_2)\phi(x_2; \mu_2^{\circ} + \gamma(x_1^{''} - \mu_1^{''}), \sigma^2) dx_2 \\
&= \int_{-\infty}^{\infty} u_2(x_1^{'} + d_1, x_2 + \gamma(d_1 - d_2))\phi(x_2; \mu_2^{\circ} + \gamma(x_1^{'} - \mu_1^{'}), \sigma^2) dx_2.
\end{aligned}
$$

Lipschitz continuity of $u_2$ implies that

$$
\begin{aligned}
|u_2(x_1^{'} + d_1, x_2 + \gamma(d_1 - d_2)) - u_2(x_1^{'}, x_2)| &\leq K_2((1 + |\gamma|) \cdot |d_1| + |\gamma| \cdot |d_2|) \\
&\leq K_2(1 + |\gamma|) \cdot (|d_1| + |d_2|),
\end{aligned}
$$

which shows $|I(x_1^{''}; \mu_1^{''}) - I(x_1^{'}; \mu_1^{'})| \leq K_2(1 + |\gamma|) \cdot (|x_1^{'} - x_1^{''}| + |x_2^{'} - x_2^{''}|)$. That is, $I$ is Lipschitz continuous.

Suppose $\mu_1 = \mu_1^{\circ} + \Delta$ for some $\Delta > 0$. Similar to the above argument bounding the first summand in $(c^{'}; \mu_1, \mu_2)$, we have

$$
\int_{-\infty}^{c^{'}} I(x_1; \mu_1)\phi(x_1; \mu_1, \sigma^2) dx_1 = \int_{-\infty}^{c^{'} - \Delta} I(x_1 + \Delta; \mu_1^{\circ} + \Delta)\phi(x_1; \mu_1^{\circ}, \sigma^2) dx_1.
$$

By Lipschitz continuity of $I$, $|I(x_1; \mu_1^{\circ}) - I(x_1 + \Delta; \mu_1^{\circ} + \Delta)| \leq 2K_2(1 + |\gamma|)\Delta$ for all $x_1 \in \mathbb{R}$. Thus we conclude

$$
\begin{aligned}
&|\int_{-\infty}^{c^{'}} I(x_1; \mu_1)\phi(x_1; \mu_1, \sigma^2) dx_1 - \int_{-\infty}^{c^{'}} I(x_1; \mu_1^{\circ})\phi(x_1; \mu_1^{\circ}, \sigma^2) dx_1| \\
&\leq 2K_2(1 + |\gamma|)\Delta + |\int_{c^{'} - \Delta}^{c^{'}} I(x_1; \mu_1^{\circ})\phi(x_1; \mu_1^{\circ}, \sigma^2) dx_1|.
\end{aligned}
$$

Since $x_1 \mapsto I(x_1; \mu_1^{\circ})$ is Lipschitz continuous, there exists $J_2 > 0$ so that $|I(x_1; \mu_1^{\circ})\phi(x_1; \mu_1^{\circ}, \sigma^2)| \leq J_2$ for all $x_1 \in \mathbb{R}$, which means $|\int_{c^{'} - \Delta}^{c^{'}} I(x_1; \mu_1^{\circ})\phi(x_1; \mu_1^{\circ}, \sigma^2) dx_1| \leq J_2\Delta$. The case of $\mu_1 = \mu_1^{\circ} - \Delta$ is symmetric and we have shown that

$$
|\int_{-\infty}^{c^{'}} I(x_1; \mu_1)\phi(x_1; \mu_1, \sigma^2) dx_1 - I(x_1; \mu_1^{\circ})\phi(x_1; \mu_1^{\circ}, \sigma^2) dx_1| \leq (2K_2(1 + |\gamma|) + J_2) \cdot |\mu_1 - \mu_1^{\circ}|.
$$

Finally, we investigate the difference in the second summand of $U(c^{'}; \mu_1, \mu_2)$ between parameters $(\mu_1, \mu_2^{\circ})$ and $(\mu_1, \mu_2)$ for $\mu_1, \mu_2 \in \mathbb{R}$. This difference is bounded by

$$
\int_{-\infty}^{c^{'}} \left| \int_{-\infty}^{\infty} u_2(x_1, x_2)\phi(x_2; \mu_2^{\circ} + \gamma(x_1 - \mu_1), \sigma^2) dx_2 - \int_{-\infty}^{\infty} u_2(x_1, x_2)\phi(x_2; \mu_2 + \gamma(x_1 - \mu_1), \sigma^2) dx_2 \right| \phi(x_1; \mu_1, \sigma^2) dx_1. \tag{5}
$$

But for every $x_1 \in \mathbb{R}$,

$$\int_{-\infty}^{\infty} u_2(x_1, x_2)\phi(x_2; \mu_2+\gamma(x_1-\mu_1), \sigma^2)dx_2 = \int_{-\infty}^{\infty} u_2(x_1, x_2+(\mu_2-\mu_2^\circ))\phi(x_2; \mu_2^\circ+\gamma(x_1-\mu_1), \sigma^2)dx_2,$$

and $|u_2(x_1, x_2 + (\mu_2 - \mu_2^\circ)) - u_2(x_1, x_2)| \le K_2|\mu_2 - \mu_2^\circ|$ by Lipschitz continuity of $u_2$. This shows that, for all values $\mu_1, \mu_2 \in \mathbb{R}$, (5) is bounded by $K_2|\mu_2 - \mu_2^\circ|$.

Applying the triangle inequality to the second term, we conclude that

$$|U(c^{'}; \mu_1, \mu_2) - U(c^{'}; \mu_1^\circ, \mu_2^\circ)| \le (K_1 + J_1)|\mu_1 - \mu_1^\circ| + (2K_2(1+|\gamma|) + J_2) \cdot |\mu_1 - \mu_1^\circ| + K_2|\mu_2 - \mu_2^\circ|.$$

So we see that setting $K = K_1 + J_1 + (2K_2(1 + |\gamma|) + J_2$ establishes the lemma. □

Now I prove Proposition 15.

*Proof.* Let $c, c^{'} \in \mathbb{R}$ be given and let $\mu_1^\circ = \mu_1^\bullet$, $\mu_2^\circ = \mu_2^*(c)$. Lemma A.3 implies there is a constant $K > 0$ so that $|U(c^{'}; \mu_1, \mu_2) - U(c^{'}; \mu_1^\circ, \mu_2^\circ)| \le K \cdot (|\mu_1 - \mu_1^\circ| + |\mu_2 - \mu_2^\circ|)$ for all $\mu_1, \mu_2 \in \mathbb{R}$, so for $\nu$ a joint distribution about the fundamentals $(\mu_1, \mu_2)$, we get

$$\begin{aligned} |\mathbb{E}_{(\mu_1,\mu_2)\sim\nu}\left[U(c^{'}; \mu_1, \mu_2) - U(c^{'}; \mu_1^\circ, \mu_2^\circ)\right]| &\le \mathbb{E}_{(\mu_1,\mu_2)\sim\nu}\left[|U(c^{'}; \mu_1, \mu_2) - U(c^{'}; \mu_1^\circ, \mu_2^\circ)|\right] \\ &\le K \cdot \mathbb{E}_{(\mu_1,\mu_2)\sim\nu}[|\mu_1 - \mu_1^\circ| + |\mu_2 - \mu_2^\circ|]. \end{aligned}$$

By Lemma 6, almost surely

$$\lim_{N\to\infty} \mathbb{E}_{(\mu_1,\mu_2)\sim g(\cdot|(X_n,Y_n)_{n=1}^N)}[|\mu_1 - \mu_1^\circ| + |\mu_2 - \mu_2^\circ|] = 0.$$

But along any realized sequence of $(X_n, Y_n)_{n=1}^\infty$ where the above limit holds, by putting $\nu = \tilde{g}_N$ we also have $\lim_{N\to\infty} |U_N(c^{'}) - U(c^{'}; \mu_1^\circ, \mu_2^\circ)| = 0$. This shows $U_N(c^{'})$ converges to $U(c^{'}; \mu_1^*(c), \mu_2^*(c))$ almost surely as $N \to \infty$. □

# B General Subjective Joint Distribution of $(X_1, X_2)$ with Method of Moments Inference

The analysis so far has assumed that both the objective joint distribution of $(X_1, X_2)$ as well as all of the subjective models of $(X_1, X_2)$ that the agent deems plausible are Gaussian. The Gaussian assumption makes the agents' inference problem analytically tractable, since the KL divergence between a pair of Gaussian distributions has a simple closed-form expression. However, the intuition that when an agent holds negatively correlated beliefs about the fundamentals, the censoring effect enables a positive feedback loop between distorted stopping rules and distorted beliefs holds more generally. To show this, I modify the inference procedure of the agents to a simpler but natural alternative: agents infer fundamentals as to match observed dataset of histories in terms of certain moments. As discussed in Remark 4, under the Gaussian assumption this method-of-moments (MOM) estimation procedure is identical to KL divergence minimization.

## B.1 Subjective Models for $(X_1, X_2)$

Each agent starts with a family of subjectively models $\{\mathbb{M}(\cdot;\theta_1,\theta_2) : \theta_1 \in \Theta_1, \theta_2 \in \Theta_2\}$ for the joint distribution of $(X_1, X_2)$, with parameter spaces $\Theta_1 \subseteq \mathbb{R}$ and $\Theta_2 \subseteq \mathbb{R}$. For each $(\theta_1\theta_2)$, $\mathbb{M}(\cdot;\theta_1,\theta_2)$ is a full-support measure on the rectangle $I_1 \times I_2$, where each $I_1, I_2$ is a possibly infinite interval of $\mathbb{R}$. By "full-support" I mean that for every open ball $B \subseteq I_1 \times I_2$, $\mathbb{M}(B;\theta_1,\theta_2) > 0$.

For each joint distribution $\mathbb{M}(\cdot;\theta_1,\theta_2)$, let $\mathbb{M}_1(\cdot;\theta_1,\theta_2)$ denote its marginal on $I_1$, and let $\mathbb{M}_{2|1}(\cdot|\theta_1,\theta_2;x_1)$ denote its conditional distribution of $X_2$ given $X_1 = x_1$. I will make the following assumptions on the family of distributions $\mathbb{M}$:

**Assumption A.1.** $\mathbb{M}_1(\cdot;\theta_1,\theta_2)$ *is only a function of* $\theta_1$ *and* $\mathbb{E}_{\mathbb{M}_1(\cdot;\theta_1,\theta_2)}[X_1]$ *is strictly increasing in* $\theta_1$*.*

In light of this assumption, the marginal distribution on $X_1$ can be just written as $\mathbb{M}_1(\cdot;\theta_1)$, omitting $\theta_2$.

**Assumption A.2.** *For each* $x_1 \in I_1$ *and* $\theta_1 \in \Theta_1$, $\mathbb{E}_{\mathbb{M}_{2|1}(\cdot;\theta_1,\theta_2|x_1)}[X_2]$ *strictly increases in* $\theta_2$*.*

**Assumption A.3.** *For any* $\theta_1 \in \Theta_1$ *and* $\theta_2 \in \Theta_2$, $\mathbb{E}_{\mathbb{M}_{2|1}(\cdot;\theta_1,\theta_2|x_1)}[X_2]$ *strictly decreases in* $x_1$*.*

Assumption A.3 is the substantive assumption capturing the gambler's fallacy psychology. Every subjective distribution in the family is such that the agent predicts a lower mean for $X_2$ after a higher realization of $X_1$.

Here are some examples satisfying these assumptions. The first example shows the family of Gaussian distributions I have been working with satisfies these assumptions.

**Example A.1.** Let $I_1 = I_2 = \mathbb{R}$ and let $\Theta_1 = \Theta_2 = \mathbb{R}$. Fixing some $\sigma^2 > 0$, $\gamma < 0$, let $\mathbb{M}(\cdot;\theta_1,\theta_2)$ be $\Xi(\theta_1,\theta_2,\sigma^2,\sigma^2;\gamma)$ for each $\theta_1,\theta_2 \in \mathbb{R}$. The marginal distribution on $X_1$ is $\mathcal{N}(\theta_1,\sigma^2)$ and does not depend on $\theta_2$. Its mean is $\theta_1$ so it strictly increases in $\theta_1$. The conditional mean of $X_2|X_1 = x_1$ is $\theta_2 + \gamma(x_1 - \theta_1)$, which is strictly increasing in $\theta_2$ and strictly decreasing in $x_1$ since $\gamma < 0$. So Assumptions A.1, A.2, and A.3 are satisfied.

The next example features bivariate exponential distributions supported on the half-line $[0, \infty)$.

**Example A.2.** Gumbel (1960) proposes the following family of bivariate exponential distributions, parametrized by $\alpha \in [-1, 1]$ : consider a joint distribution with the density function $(\tilde{x}_1, \tilde{x}_2) \mapsto e^{-\tilde{x}_1 - \tilde{x}_2} \cdot [1 + \alpha(2e^{-\tilde{x}_1} - 1) \cdot (2e^{-\tilde{x}_2} - 1)]$ on $\tilde{x}_1, \tilde{x}_2 \geq 0$. If $(\tilde{X}_1, \tilde{X}_2)$ are random variables with this density, then they have full support on $[0, \infty) \times [0, \infty)$, each $\tilde{X}_j$ has the marginal distribution of an exponential random variable with mean 1, $\mathbb{E}[\tilde{X}_2 | \tilde{X}_1 = \tilde{x}_1] = 1 - \frac{1}{2}\alpha - \alpha e^{-\tilde{x}_1}$. The correlation between $\tilde{X}_1$ and $\tilde{X}_2$ is $\alpha/4$.

Let $I_1 = I_2 = [0, \infty)$ and let $\Theta_1 = \Theta_2 = (0, \infty)$. Fixing some $-1 \leq \alpha < 0$, let $\mathbb{M}(\cdot; \theta_1, \theta_2)$ be the joint distribution generated by $X_1 = \theta_1 \cdot \tilde{X}_1$ and $X_2 = \theta_2 \cdot \tilde{X}_2$ where $(\tilde{X}_1, \tilde{X}_2)$ have the Gumbel bivariate distribution with parameter $\alpha$. Since $(\tilde{X}_1, \tilde{X}_2)$ have full support on $I_1 \times I_2$, the same holds for $(X_1, X_2)$ for any $\theta_1, \theta_2 > 0$. The marginal distribution of $X_1$ is exponential with a mean of $\theta_1$, so Assumption A.1 is satisfied. The conditional mean of $X_2 | X_1 = x_1$ is given by $\mathbb{E}[\theta_2 \tilde{X}_2 | \theta_1 \tilde{X}_1 = x_1] = \theta_2 \cdot \mathbb{E}\left[\tilde{X}_2 | \tilde{X}_1 = \frac{x_1}{\theta_1}\right] = \theta_2 \cdot \left(1 - \frac{1}{2}\alpha - \alpha e^{-(x_1/\theta_1)}\right)$. As $\alpha < 0$, the term inside the bracket is strictly positive. So this conditional expectation is strictly increasing in $\theta_2$, showing that Assumption A.2 is satisfied. Also, since $\theta_1, \theta_2 > 0$, $x_1 \mapsto -\alpha\theta_2 e^{-(x_1/\theta_1)}$ is strictly decreasing and so Assumption A.3 is satisfied.

Finally I give another example where $I_1 = I_2 = [0, 1]$ are bounded intervals.

**Example A.3.** Let $\Theta_1 = \Theta_2 = (0, \infty)$ and consider the family of distribution $\mathbb{M}(\cdot; \theta_1, \theta_2)$ such that under parameters $(\theta_1, \theta_2)$, $X_1 \sim \text{Beta}(\theta_1, 1)$ and $X_2 | X_1 = x_1 \sim \text{Beta}((1 - x_1)\theta_2, 1)$. For any values of $\theta_1, \theta_2 > 0$, $X_1$ has full support on $[0, 1]$. Conditional on any $x_1 \in (0, 1)$, $X_2$ has full support on $[0, 1]$. This shows the distribution $\mathbb{M}(\cdot; \theta_1, \theta_2)$ has full-support on $[0, 1]^2$ for every $(\theta_1, \theta_2) \in \Theta_1 \times \Theta_2$. The mean of $X_1$ is $\frac{\theta_1}{\theta_1 + 1}$, which only depends on $\theta_1$ and is strictly increasing in it. So Assumption A.1 is satisfied. The conditional mean of $X_2 | X_1 = x_1$ is $\frac{(1 - x_1)\theta_2}{(1 - x_1)\theta_2 + 1}$, which is strictly increasing in $\theta_2$ and strictly decreasing in $x_1$. So, Assumptions A.2 and A.3 are satisfied.

Finally, I give a general class of examples that allows any pair of given marginal distributions for $X_1$ and $X_2$ to be joined together using a copula as to induce negative dependence for the joint distribution.

**Example A.4.** Consider two families of distribution functions $Q_1(\cdot; \theta_1) : I_1 \to [0, 1]$, $Q_2(\cdot; \theta_2) : I_2 \to [0, 1]$, such $Q_1$ and $Q_2$ are supported on $I_1, I_2$ respectively under all parameters. Suppose the mean of $Q_1$ is increasing in $\theta_1$, and $Q_2$ is increasing in stochastic dominance order for $\theta_2$. Fix a differentiable copula: that is, a differentiable function $\mathcal{C} : [0, 1]^2 \to [0, 1]$ so that $\mathcal{C}(u, 0) = \mathcal{C}(0, v) = 0$, $\mathcal{C}(u, 1) = u$, $\mathcal{C}(1, v) = v$ for all $u, v \in [0, 1]$, and so that for $u_1 \leq u_2$, $v_2 \leq v_2 \in [0, 1]$, we get $\mathcal{C}(u_2, v_2) - \mathcal{C}(u_2, v_1) - \mathcal{C}(u_1, v_2) - \mathcal{C}(u_1, v_1) \geq 0$. Consider the family of joint distributions $\mathbb{M}(\cdot; \theta_1, \theta_2)$ generated by joining together $Q_1(\cdot; \theta_1)$ with $Q_2(\cdot; \theta_2)$

using the copula $\mathcal{C}$, namely

$$\mathbb{M}((-\infty, x_1] \times (-\infty, x_2]; \theta_1, \theta_2) = \mathcal{C}(Q_1^{-1}(x_1|\theta_1), Q_2^{-1}(x_2|\theta_2)).$$

Then $\mathbb{M}(\cdot; \theta_1, \theta_2)$ has marginal distributions on $X_1$ and $X_2$ given by distribution functions $Q_1(\cdot; \theta_1)$, $Q_2(\cdot; \theta_2)$, and:

**Lemma A.4.** *Provided* $\frac{\partial \mathcal{C}}{\partial u}(u, v)$ *is an increasing function in* $u$*, Assumptions A.1, A.2, and A.3 are satisfied for the family of distributions* $\mathbb{M}(\cdot; \theta_1, \theta_2)$.

The condition that $\frac{\partial \mathcal{C}}{\partial u}(u, v)$ increases in $u$ is satisfied by, for example, the Gaussian copula with any negative correlation. The derivative of the Gaussian copula is given by $\frac{\partial \mathcal{C}}{\partial u}(u, v) = \mathbb{P}[X_2 \leq \Phi^{-1}(v)|X_1 = \Phi^{-1}(u)]$ where $\Phi$ is the standard Gaussian distribution function and $(X_1, X_2)$ are jointly Gaussian with correlation $-1 < \rho < 0$ and each with an unconditional variance of 1. As it is known that $X_2|X_1 = x_1 \sim \mathcal{N}(\rho x_1, 1 - \rho^2)$, it is clear that $X_2|X_1 = x_1$ decreases in FOSD order as $x_1$ increases, so for any $v$ we have $\mathbb{P}[X_2 \leq \Phi^{-1}(v)|X_1 = \Phi^{-1}(u)]$ increases in $u$.

## B.2 Method of Moments Inference

For a distribution $\mathcal{H}$ on the space of histories $H$, let $m_1[\mathcal{H}]$ represent the average first-period draw under this distribution and let $m_2[\mathcal{H}]$ represent the average second-period draw (when uncensored). Suppose that objectively $X_1, X_2$ are drawn from two independent distributions and denote the true distribution of histories under censoring by cutoff stopping rule $c \in \mathbb{R}$ as $\mathcal{H}^\bullet(c\uparrow)$. Then by independence, $m_1[\mathcal{H}^\bullet(c\uparrow)]$ and $m_2[\mathcal{H}^\bullet(c\uparrow)]$ do not in fact depend on $c$.

Given the family of subjective models $\{\mathbb{M}(\cdot; \theta_1, \theta_2) : \theta_1 \in \Theta_1, \theta_2 \in \Theta_2\}$ about the joint distribution of $(X_1, X_2)$, let $\mathcal{H}(\theta_1, \theta_2; c\uparrow) := \mathcal{H}(\mathbb{M}(\cdot; \theta_1, \theta_2); c\uparrow)$ denote the distribution on histories under joint distribution $\mathbb{M}(\cdot; \theta_1, \theta_2)$ and censoring cutoff $c$. I now define the method of moments estimator.

**Definition A.1.** The **method-of-moments (MOM) estimator** derived from an infinite dataset with history distribution $\mathcal{H}^\bullet(c\uparrow)$ is any pair $(\theta_1^M, \theta_2^M) \in \Theta_1 \times \Theta_2$ such that:

1. $m_1[\mathcal{H}(\theta_1^M, \theta_2^M; c\uparrow)] = m_1[\mathcal{H}^\bullet(c\uparrow)]$
2. $m_2[\mathcal{H}(\theta_1^M, \theta_2^M; c\uparrow)] = m_2[\mathcal{H}^\bullet(c\uparrow)]$

I will sometimes write $\theta_1^M(c)$, $\theta_2^M(c)$ to emphasize the dependence of the MOM estimators on the censoring threshold $c$. The MOM estimator need not exist — for example, if all values of $\theta_1 \in \Theta_1$ generate a marginal distribution on $X_1$ that is smaller than $m_1[\mathcal{H}^\bullet(c\uparrow)]$. However, when it exists, it is unique under the assumptions I made.

**Lemma A.5.** *When the family of distributions $\mathbb{M}$ satisfies Assumptions A.1, A.2, and A.3, the MOM estimator is unique when it exists.*

Now I show the MOM estimators have the same monotonicity property in $c$ as the pseudo-true fundamentals minimizing KL divergence for Gaussian distributions, a result comparable with the final statement of Proposition 1. So, this key ingredient driving the positive feedback cycle in biased agents' learning does not depend on the Gaussian assumption per se. Rather, the crucial assumption is the generalized notion of negative dependence between $X_1$ and $X_2$, as articulated by Assumption A.3 for arbitrary joint distributions. Along with regularity conditions in Assumptions A.1 and A.2, agents using a natural method of moments procedure will end up with more pessimistic beliefs about the second-period fundamental when the dataset is more severely censored.

**Proposition A.1.** *Suppose Assumptions A.1, A.2, and A.3 hold. Suppose $c^{'} < c^{''}$ are two different interior values in $I_1$ and that MOM estimators $(\theta_1^M(c^{'}), \theta_2^M(c^{'}))$ and $(\theta_1^M(c^{''}), \theta_2^M(c^{''}))$ exist. Then $\theta_1^M(c^{'}) = \theta_1^M(c^{''})$ and $\theta_2^M(c^{'}) < \theta_2^M(c^{''})$.*

As a corollary, I characterize the large-generations learning dynamics for method of moments agents using a general class of subjective models about $(X_1, X_2)$. The key idea is that the positive feedback between distorted stopping rules and distorted beliefs continue to hold, with the parametric version of gambler's fallacy interpreted as $\gamma < 0$ in a specific Gaussian setup replaced with the general notion of negative dependence as in Assumption A.3.

One caveat: we must now ensure the MOM estimator exists in each generation when the previous generation uses any cutoff stopping rule that has a positive probability of continuing into the next period. To guarantee existence, I impose an additional restriction on how $\mathbb{M}$ relates to the true distribution of $(X_1, X_2)$.

**Assumption A.4.** *(a) The objective supports of $X_1$ and $X_2$ are $I_1$ and $I_2$, respectively.*

*(b) The range of $\theta_1 \mapsto \mathbb{E}_{\mathbb{M}_1(\cdot;\theta_1)}[X_1]$ is $I_1$.*

*(c) For every $\theta_1 \in \Theta_1$ and $x_1 \in I_1$, the range of $\theta_2 \mapsto \mathbb{E}_{\mathbb{M}_{2|1}(\cdot;\theta_1,\theta_{2|x_1})}[X_2]$ is $I_2$.*

Assumption A.4(a) is a consistency requirement, saying that the supports for the objective distributions of $X_1$ and $X_2$ match their supports under the agents' subjective models. Assumption A.4(b) and Assumption A.4(c) ensures the agents can always match the two moment conditions. It is easily verified that Examples A.1, A.2, and A.3 satisfy Assumption A.4 when the true joint distribution of $(X_1, X_2)$ is supported on $\mathbb{R}^2$, $[0, \infty)^2$, and $[0, 1]^2$ respectively.

**Corollary A.1.** *Fix some objective, independent distribution for $(X_1, X_2)$ and suppose agents have a family of subjective models $\{\mathbb{M}(\cdot;\theta_1,\theta_2) : \theta_1 \in \Theta_1, \theta_2 \in \Theta_2\}$ satisfying Assumptions A.1, A.2, A.3, and A.4. Suppose the stopping problem is such that the payoff function $u_2(x_1, x_2)$ is linear in $x_2$. Suppose each generation $t \geq 1$ believes in the model estimated using MOM from previous generation's cutoff stopping strategy $c_{[t-1]} \uparrow$, namely $\mathbb{M}(\cdot;\theta_1^M(c_{[t-1]}), \theta_2^M(c_{[t-1]}))$, provided $c_{[t-1]} > \inf(I_1)$.*

*Let the 0th generation choose an arbitrary cutoff $c_{[0]}$ in the interior of $I_1$. Then, up until the first period $T$ where $c_{[T]} \leq inf(I_1)$ (and MOM becomes undefined subsequently), beliefs and cutoff rules $(\mu^M_{1,[t]})_{t=1}^T$, $(\mu^M_{2,[t]})_{t=1}^T$, and $(c_{[t]})_{t=1}^T$ form monotonic sequences.*

This corollary establishes the monotonicity of the beliefs and cutoffs up until when some generation decides to always stop. If this happens, MOM is no longer well-defined since second-period draw is never observed.

# C Optimal-Stopping Problems with $L$ Periods

## C.1 An $L$-Periods Model of the Gambler's Fallacy

In an optimal-stopping problem with $L$ periods, the agent observes a draw $x_\ell \in \mathbb{R}$ in each period $1 \leq \ell \leq L$. At the end of period $\ell$, the agent must decide between stopping and receiving a payoff $u_\ell(x_1, ..., x_\ell)$ that depends on the profile of draws $(x_i)_{i=1}^\ell$ observed so far, or continuing into the next period. If the agent continues into period $L$ without stopping, then his payoff will be $u_L(x_1, ..., x_L)$.

I first introduce notation for a class of joint distributions of the $L$ possible draws $(X_i)_{i=1}^L$.

**Definition A.2.** Let $\sigma^2 > 0$ be fixed. For every vector $\boldsymbol{\mu} = (\mu_i)_{i=1}^L$ and triangular array $\boldsymbol{\gamma} = (\gamma_{i,j})_{2\leq i\leq L, 1\leq j\leq i-1}$ with each $\gamma_{i,j} \in \mathbb{R}$, the **subjective model** $\Xi(\boldsymbol{\mu};\boldsymbol{\gamma})$ denotes the joint distribution of $(X_i)_{i=1}^L$ where $X_1 \sim \mathcal{N}(\mu_1, \sigma^2)$ and, for all $i \geq 2$ and $(x_j)_{j=1}^{i-1} \in \mathbb{R}^{i-1}$,

$$X_i|(X_1 = x_1, ..., X_{i-1} = x_{i-1}) \sim \mathcal{N}(\mu_i + \sum_{j=1}^{i-1} \gamma_{i,j} \cdot (x_j - \mu_j), \sigma^2).$$

Under $\Xi(\boldsymbol{\mu};\boldsymbol{\gamma})$, $(X_i)_{i=1}^L$ are jointly Gaussian,[17] such that the conditional mean of $X_i$ given the previous draws $X_1 = x_1, ..., X_{i-1} = x_{i-1}$ depends linearly on these realizations. I consider agents who entertain a set of subjective models, $\{\Xi(\boldsymbol{\mu};\boldsymbol{\gamma}) : \boldsymbol{\mu} \in \mathbb{R}^L\}$ for a fixed array $\boldsymbol{\gamma}$ where each $\gamma_{i,j} < 0$. The negative $\gamma_{i,j}$ capture the gambler's fallacy, as higher realizations of earlier

[17] An equivalent description of the subjective model $\Xi(\boldsymbol{\mu};\boldsymbol{\gamma})$ is to consider a set of $L$ independent Gaussian random variables $Z_i \sim \mathcal{N}(\mu_i, \sigma^2)$ for $1 \leq i \leq L$. Let $X_1 = Z_1$ and iteratively define $X_i = Z_i + \sum_{j=1}^{i-1} \gamma_{i,j}(X_j - \mu_j)$. Using induction, one can show that every $X_i$ is a linear function of the $Z_i$'s, so they are jointly Gaussian.

draws lead agents to predict lower means for future draws. The greater the magnitude of $\gamma_{i,j}$, the more that the agent's prediction of $X_i$ depends on realization of $X_j$. Agents hold a dogmatic belief in the correlation structure between $(X_i)_{i=1}^L$, but can flexibly estimate $(\mu_i)_{i=1}^L$, the **fundamentals** of the environment. Objectively, $(X_i)_{i=1}^L$ are independent, so the true joint distribution is $\Xi^\bullet = \Xi(\boldsymbol{\mu}^\bullet; \mathbf{0})$ for some $(\mu_i^\bullet)_{i=1}^L$.

A useful functional form to keep in mind is $\gamma_{i,j} = -\alpha \cdot \delta^{i-j-1}$ for $\alpha > 0$, $0 \le \delta \le 1$, which corresponds to Rabin and Vayanos (2010)'s specification of gambler's fallacy in multiple periods. Here, $\alpha$ relates to the severity of the bias and $\delta$ captures how quickly the influence of past observations decay in predicting future draws.

## C.2 Inference from Censored Datasets in $L$ Periods

In general, a stopping strategy in an optimal-stopping problem over $L$ periods is a set of functions $s_i : \mathbb{R}^i \to \{\text{Stop}, \text{Continue}\}$ for $1 \le i \le L-1$, where $s_i(x_1, ..., x_i)$ maps the realizations of the first $i$ draws to a stopping decision. I consider stopping strategies where $s_i$ is a cutoff rule in $x_i$ after each partial history $(x_1, ..., x_{i-1})$, that is there exist $(c_i)_{i=1}^{L-1}$ with $c_1 \in \mathbb{R}$ and for $i \ge 2$, $c_i(x_1, ..., x_{i-1}) \in \mathbb{R}$ for every $(x_1, ..., x_{i-1}) \in \mathbb{R}^{i-1}$, so that the agent stops after $(x_1, ...x_i)$ if and only if $x_i \ge c_i(x_1, ..., x_{i-1})$. A stopping strategy with stopping regions characterized by a profile of cutoff rules $\boldsymbol{c} = (c_i)_{i-1}^{L-1}$ will be abbreviated as $\boldsymbol{c}\uparrow$.

For subjective model $\Xi$ and cutoff rule $\boldsymbol{c}\uparrow$, let $\mathcal{H}(\Xi; \boldsymbol{c}\uparrow)$ represent the distribution of histories when applying rule $\boldsymbol{c}\uparrow$ to draws $(X_i) \sim \Xi$. More precisely, consider a procedure where $X_1, X_2, ..., X_L$ is drawn according to $\Xi$ and revealed one at a time. At the earliest $1 \le \bar{i} \le L-1$ such that $X_{\bar{i}} \ge c_{\bar{i}}(X_1, ..., X_{\bar{i}-1})$, the process stops and the history records $(X_1, ..., X_{\bar{i}}, \varnothing, ..., \varnothing)$, with $L - \bar{i}$ instances of the censoring indicator $\varnothing$ replacing the unobserved subvector $(X_{\bar{i}+1}, ..., X_L)$. If no such $\bar{i}$ exists, then history records the entire profile of draws, $(X_1, ..., X_L)$. The distribution of histories generated this way is denoted $\mathcal{H}(\Xi; \boldsymbol{c}\uparrow)$.

**Definition A.3.** For cutoff strategy $\boldsymbol{c}\uparrow$ and fundamentals $\hat{\boldsymbol{\mu}}$, the **KL divergence** between objective distribution of histories and the predicted distribution under censoring is the sum of $L$ integrals,

$$D_{KL}(\ \mathcal{H}(\Xi^\bullet; \boldsymbol{c}\uparrow) \ || \ \mathcal{H}(\Xi(\boldsymbol{\mu}; \boldsymbol{\gamma}); \boldsymbol{c}\uparrow)\ ) := \sum_{i=1}^{L} I_i,$$

where

$$I_1 = \int_{c_1}^{\infty} \phi(x_1; \mu_1^\bullet, \sigma^2) \ln\left(\frac{\phi(x_1; \mu_1^\bullet, \sigma^2)}{\phi(x_1; \mu_1, \sigma^2)}\right) dx_1,$$

and for $2 \le i \le L-1$, integral $I_i$ is

$$\int_{-\infty}^{c_1} ... \int_{-\infty}^{c_{i-1}(x_1,...,x_{i-2})} \int_{c_i(x_1,...,x_{i-1})}^{\infty} \prod_{k=1}^{i} \phi(x_k; \mu_k^\bullet, \sigma^2) \ln\left(\frac{\prod_{k=1}^{i} \phi(x_k; \mu_k^\bullet, \sigma^2)}{\prod_{k=1}^{i} \phi(x_k; \mu_k + \sum_{j=1}^{k-1} \gamma_{k,j} \cdot (x_j - \mu_j), \sigma^2)}\right) dx_i ... dx_1.$$

Finally, $I_L$ is given by

$$\int_{-\infty}^{c_1} \cdots \int_{-\infty}^{c_{L-1}(x_1,...,x_{L-2})} \int_{-\infty}^{\infty} \cdot \prod_{k=1}^{i} \phi(x_k;\mu_k^\bullet,\sigma^2) \ln\left(\frac{\prod_{k=1}^{i}\phi(x_k;\mu_k^\bullet,\sigma^2)}{\prod_{k=1}^{i}\phi(x_k;\mu_k+\sum_{j=1}^{k-1}\gamma_{k,j}\cdot(x_j-\mu_j),\sigma^2)}\right) dx_i...dx_1.$$

To interpret, consider a history $h=(x_1,...,x_i,\varnothing,...,\varnothing)$ where $x_k<c_k(x_1,...,x_{k-1})$ for all $k\le i-1$ and $x_i\ge c_i(x_1,...,x_{i-1})$. This history is possible under the stopping strategy $\boldsymbol{c}\uparrow$. It has a likelihood of $\Pi_{k=1}^{i}\phi(x_k;\mu_k^\bullet,\sigma^2)$ under $\Xi^\bullet$ and a likelihood of $\Pi_{k=1}^{i}\phi(x_k;\mu_k+\sum_{j=1}^{k-1}\gamma_{k,j}\cdot(x_j-\mu_j),\sigma^2)$ under $\Xi(\boldsymbol{\mu};\boldsymbol{\gamma})$. So, the integral $I_i$ calculates the contribution of all possible histories of length $i$ to the KL divergence from $\mathcal{H}(\Xi(\boldsymbol{\mu};\boldsymbol{\gamma});\boldsymbol{c}\uparrow)$ to $\mathcal{H}(\Xi^\bullet;\boldsymbol{c}\uparrow)$. In the case of $L=2$, this definition reduces to Definition 5, the KL divergence in the two-periods baseline model, with $\gamma=\gamma_{2,1}$ and $c_1\in\mathbb{R}$ as the censoring threshold.

The KL-divergence minimizers

$$\min_{\boldsymbol{\mu}\in\mathbb{R}^L} D_{KL}(\ \mathcal{H}(\Xi^\bullet;\boldsymbol{c}\uparrow)\ ||\ \mathcal{H}(\Xi(\boldsymbol{\mu};\boldsymbol{\gamma});\boldsymbol{c}\uparrow)\ )$$

are the **pseudo-true fundamentals** with respect to stopping strategy $\boldsymbol{c}\uparrow$. The next proposition gives an explicit characterization of them.

**Proposition A.2.** *Let stopping strategy $\boldsymbol{c}\uparrow$ be given. For each $i\ge1$, let $R_i$ represent the region*

$$\{(x_1,...,x_i): x_1<c_1, x_2<c_2(x_1),...,x_i<c_i(x_1,..,x_{i-1})\}\subseteq\mathbb{R}^i.$$

*The pseudo-true fundamentals with respect to $\boldsymbol{c}\uparrow$ are $\mu_1^*=\mu_1^\bullet$ and, iteratively,*

$$\hat{\mu}_i^*=\mu_i^\bullet+\sum_{j=1}^{i-1}\gamma_{i,j}\cdot(\mu_j^*-\mathbb{E}_{\Xi^\bullet}[X_j|(X_k)_{k=1}^{i-1}\in R_{i-1}]).$$

The expression for $\mu_i^*$ in the general $L$-periods setting resembles the expression for $\mu_2^*$ in the two-period setting. Relative to the truth $\mu_i^\bullet$, the estimate $\mu_i^*$ is distorted by the fact that $X_i$ is only observed when previous draws $(X_1,...,X_{i-1})$ fall into the continuation region $R_{i-1}\subseteq\mathbb{R}^{i-1}$ associated with $\boldsymbol{c}\uparrow$. The agent uses this censored empirical distribution of $(X_1,...,X_{i-1},X_i)$ to infer the period-$i$ fundamental, under a dogmatic belief about the correlation structure between the draws given by $\boldsymbol{\gamma}$. Importantly, whether a certain realization $X_j$ for $j<i$ should be judged as below-average (and thus predict a higher $X_i$) or above-average (and thus predict a lower $X_i$) depends on agent's belief about the period $j$ fundamental, $\mu_j^*$, which gives the iterative structure of the expression for $\hat{\mu}_i^*$.

The proof of this result follows two steps. First, recall that $D_{KL}(\mathcal{H}(\Xi^\bullet;\boldsymbol{c}\uparrow)||\mathcal{H}(\Xi(\boldsymbol{\mu};\boldsymbol{\gamma});\boldsymbol{c}\uparrow))$ is defined as the sum $\sum_{i=1}^{L}I_i$, where $I_i$ is the KL-divergence contribution from histories with length $i$. I rewrite this expression as the sum of $L$ different integrals, $\sum_{i=1}^{L}J_i$, where

$J_i$ is the KL-divergence contributions from histories containing $X_i$. So, $J_i$ is a function of $\mu_1, ..., \mu_i$. The second step is similar to the proof of Proposition 1, where I show $\frac{\partial J_i}{\partial \mu_j}$ is a linear multiple of $\frac{\partial J_i}{\partial \mu_i}$ whenever $j < i$. First-order condition at $\boldsymbol{\mu}^*$ allows for a telescoping rearrangement, yielding $\frac{\partial J_i}{\partial \mu_i}(\boldsymbol{\mu}^*) = 0$ for every $i$. The proposition readily follows.

Now I turn to a special class of cutoff-based stopping rules where $c_k$ is independent of history. So, a stopping rule of this kind $\boldsymbol{c}\uparrow$ can be viewed simply as a list of $L$ constants, $c_1, ..., c_L \in \mathbb{R}$, such that the agent stops after the draw $X_\ell = x_\ell$ if and only if $x_\ell < c_\ell$. I show that the expression for the pseudo-true fundamentals greatly simplifies and admits a path-counting interpretation.

**Definition A.4.** For $1 \leq j < i \leq L$, a **path** $p$ from $i$ to $j$ is a sequence of pairs $p = ((i_0, i_1), ..., (i_{M-1}, i_M))$ with $M \geq 1$, $i_0 = i$, $i_M = j$, and $i_{m+1} < i_m$ for all $m = 0, 1, ..., M-1$. The length of $p$ is $\#(p) := M$. The weight of $p$ is $W(p) := \Pi_{0 \leq m \leq M-1} \gamma_{i_\ell, i_{\ell+1}}$. Denote the set of all paths from $i$ to $j$ as $P[i \to j]$.

That is, we may imagine a network with $L$ nodes, one per period of the optimal-stopping problem. There is a directed edge with weight $\gamma_{i,j}$ for all pairs $i > j$. A path from $i$ to $j$ is a concatenation of edges, starting with $i$ and ending with $j$. Its weight is the product of the weights of all the edges used.

The next proposition differs from Proposition A.2 in that the expression for the pseudo-true fundamental $\mu_i^*$ does not involve other pseudo-true fundamentals $\mu_j^*$. It shows that the distortion of $\mu_i^*$ from the true value $\mu_i^\bullet$ depends on terms $\mu_j^\bullet - \mathbb{E}_{\Xi^\bullet}[X_j | X_j \leq c_j]$ and the total number of paths from $i$ to $j$ in the network that $\boldsymbol{\gamma}$ defines.

**Proposition A.3.** *For stopping strategy $\boldsymbol{c}\uparrow = (c_1, ..., c_L) \in \mathbb{R}^L$, the pseudo-true fundamentals are given by*

$$\mu_i^* = \mu_i^\bullet + \sum_{j=1}^{i-1} \left( \sum_{p \in P[i \to j]} W(p) \right) \cdot \left( \mu_j^\bullet - \mathbb{E}[X_j | X_j \leq c_j] \right).$$

As a corollary, suppose $L \geq 3$ and $\boldsymbol{\gamma}$ have the Rabin and Vayanos (2010) functional form of $\gamma_{i,j} = -\alpha \cdot \delta^{i-j-1}$ for $\alpha > 0$, $0 \leq \delta \leq 1$. I show that all pseudo-true fundamentals are too pessimistic in every dataset censored with $\boldsymbol{c}\uparrow = (c_1, ..., c_L) \in \mathbb{R}^L$ if and only if $\delta > \alpha$. The idea is the influence of the gambler's fallacy psychology must not decay "too quickly" relative to the influence of the most recent observation. This condition is satisfied in all the calibration exercises in Rabin and Vayanos (2010) and in the structural estimations of Benjamin, Moore, and Rabin (2017).

**Corollary A.2.** *Suppose $L \geq 3$ and $\gamma_{i,j} = -\alpha \cdot \delta^{i-j-1}$ for $\alpha > 0$, $0 \leq \delta \leq 1$. If $\delta > \alpha$, then for all stopping strategies $\boldsymbol{c}\uparrow = (c_1, ..., c_L) \in \mathbb{R}^L$, the pseudo-true fundamentals satisfy*

$\mu_i^* < \mu_i^\bullet$ *for all* $i$*. If* $\delta < \alpha$*, then there exists a stopping strategy* $\boldsymbol{c} \uparrow= (c_1, ..., c_L) \in \mathbb{R}^L$ *such that* $\mu_i^* > \mu_i^\bullet$ *for at least one* $i$*.*

To understand the intuition, consider an example that violates the condition of the corollary, $\alpha = 0.5$, $\delta = 0$, so that $\gamma_{2,1} = -0.5$, $\gamma_{3,2} = -0.5$, and $\gamma_{3,1} = 0$. The agent expects reversals between the pairs $(X_1, X_2)$ and $(X_2, X_3)$, but his expectation for $X_3|(X_1 = x_1, X_2 = x_2)$ does not vary with $x_1$. By the same logic as the two-periods censoring effect, inference about the second-period fundamental $\mu_2^*$ decreases as $c_1$ decreases, with $\lim_{c_1 \to -\infty} \mu_2^*(c_1) = -\infty$. This has an important indirect effect on $\mu_3^*$, since a very pessimistic $\mu_2^*$ leads the agent to interpret objectively typical draws of $X_2$ as greatly above average. Expecting low values of $X_3$ after these surprisingly high draws of $X_2$, the agent infers the fundamental $\mu_3^*$ to be above the sample mean of $X_3$ in the dataset, hence overestimating it as $c_1 \to -\infty$. When $\delta$ is strictly positive, however, there is an opposite effect where lower sample mean of $X_1$ in observations containing uncensored $X_3$ lead to more pessimistic inference about the third-period fundamental. When $\delta > 0.5$, overoptimistic inference never happens because this second effect dominates.

# D The Censoring Effect in a Finite-Urn Model

Rabin (2002) Section 7 discusses an example with endogenous observations. There is an infinite population of financial analysts, each with quality $\theta \in \{\frac{1}{4}, \frac{1}{2}, \frac{3}{4}\}$. Conditional on quality $\theta$, an analyst generates either a good (signal $a$) or bad (signal $b$) return each period, with probabilities $\theta$ and $1-\theta$ and independently across periods. The agent, however, believes successive returns from the same analyst are generated through a finite-urn model. Consider an urn with $N$ balls where $N$ is a multiple of 4. For an analyst with quality $\theta$, initialize the urn with $\theta N$ balls labeled "$a$" and $(1-\theta)N$ balls labeled "$b$". Successive returns are successive draws without replacement from the urn. The urn is refreshed every two draws. Rabin (2002) calls an agent with this finite-urn model an "$N$-Freddy". Since the urn is not refreshed between draws $2k-1$ and $2k$ for $k = 1, 2, 3, ...$, such pairs of draw exhibit negative correlation in agent's subjective model, generating the gambler's fallacy.

Returning to Rabin (2002) Section 7's example, objectively all financial analysts have quality $\theta = \frac{1}{2}$. The agent samples a financial analyst at random and observes his returns over two periods. Depending on the realizations of these two returns, the agent either observes the same analyst for two more periods before sampling a new analyst, or immediately samples a new analyst. This procedure is infinitely repeated. Rabin (2002) investigates a 4-Freddy agent's long-run belief about the proportions of analysts with the three levels of quality in the population.

The endogenous observation in the example is distinct from what I term the "censoring effect" in this paper. The main mechanism behind the censoring effect is that the some rows of the dataset omits signals ($X_2$) which the biased agent judges to be negatively correlated with signals that are always observed ($X_1$). This then leads to distorted inference. However, in Rabin (2002)'s finite-urn model, the urn is refreshed every two periods. This means an $N$-Freddy agent judges the part of the data that is sometimes censored (the analyst's returns in periods 3 and 4) to be independent of the part of the data that is always observed (the analyst's returns in periods 1 and 2). Therefore the driving force behind Rabin (2002) Section 7's example is not the interaction between censoring and the gambler's fallacy, but rather between censoring and the Bayesian aspect of $N$-Freddy's quasi-Bayesian inference.

In this section, I study a related problem where an $N$-Freddy agent observes each analyst for either one or two periods, depending on whether the analyst generates a bad first-period return. This setup features the censoring effect, because the finite-urn model generates negative correlation between the first and second draws from each urn. I find that the agent's inference under this censoring structure tends to be too optimistic. This conclusion is in line with predictions about the censoring effect in the baseline model of this paper, for the basic inference result in Proposition 1 shows that when the dataset is censored in the opposite way (i.e. censored when the first draw is good), the resulting inference is too pessimistic[18]. That is, I demonstrate the robustness of my censoring effect to an alternative model of the gambler's fallacy in a binary-signals setting, showing that it is not an artifact of the continuous-signals setup in my baseline model.

Table A.1 displays the likelihood of all signals of length 2 for the 4-Freddy and 8-Freddy agents, for different values of $\theta \in \{\frac{1}{4}, \frac{1}{2}, \frac{3}{4}\}$. The last row of each table also shows the likelihoods of simply observing the signal $b$ in the first period, under the censoring rule that stops observing an analyst if his first return is bad.

I first discuss inference without censoring. After $aa$, Freddy exaggerates the relative likelihood of $\theta = \frac{3}{4}$ to $\theta = \frac{1}{2}$ compared to a Bayesian, whereas after $ab$ Freddy's relative likelihoods of these two qualities are the same as a Bayesian's. Overall, given a sample with an equal number of $aa$ and $ab$ signals, Freddy exaggerates the relative likelihood of $\theta = \frac{3}{4}$ to $\theta = \frac{1}{2}$. This phenomenon is analogous to the continuous version of gambler's fallacy where a biased observer "partially forgives" a mediocre outcome following an outstanding outcome. Here, even though the average outcome in the second period is mediocre, the fact that they follow the best possible outcome $a$ in the first period lead to an overly optimistic estimate about the analyst's ability. By the same logic, observing an equal number of $ba$ and $bb$ signals would lead to exaggeration of the likelihood of $\theta = \frac{1}{4}$ relative to $\theta = \frac{1}{2}$.

[18]Proposition OA.5 in the Online Appendix shows that when the dataset is censoring using a strategy that stops when $X_1 \le c$ for some $c \in \mathbb{R}$, inference about second-period fundamental is always too high.

| 4-Freddy | $\theta = \frac{1}{4}$ | $\theta = \frac{1}{2}$ | $\theta = \frac{3}{4}$ | 8-Freddy | $\theta = \frac{1}{4}$ | $\theta = \frac{1}{2}$ | $\theta = \frac{3}{4}$ |
|---|---|---|---|---|---|---|---|
| $aa$ | 0 | $\frac{1}{6}$ | $\frac{1}{2}$ | $aa$ | $\frac{1}{28}$ | $\frac{6}{28}$ | $\frac{15}{28}$ |
| $ab$ | $\frac{1}{4}$ | $\frac{1}{3}$ | $\frac{1}{4}$ | $ab$ | $\frac{6}{28}$ | $\frac{8}{28}$ | $\frac{6}{28}$ |
| $ba$ | $\frac{1}{4}$ | $\frac{1}{3}$ | $\frac{1}{4}$ | $ba$ | $\frac{6}{28}$ | $\frac{8}{28}$ | $\frac{6}{28}$ |
| $bb$ | $\frac{1}{2}$ | $\frac{1}{6}$ | 0 | $bb$ | $\frac{15}{28}$ | $\frac{6}{28}$ | $\frac{1}{28}$ |
| $b\varnothing$ | $\frac{3}{4}$ | $\frac{1}{2}$ | $\frac{1}{4}$ | $b\varnothing$ | $\frac{3}{4}$ | $\frac{1}{2}$ | $\frac{1}{4}$ |

Table A.1: The likelihoods of observations under different analyst qualities, for 4-Freddy and 8-Freddy agents.

However, now suppose the second observation is censored when the first observation is $b$. The otherwise symmetric situation becomes asymmetric. Following the observation of $b\varnothing$ (where the second draw is censored), Freddy's inference is the same as a Bayesian's. So we have turned off the channel that exaggerates the probability of $\theta = \frac{1}{4}$ but kept the channel that exaggerates the probability of $\theta = \frac{3}{4}$. This is analogous to the censoring effect in my model, where censoring second period draw following unfavorable first period draws would lead to overly optimistic beliefs.

In the long-run, the agent observes a distribution of returns across different analysts: 25% of the time $aa$ is observed, 25% of the time $ab$ is observed, and 50% of the time $b\varnothing$ is observed. To calculate the agent's long-run beliefs, first suppose Freddy's prior specifies either all analysts have $\theta = \frac{1}{4}$ or all analysts have $\theta = \frac{3}{4}$. Then Freddy's long-run inference is given by the parameter maximizing expected log-likelihood of the data. For 4-Freddy, the log-likelihood likelihood under $\theta = \frac{1}{4}$ is $-\infty$. For 8-Freddy, The log-likelihood under $\theta = \frac{1}{4}$ is

$$\frac{1}{4}\ln(1/28) + \frac{1}{4}\ln(6/28) + \frac{1}{2}\ln(3/4) \approx -1.362$$

and the log-likelihood under $\theta = \frac{3}{4}$ is

$$\frac{1}{4}\ln(\frac{15}{28}) + \frac{1}{4}\ln(\frac{6}{28}) + \frac{1}{2}\ln(1/4) \approx -1.234.$$

So in both cases, Freddy will come to believe $\theta = \frac{3}{4}$ over $\theta = \frac{1}{4}$ for all analysts.

Now consider a 4-Freddy who dogmatically believes some $1 - \kappa \in (0, 1)$ fraction of the analysts have $\theta = \frac{1}{2}$, but the remaining analysts either have $\theta = \frac{1}{4}$ or $\theta = \frac{3}{4}$. So, the agent estimates $q_a \in [0, 1 - \kappa]$, the fraction of analysts who have $\theta = \frac{3}{4}$. Straightforward algebra shows that the $q_a^*$ maximizing expected log-likelihood of the data is $q_a^* = \frac{7}{18}\kappa + \frac{1}{9}$ for $\kappa \geq \frac{2}{11}$, $q_a^* = \kappa$ otherwise. Since $\frac{7}{18}\kappa + \frac{1}{9} > \frac{1}{2}\kappa$ for all $\kappa \in (\frac{2}{11}, 1)$, we see that no matter what fraction of analysts 4-Freddy believes to be average, he will end up believing there are more above-average than below-average analysts in the population. That is, his overall belief will be too optimistic.

# E The Gambler's Fallacy and Attentional Stability

Many studies of learning with behavioral agents, including this paper, can be phrased as agents with a prior (or "misspecified theory") over states of the world whose support excludes the true, data-generating state. Agents in my baseline model, for example, start with a prior supported on the class of subjective models $\{\Xi(\mu_1, \mu_2, \sigma^2, \sigma^2; \gamma) : \mu_1, \mu_2 \in \mathbb{R}\}$ for some $\gamma < 0$, with different models viewed as different states of the world. But the true state of the world is the objective distribution $(X_1, X_2) \sim \Xi(\mu_1^\bullet, \mu_2^\bullet, \sigma^2, \sigma^2; 0)$, which does not belong to the previous set. As the agent's data size grows, the misspecified theory can appear infinitely less likely in the limit than an alternative prior belief (or "light-bulb theory") that includes the true state in its support.

Gagnon-Bartsch, Rabin, and Schwartzstein (2018) offer an explanation for why such misspecified theories persist with learning – attentional stability. Under a misspecified theory, some coarsened information may be sufficient for decision-making. When agents only pay attention to this coarsened information, the part of the data that they attend to may be so coarse that their misspecified theory no longer appears infinitely less likely than the light-bulb theory.

In this section, I investigate attentional stability of the gambler's fallacy bias in my learning setting. The main intuition is that when agents are dogmatic about $\gamma$, they are dogmatic about the negative correlation between $X_1$ and $X_2$. As such, under their misspecified theory agents do not find it necessary to separately keep track of the conditional distributions $X_2|(X_1 = x_1)$ for different values of $x_1$. Agents believe certain moments of the dataset are sufficient for decision-making, and this process of compressing the entire dataset into these sufficient statistics removes aspects of the dataset that would have led the agents to question the validity of their theory.

My setting differs slightly from the setting of Gagnon-Bartsch, Rabin, and Schwartzstein (2018) as each of my agents acts once after observing an infinitely large dataset, while their agents observe one signal each period over an infinite number of periods. So, I begin by defining the key concepts surrounding attentional stability in my setting.

## E.1 A Definition of Attentional Stability for Large Datasets

Recall that each agent in generation $t \geq 1$ observes an infinite dataset of histories $(h_n)_{n\in[0,1]}$ from the previous generation, where each $h_n \in H = \mathbb{R} \times (\mathbb{R} \cup \{\varnothing\})$ is the history of a predecessor. Since there is a continuum of predecessors, I will think of each agent as directly observing a distribution on $H$, i.e. directly observing $\mathcal{H}(\Xi^{\bullet}; c\uparrow)$ when their predecessors use the stopping rule $c\uparrow$.

**Definition A.5.** Let $\pi, \lambda$ be beliefs over the joint distribution of $(X_1, X_2)$. Say $\pi$ is **inexplicable** relative to $\lambda$, conditional on the objective model $\Xi^{\bullet}$ and cutoff rule $c\uparrow$, if $\mathcal{H}(\Xi^{\bullet}; c\uparrow) = \mathcal{H}(\Xi; c\uparrow)$ for some $\Xi \in \text{supp}(\lambda)$, but $\mathcal{H}(\Xi^{\bullet}; c\uparrow) \neq \mathcal{H}(\Xi; c\uparrow)$ for any $\Xi \in \text{supp}(\lambda)$.

That is, each subjective model $\Xi$ and cutoff rule $c\uparrow$ induces a distribution over histories, $\mathcal{H}(\Xi; c\uparrow)$. When predecessors in generation $t-1$ use the stopping rule $c\uparrow$, agents in generation $t$ observe an infinite dataset of histories with the distribution $\mathcal{H}(\Xi^{\bullet}; c\uparrow)$. If this distribution can be explained by some subjective model of $(X_1, X_2)$ in the support of the light-bulb theory $\lambda$, but not by any distribution in the support of the misspecified theory $\pi$, then I call $\pi$ inexplicable.

I now define a particular kind of limited attention. Given an infinite dataset (i.e. a distribution over histories), the agent maps the entire distribution to a finite number of real numbers, an extreme form of data coarsening. If there is a strategy optimal under the misspecified theory $\pi$ that only makes use of these finitely many statistics, then we have a sufficient-statistics strategy.

**Definition A.6.** A **sufficient-statistics strategy (SSS) in large datasets** consists of a statistics map $S : \Delta(H) \to \mathbb{R}^K$ for some finite $K < \infty$ and a cutoff map $\sigma : \text{Im}(S) \times \mathbb{R} \to \mathbb{R}$, such that agents in every generation $t \geq 1$ find it optimal (under the prior $\Xi \sim \pi$) to use the stopping strategy with cutoff $\sigma(S(\mathcal{H}), c_{[t-1]})$ when they observe a dataset $(h_n)_{n\in[0,1]}$ from their predecessors censored using their stopping threshold $c_{[t-1]}$.

An agent following the strategy $(S, \sigma)$ first extracts $K$ statistics (i.e. real numbers) from the infinite dataset $(h_n)_{n\in[0,1]}$ of predecessor histories. Then, she applies $\sigma$ to choose a cutoff strategy that only depends on the dataset $(h_n)_{n\in[0,1]}$ through its $K$ extracted statistics $S((h_n)_{n\in[0,1]})$. The idea is that the agent only pays attention to the finitely many statistics, which is perhaps more realistic than paying full attention to the entire infinite dataset containing the histories from a continuum of agents. If such a strategy is optimal for an agent believing the true joint distribution of $(X_1, X_2)$ is drawn according to her (misspecified) prior $\Xi \sim \pi$, I call the pair $(S, \sigma)$ an SSS.

A related definition of sufficiency works with finite datasets instead of infinite datasets. While I will mostly work with these large-dataset concepts, the finite-dataset analog of a large-dataset SSS I look at in the next subsection is optimal in the way I formalize here.

**Definition A.7.** A **sufficient-statistics strategy (SSS) in datasets of size** $N < \infty$ consists of a statistics map $S_N : H^N \to \mathbb{R}^K$ for some finite $K < \infty$ and a cutoff map

$\sigma_N : \text{Im}(S_N) \times \mathbb{R} \times \mathbb{N} \to \mathbb{R}$, such that the subjectively optimal cutoff threshold (under the prior $\pi$ over $\Xi$, given by a prior density $g : \mathbb{R}^2 \to \mathbb{R}_{++}$ over $(\mu_1, \mu_2)$) is $\sigma_N(S_N((h_n)_{n=1}^N), c, N_1)$ after observing a dataset $(h_n)_{n=1}^N$ censored using cutoff strategy $c \uparrow$, where $N_1 \le N$ is the number of histories with an uncensored second-period draw.

Finally, I combine these concepts to define attentional stability. Roughly speaking, the theory $\pi$ is attentionally stable if we can find an $(S, \sigma)$ pair that pays "fine" enough attention to be an SSS under $\pi$, but "coarse" enough attention so that the resulting statistics can be explained by some model in the support of $\pi$.

**Definition A.8.** Theory $\pi$ is **attentionally stable**, conditional on the objective model $\Xi^\bullet$ and cutoff rule $c \uparrow$, if there exists an SSS $(S, \sigma)$ such that $S(\mathcal{H}(\Xi^\bullet; c \uparrow)) = S(\mathcal{H}(\Xi; c \uparrow))$ for some $\Xi$ in the support of $\pi$.

## E.2 Two Sufficient-Statistics Strategy for Gambler's Fallacy

Fix $\gamma < 0$. Let $\pi$ be any full-support belief over $\{\Xi(\mu_1, \mu_2, \sigma^2, \sigma^2; \gamma) : \mu_1, \mu_2 \in \mathbb{R}\}$. In the baseline model, I consider agents with a prior density $g : \mathbb{R}^2 \to \mathbb{R}_{++}$ over $(\mu_1, \mu_2)$ that is everywhere strictly positive, which clearly induces such a $\pi$. Let $\lambda$ be any belief with $\Xi^\bullet = \Xi(\mu_1^\bullet, \mu_2^\bullet, \sigma^2, \sigma^2; 0)$ in its support. I first show without channeled attention, agents will come to realize that their misspecified theory $\pi$ is wrong after seeing a large dataset.

*Claim* A.1. $\pi$ is inexplicable relative to $\lambda$ conditional on $\Xi^\bullet$ and any stopping rule $c \uparrow$.

*Proof.* This is because $\Xi^\bullet \in \text{supp}(\lambda)$ but every $\Xi \in \text{supp}(\pi)$ has KL divergence bounded away from 0 relative to $\Xi^\bullet$ in terms of the histories they generate under $c \uparrow$ censoring, that is to say

$$\begin{aligned}
&\min_{\Xi \in \text{supp}(\pi)} D_{KL}(\mathcal{H}(\Xi^\bullet; c \uparrow) || \mathcal{H}(\Xi; c \uparrow)) \\
&= \min_{\mu_1, \mu_2 \in \mathbb{R}} D_{KL}(\mathcal{H}(\Xi^\bullet; c \uparrow) || \mathcal{H}(\Xi(\mu_1, \mu_2, \sigma^2, \sigma^2; \gamma); c \uparrow)) > 0.
\end{aligned}$$

Next, I exhibit two different SSS that establish the attentional stability of the gambler's fallacy psychology in my setting. Both SSS have the additional property that they lead agents to believe the pseudo-true fundamentals and hence use the same cutoff strategy in large datasets as the full-attention Bayesian agents in the baseline model. So, not only do these SSS provide a justification for agents not discarding their misspecified theory after seeing large datasets, they also justify the learning dynamics that I investigate in the main text of the paper. ☐

### E.2.1 First SSS: Sample Means of Each Period

The first SSS relates to the method-of-moments interpretation of the pseudo-true fundamentals from Remark 4. Suppose $K = 2$ and let $S(\mathcal{H})$ return the sample mean of $X_1$ and the uncensored sample mean of $X_2$ in the history distribution $\mathcal{H}$. Let $\sigma(s_1, s_2, c) = C(s_1, s_2 + \gamma(s_1 - \mathbb{E}[\tilde{X}_1|\tilde{X}_1 \leq c]))$, where $\tilde{X}_1 \sim \mathcal{N}(s_1, \sigma^2)$ and $C(\mu_1, \mu_2)$ is the indifference threshold with dogmatic belief in $\Xi(\mu_1, \mu_2; \gamma)$. To see that this is an SSS, note that for any $c \in \mathbb{R}$, $\Xi \mapsto S(\mathcal{H}(\Xi; c))$ is a one-to-one function on the support of $\pi$. The unique $\mu_1, \mu_2$ generating the two moments $s_1, s_2$ are $\mu_1 = s_1$, $\mu_2 = s_2 + \gamma(s_1 - \mathbb{E}[\tilde{X}_1|\tilde{X}_1 \leq c])$. So an agent who thinks only models in the support of $\pi$ are feasible will believe she exactly identifies the data-generating model using just the two statistics given by $S(\mathcal{H})$, and subsequently use the subjectively optimal cutoff strategy $C(s_1, s_2 + \gamma(s_1 - \mathbb{E}[\tilde{X}_1|\tilde{X}_1 \leq c]))$.

This SSS makes $\pi$ attentionally explicable conditional on the objective model $\Xi^\bullet$ and stopping strategy $c\uparrow$, because

$$S(\mathcal{H}(\Xi^\bullet; c\uparrow)) = S(\mathcal{H}(\Xi(\mu_1^\bullet, \mu_2^*(c)); c\uparrow)),$$

where $\mu_2^*(c) \in \mathbb{R}$ is the pseudo-true fundamental with censoring $c\uparrow$, by the method-of-moments interpretation of pseudo-true fundamentals.

### E.2.2 Second SSS: Sample Mean of Re-Centered Second-Period Draws

The second SSS has an even stronger sufficiency property, in the sense that its finite-sample analog is SSS in finite datasets.

Again let $K = 2$. In a dataset of size $N$, let

$$S_N((h_n)_{n=1}^N) = \left( \frac{1}{N} \sum_{n=1}^N h_{1,n},\ \frac{1}{\#(n : h_{2,n} \neq \varnothing)} \sum_{n: h_{2,n} \neq \varnothing} (h_{2,n} - \gamma h_{1,n}) \right).$$

The first statistic is the sample mean of the first-period draws. The interpretation of the second statistic is that the agent forms the "re-centered" observation $w_n := h_{2,n} - \gamma h_{1,n}$ for each history $h_n$ where $h_{2,n} \neq \varnothing$. The agent only pays attention to the sample averages of $x_{1,n} = h_{1,n}$ and $w_n$. Under the subjective model $\Xi(\mu_1, \mu_2; \gamma)$, we may write the distributions of $X_1, X_2$ as

$$\begin{aligned} X_1 &= \mu_1 + \epsilon_1 \\ X_2 &= \mu_2 + \gamma\epsilon_1 + z_2 \end{aligned}$$

where $\epsilon_1, z_2 \sim \mathcal{N}(0, \sigma^2)$, are independent. Defining $W := X_2 - \gamma X_1$, we see that under

$\Xi(\mu_1, \mu_2; \gamma)$, $W = \mu_2 - \gamma\mu_1 + z_2$. So, observations of first-period draws are signals about $\mu_1$, while observations of re-centered second-period $W$ are signals about $\mu_2 - \gamma\mu_1$.

*Claim* A.2. $S_N$ is part of an SSS in datasets of size $N$.

Roughly speaking ,this is because the subjective joint distribution between $(X_1, W)$ is Gaussian and the mean of a sequence of Gaussian random variables is a sufficient statistic for the likelihood of the entire sequence.

Consider now large-sample analog of $S_N$. Again with $K = 2$, the statistic map $S$ sends each distribution $\mathcal{H}$ to $\mathbb{E}_{h\sim\mathcal{H}}[h_{i,1}]$ and $\mathbb{E}_{h\sim\mathcal{H}}[h_{i,2} - \gamma h_{i,1} | h_{i,2} \neq \varnothing]$. I first show that $S$ makes $\pi$ attentionally explicable.

*Claim* A.3. For any censoring threshold $c \in \mathbb{R}$, $S(\mathcal{H}(\Xi^\bullet; c\uparrow)) = S(\mathcal{H}(\Xi(\mu_1^\bullet, \mu_2^*(c)); c\uparrow))$, with $S_2(\mathcal{H}(\Xi^\bullet; c\uparrow)) = \mu_2^\bullet - \gamma\mathbb{E}[X_1 | X_1 \leq c]$.

Under the theory $\Xi = \Xi(\mu_1, \mu_2; \gamma)$, agents expect the second statistic to be

$$\begin{aligned}
\mathbb{E}_\Xi[X_2 - \gamma X_1 | X_1 \leq c] &= \mathbb{E}_\Xi[X_2 | X_1 \leq c] - \gamma\mathbb{E}_\Xi[X_1 | X_1 \leq c] \\
&= \mathbb{E}_\Xi[\mu_2 + \gamma(X_1 - \mu_1) | X_1 \leq c] - \gamma\mathbb{E}_\Xi[X_1 | X_1 \leq c] \\
&= \mu_2 - \gamma\mu_1.
\end{aligned}$$

It is clear then $\Xi \mapsto S(\mathcal{H}(\Xi; c))$ is a one-to-one function on the support of $\pi$, and that we may put $\sigma(s_1, s_2) = C(s_1, s_2 + \gamma s_1)$ to make $(S, \sigma)$ an SSS. By Claim A.3, $S_2(\mathcal{H}(\Xi^\bullet; c\uparrow)) = \mu_2^\bullet - \gamma\mathbb{E}[X_1 | X_1 \leq c]$, so this means the agents will play

$$C(\mu_1^\bullet, \mu_2^\bullet - \gamma\mathbb{E}[X_1 | X_1 \leq c] + \gamma\mu_1^\bullet) = C(\mu_1^\bullet, \mu_2^\bullet + \gamma(\mu_1^\bullet - \mathbb{E}[X_1 | X_1 \leq c])) = C(\mu_1^\bullet, \mu_2^*(c))$$

when faced with the dataset $\mathcal{H}^\bullet(c)$, same as full-attention Bayesian agents in the baseline model.