Mislearning from Censored Data:  
The Gambler’s Fallacy in Optimal-Stopping Problems  

Kevin He*  

First version: March 21, 2018  
This version: August 15, 2019  

Abstract  

I study endogenous learning dynamics for people expecting systematic reversals from random sequences — the “gambler’s fallacy.” Biased agents face an optimal-stopping problem, such as managers conducting sequential interviews. They are uncertain about the underlying distribution (e.g. talent distribution in the labor pool) and learn its parameters from their predecessors. Agents stop when early draws are deemed “good enough,” so predecessors’ experience contain negative streaks but not positive streaks. Since biased agents understate the likelihood of consecutive below-average draws, society converges to over-pessimistic beliefs about the distribution’s mean. When early agents decrease their acceptance thresholds due to pessimism, later agents will become more surprised by the lack of positive reversals in their predecessors’ histories, leading to more pessimistic inferences and lower acceptance thresholds — a positive-feedback cycle. Agents who are additionally uncertain about the distribution’s variance believe in fictitious variation (exaggerated variance) to an extent depending on the severity of data censoring.

*California Institute of Technology and University of Pennsylvania. Email: hesichao@gmail.com. I am indebted to Drew Fudenberg, Matthew Rabin, Tomasz Strzalecki, and Ben Golub for their guidance and support. I thank Isaiah Andrews, Ruiqing Cao, In-Koo Cho, Martin Cripps, Krishna Dasaratha, Jetlir Duraj, Ben Enke, Ignacio Espona, Jiacheng Feng, Mira Frick, Tristan Gagnon-Bartsch, Ashvin Gandhi, Oliver Hart, Johannes Hörner, Alice Hsiaw, Ryota Iijima, Yuhta Ishii, Lawrence Jin, Yizhou Jin, Michihiro Kandori, Max Kasy, Shengwu Li, Jonathan Libgober, Matthew Lilley, George Mailath, Eric Maskin, Weicheng Min, Xiaosheng Mu, Andy Newman, Harry Pei, Joshua Schwartzstein, Roberto Serrano, Philipp Strack, Elie Tamer, Omer Tamuz, Michael Thaler, Linh T. Tô, Maria Voronina, Yuichi Yamamoto, and my seminar participants for their insightful comments.


1 Introduction

The gambler’s fallacy is widespread. Many people believe that 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 birth to consecutive daughters, and, in general, expect too much reversal from sequential realizations of independent random events. Studies have documented this bias in settings where it is strictly costly, such as state lotteries with pari-mutuel payouts (Terrell, 1994; Suetens, Galbo-Jørgensen, and Tyran, 2016) and incentivized lab experiments (Benjamin, Moore, and Rabin, 2017). The same bias also affects experienced decision-makers in high-stakes environments, including immigration judges (Chen, Moskowitz, and Shue, 2016). Section 1.3 surveys more of this empirical literature.

This paper highlights novel implications of the gambler’s fallacy in optimal-stopping problems when a society of biased agents learns about the underlying distributions. As a running example, consider a junior HR manager who sequentially interviews candidates for a single job opening. In deciding whether to hire a candidate or to keep searching, the junior manager must form a belief about the distribution of potential future applicants should she keep the position open. She consults with senior managers and adopts their belief about the labor pool based on their recruiting experience for similar positions in the past. The junior manager then implements a stopping strategy for her own recruiting problem, updates her belief at the end of the hiring season, and shares this new belief with future managers. Suppose all managers commit the gambler’s fallacy — that is, they exaggerate how unlikely it is to get consecutive above-average or consecutive below-average applicants (relative to the labor pool mean). This error stems from the same psychology that leads people to exaggerate how unlikely it is to get consecutive heads or consecutive tails when tossing a fair coin. How does this bias influence the managers’ beliefs and behavior over time?

In this example and other natural optimal-stopping problems, agents tend to stop when early draws are deemed “good enough,” leading to an asymmetric truncation of experience. When a manager discovers a sufficiently strong candidate early in the hiring cycle, she stops her recruitment efforts and does not observe what additional candidates would have been found for the same job opening with a longer search. This endogenous censoring effect on histories interacts with the gambler’s fallacy bias and leads to pessimistic inference about the labor pool. Managers continue searching only when their early candidates are below-average. They misinterpret subsequent above-average candidates as the expected positive reversal after bad initial outcomes, not as strong signals about the labor pool. On the other hand, they are surprised by subsequent below-average candidates since they understate the likelihood of bad streaks, misreading consecutive bad draws as very strong negative signals about the pool. That is, after bad early draws, managers under-infer from subsequent good draws but over-infer from subsequent bad draws. On average, they communicate an over-pessimistic impression of the labor pool to today’s junior manager. This pessimism informs
the junior manager’s stopping strategy and affects the kind of censored history she observes and the new belief she communicates to future managers.

This paper examines the endogenous learning dynamics of a society of agents believing in the gambler’s fallacy. All agents face a common stage game: an optimal-stopping problem with draws in different periods independently generated from fixed yet unknown distributions. They take turns playing the stage game, with each agent’s payoff determined by the game’s outcome. Agents are Bayesians except for the statistical bias. That is, they start with a prior belief supported on a class of feasible models about the joint distribution of draws. Feasible models are symmetric, log-concave distributions indexed by different unconditional means (the fundamentals). I study the gambler’s fallacy as a misspecified prior: all feasible models specify that better earlier draws tend to lead to worse later draws, and vice versa. The feasible models exclude the true distribution where draws are independent, so agents undertake misspecified Bayesian learning.

I consider two social-learning environments. In the first environment, agents play the stage game one at a time. Before playing her own game, each agent adopts the final belief of her immediate predecessor as her prior belief and formulates a stopping strategy. At the end of her game, she updates her belief about the fundamentals by applying the Bayes’ rule to her stage-game history, then passes on her posterior belief to her successor. I show that the stochastic processes of the agents’ beliefs and behavior almost surely converge to a unique steady state in which agents are over-pessimistic about the fundamentals and stop too early relative to the objectively optimal strategy.

In the second environment, agents arrive in large generations with everyone in the same generation playing simultaneously after observing all predecessors’ histories. Society converges to the same steady state as the previous environment. This large-generations model illustrates a positive-feedback cycle between distorted beliefs and distorted stopping strategies. More severely censored datasets lead to more pessimistic beliefs, while more pessimistic beliefs lead to earlier stopping and, as a consequence, heavier history censoring. Mapping back to the recruiting example, suppose a firm appoints HR managers in cohorts. Upon arrival, each junior manager learns the recruiting experience of all previous managers. If managers in the first cohort start with the correct stopping strategy, then average hiring outcome monotonically deteriorates across all future cohorts. After today’s cohort observes predecessors’ histories and makes an over-pessimistic inference, this belief leads them to act less “choosy” and only keep searching if their early candidates prove to be truly un-

---

1 Mueller, Spinnewijn, and Topa (2018) find evidence consistent with people exhibiting the gambler’s fallacy in an optimal-stopping problem. They show job seekers’ beliefs about the probability of finding a job in the near future increase significantly over the course of the unemployment spell, after controlling for individual fixed effects. These beliefs contrast with theories that predict decreasing job-finding rates (e.g., human capital depreciation) and with the authors’ structural estimation that suggests constant rates. One application of my work is studying how a society of such biased job seekers make inferences about job-finding rates from others’ job-search experience.
satisfactory. On average, early applicants rejected by today’s managers are worse than the early applicants rejected by previous cohorts. Since biased learners expect greater positive reversal following worse early outcomes, they expect an improved distribution of later candidates compared to what happened in previous generations. When this improvement fails to materialize, biased managers become disappointed and view it as a negative signal about the average labor-pool quality. The pessimistic inference of today’s cohort thus becomes amplified in the next generation, leading to a further lowering of acceptance thresholds and a further decrease in the average quality of the hired candidate.

The endogenous-data setting leads to novel comparative statics predictions about how the economic environment affects learning outcomes under the gambler’s fallacy. For instance, suppose a new government policy subsidizes longer search in the optimal-stopping problem, so that agents incur a lower continuation cost when they decide to keep searching. This policy does not change long-run learning outcomes if data is exogenous or if agents are correctly specified. When biased agents learn from endogenously censored histories, however, lower continuation costs in the stage game moderate steady-state pessimism about the fundamentals. This result is another expression of the positive feedback between actions and beliefs. Under the new policy, agents use higher acceptance thresholds and generate less censored histories, which lead to less pessimistic beliefs for their successors. An increase in acceptance thresholds today (due to the subsidy) gets compounded with further increases in the future (due to more optimistic beliefs). In the long run, the government subsidy — a policy intervention implementable without knowing the true distribution — partially corrects society’s beliefs.

Finally, I specialize to the Gaussian case, expand the set of feasible models, and consider agents who are uncertain about both the means and variances of the distributions. In this joint estimation, agents misinfer means by the same amount as in the baseline model, and exaggerate variances. In the recruiting example, biased managers wrongly believe that applicants for different vacancies come from different labor pools that vary in average quality, when in reality applicants for all vacancies originate from the same pool with a fixed quality distribution. This belief in vacancy-specific fixed effects helps biased managers better explain histories containing consecutive below-average candidates, reasoning that it must have been especially difficult to find good applicants for these specific job openings. The degree of belief in this fictitious variation both depends on severity of history censoring and influences the managers’ stopping strategy. I derive two results that illustrate how this belief in fictitious variation interacts with endogenous learning. First, when the stage-game payoff function is convex in draws (such as when previously rejected candidates can be recalled with some probability in the sequential interviewing game), the positive-feedback cycle of the baseline environment strengthens. More severely censored histories not only make agents more pessimistic about the fundamentals by the usual censoring effect, but also decrease their belief in fictitious variation. Both forces encourage earlier stopping due to
the convexity of the optimal-stopping problem, so subsequent agents will face even heavier data censoring. Second, a society where agents are uncertain about the variances can end up with a different long-run belief about the means than another society where agents know the correct variances. This is despite the fact that agents in both societies would make the same (mis)inference about the means given the same data.

I study a number of extensions in the Online Appendix, showing robustness of the results to a range of alternative specifications. The paper focuses on (misspecified) Bayesian agents, but the over-pessimism result and the positive-feedback loop continue to obtain under a non-Bayesian method-of-moments inference procedure (Online Appendix OA 5). For simplicity I consider a two-period optimal-stopping problem as the stage game, but the combination of the gambler’s fallacy and history truncation after good outcomes still produces over-pessimistic inferences in stage games of arbitrary length (Online Appendix OA 2). I assume all agents have the gambler’s fallacy. The presence of a subpopulation of unbiased agents or agents suffering from additional behavioral biases may mitigate the extent of over-pessimism (Online Appendix OA 8.2), but does not eliminate it.

1.1 Contributions

This work contributes to two strands of literature: the behavioral economics literature on inference mistakes for biased learners, and the theoretical literature on the dynamics of misspecified endogenous learning.

As a contribution to behavioral economics, I highlight a novel channel of misinference for behavioral agents — the interaction between psychological bias and data censoring. In many natural environments, agents learn from censored data. The economics literature has recently focused on the learning implications of selection neglect in these settings, where agents act as if their dataset is not censored. This work points out that other well-documented behavioral biases can also interact with data censoring to produce new implications. Mislearning stems precisely from this interaction, not from either censored data or the gambler’s fallacy alone. Agents who do not suffer from the statistical bias learn the fundamentals correctly even from censored histories. On the other hand, if we removed censoring by having agents observe ex-post what would have been drawn in each period of the optimal-stopping problem, then even biased agents would learn the fundamentals correctly. The intuition is that the gambler’s fallacy is a “symmetric” bias. The “asymmetric” outcome of over-pessimism only occurs when the bias interacts with an (endogenous) asymmetric censoring mechanism that tends to produce data containing negative streaks but not positive streaks. Environments that feature different censoring patterns (e.g., strategies that produce positive streaks) or other behavioral biases would produce different predictions, but again through the same basic mechanism — interaction between censoring and bias.

2See, for example, Enke (2019) and Jehiel (2018).
As a theoretical contribution, I prove convergence of beliefs and behavior in a non-self-confirming misspecified setting with a continuum of states of the world. Economists study many kinds of misspecifications with the property that even the “best-fitting” feasible belief does not match data exactly — that is to say, no feasible self-confirming belief exists. The gambler’s fallacy belongs to this family, since all feasible models imply a negative correlation absent in the data. This paper analyzes the stochastic processes of belief and behavior under this statistical bias, proving their global almost-sure convergence to a unique steady state. In related work, Heidhues, Koszegi, and Strack (2018) study learning dynamics under overconfidence about own ability. Despite being biased, agents in their setting always have some feasible belief that exactly rationalizes data, so their learning steady-state is a self-confirming equilibrium. Another difference is that I establish my convergence result in a setting with multiple dimensions of uncertainty (the distributional parameters for different periods of the stage game), whereas Heidhues, Koszegi, and Strack (2018) consider convergence of misspecified learning with one-dimensional uncertainty. Fudenberg, Romanyuk, and Strack (2017) study a continuous-time model of active learning under misspecification, but their learning problem only involves two feasible models. In this work, agents’ prior belief about each distributional parameter is supported on a continuum of possible values.

As another contribution to the theoretical literature on misspecified learning dynamics, this project studies a new source of endogeneity: the censoring effect in a dynamic stage game. The dynamic stage game is both essential for studying learning under the gambler’s fallacy — a behavioral bias concerning the serial correlation of data — and crucial for the censoring effect. In my setting, the type of data that an agent generates depends on her beliefs. To understand the distinction from the existing literature, consider the classic paper in this area, Nyarko (1991), who studies a monopolist setting a price on each date and observing the resulting sales. No matter what action the monopolist takes, she observes the same type of data: quantity sold. Similarly, the agent in Fudenberg, Romanyuk, and Strack (2017) always observes payoffs and the agent in Heidhues, Koszegi, and Strack (2018) always observes output levels, after any action. Endogenous learning in these other papers takes the form of agents attributing different meanings to the same data, when interpreted through the lenses of different actions. On the other hand, we may think of stage-game histories censored with different thresholds as different types of data that, by themselves, lead to different beliefs about the fundamentals for biased learners. Actions play no role in inference except to generate these different types of data, as the likelihood of a (feasible) history does not depend on the censoring threshold that produced it.

\[^{3}\text{Their follow-up work Heidhues, Koszegi, and Strack (2019) also focuses on mislearning with one-dimensional uncertainty in a self-confirming setting.}\]
1.2 Other Related Theoretical Work

Rabin (2002) and Rabin and Vayanos (2010) are the first to study the inferential mistakes implied by the gambler’s fallacy. Except for an example in Rabin (2002), discussed below, all such investigations focus on passive inference, whereby learners observe an exogenous information process. By contrast, this paper examines an endogenous learning setting where actions affect observables. This setting allows me to study how actions and beliefs co-evolve and whether the feedback cycle between them attenuates or exaggerates distortions over the course of learning. Another distinction is that the present paper focuses on the dynamics of mislearning under censoring and works with the stochastic processes of beliefs and behavior.

Section 7 of Rabin (2002) discusses an example of endogenous learning with a finite-urn model of the gambler’s fallacy. The nature of Rabin (2002)’s endogenous data, however, is unrelated to the censoring effect central to the current paper. In Online Appendix OA 6, I modify Rabin’s example to induce the censoring effect. His finite-urn model then delivers a misinference result analogous to the results in this paper, which are derived in a different setting with continuously-valued draws. This exercise shows the robustness of my results within different modeling frameworks of the same statistical bias.

Steady state in this work corresponds to Esponda and Pouzo (2016)’s Berk-Nash equilibrium. Rather than focusing only on equilibrium analysis, however, I study non-equilibrium learning dynamics and prove global convergence of behavior. This paper also contains more specific results: I emphasize the interaction between censoring and bias as the driver of mislearning, discuss how changing the stage game affects long-run beliefs, and relate my results to previous findings on inference under the gambler’s fallacy (e.g., fictitious variation in an endogenous-data setting).

Although my learning framework involves a sequence of short-lived agents, the social-learning aspect of the framework is not central to the results. In fact, the environment where a sequence of short-lived agents acts one at a time is equivalent to an environment where a single long-lived agent plays the stage game repeatedly, myopically maximizing her expected payoff in each iteration of the stage game. In the growing literature on social learning with misspecified Bayesians (e.g., Eyster and Rabin (2010); Gaurino and Jehiel (2013); Bohren (2016); Bohren and Hauser (2018); Frick, Iijima, and Ishii (2019)), agents observe their predecessors’ actions but make errors when inverting these actions to deduce said predecessors’ information. This kind of action inversion does not take place here: later

---

4In Rabin (2002)’s example, biased agents (correctly) believe that the part of the data which is always observable is independent of the part of the data which is sometimes missing. However, what I term the “censoring effect” is about misinference resulting from agents wrongly believing in negative correlation between the early draws that are always observed and the later draws that may be censored, depending on the realizations of the early draws.

5Esponda, Pouzo, and Yamamoto (2019)’s work-in-progress considers misspecified learning environments with finite action sets and studies the convergence of empirical action frequencies. Their techniques and notion of convergence do not seem to apply to a setting with a continuum of actions.
agents inherit all the information that their predecessors have seen, either by adopting their beliefs or by observing their histories, so predecessors’ actions are uninformative.

The econometrics literature has also studied data-generating processes with censoring — for example, the Tobit model and models of competing risks. This literature has primarily focused on the issue of model identification from censored data (Cox, 1962; Tsiatis, 1975; Heckman and Honoré, 1989). In my setting, there is no identification problem for correctly specified agents. Instead, I study how agents make wrong parameter estimates from censored data when they infer using a family of misspecified models. Another contrast is that the econometrics literature has focused on exogenous data-censoring mechanisms, but censoring is endogenous in this paper and depends on the beliefs of previous agents. This endogeneity is central to the results, as discussed before.

1.3 Empirical Evidence on the Gambler’s Fallacy

Bar-Hillel and Wagenaar (1991) review classical psychology studies on the gambler’s fallacy. The earliest lab evidence involves two types of tasks. In “production tasks,” subjects are asked to write down sequences using a given alphabet, with the goal of generating sequences that resemble the realizations of an i.i.d. random process. Subjects tend to produce sequences with too many alternations between symbols, as they attempt to locally balance out symbol frequencies. In “judgment tasks” where people are asked to identify which sequence of binary symbols appears most like consecutive tosses of a fair coin, subjects routinely judge sequences with an alternation rate of 60% as “more random” than those with an alternation rate of 50%. While most of these studies are unincentivized, Benjamin, Moore, and Rabin (2017) have found the gambler’s fallacy with strict monetary incentives, where a bet on a fair coin continuing its streak pays strictly more than the bet on the streak reversing. Barron and Leider (2010) have shown 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.

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 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

References can be found in Amemiya (1985) and Crowder (2001).
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. Finally, using field data from asylum judges, loan officers, and baseball umpires, 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 the gambler’s fallacy, as the fallacy leads to the belief that the objectively “correct” decision is negatively auto-correlated across a sequence of decision problems. The authors rule out several 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 will attribute the lack of reversals in the data to the fundamental probabilities of the randomizing device. One implication is that the gambler’s fallacy generates the hot-hand fallacy: in settings where it is plausible that the fundamental probabilities fluctuate, biased agents may come to believe in such fluctuations even if they do not exist. For instance, a sports fan who under-estimates the likelihood of making four free throws in a row with a 70% accuracy rate reasons that anyone on a four-hits streak must be in a “hot” state with a temporarily elevated accuracy rate far above 70%. This misinference leads the fan to over-estimate the probability that the said player will also make the next free throw. The same kind of 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 (Narayan 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. Similarly, a number that has been drawn more often in the past six weeks, excluding the most recent past week, gets more bets in the Denmark lottery (Suetens, Galbo-Jørgensen, and Tyran, 2016). This kind of overinference 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 studies about misinferences under a primitive gambler’s fallacy bias, I consider agents who believe in reversals \textit{conditional} on the underlying fundamentals and mislearn some parameters of the world as a result. But, the misinference mechanism in this paper is further complicated by the presence of endogenous data censoring.

2 Model and Overview of Results

This section presents the basic elements of the model, previews the main results, and provides intuition for how the censoring effect drives the conclusions. I describe the (single-player) \textit{stage game}, an optimal-stopping problems satisfying some conditions. Agents are uncertain about the distribution of draws in the stage game. They entertain a prior belief over a family of \textit{feasible models} of how draws are generated. All feasible models specify the same negative correlation between draws, though they are objectively independent — an error that reflects the gambler’s fallacy. Sections 3 and 4 embed these model elements into social-learning
environments and derive learning dynamics. Section 5 contains a number of extensions that verify robustness of the main results.

2.1 Basic Elements of the Model

2.1.1 Optimal-Stopping Problem as a Dynamic Stage Game

The stage game is a two-period optimal-stopping problem. In the first period, the agent draws \( x_1 \in \mathbb{R} \) and decides whether to stop. If she stops, her payoff is \( u_1(x_1) \) and the stage game ends. Otherwise, she continues to the second period and draws \( x_2 \in \mathbb{R} \). The stage game then ends with the payoff \( u_2(x_1, x_2) \).

The payoff functions \( u_1 : \mathbb{R} \to \mathbb{R} \) and \( u_2 : \mathbb{R}^2 \to \mathbb{R} \) satisfy some regularity conditions to be introduced in Assumption 1. The following example satisfies Assumption 1 and will be used to illustrate my results throughout this paper.

Example 1 (search with \( q \) probability of recall). Many industries have an annual hiring cycle. Consider a firm in such an industry and an HR manager who must fill a job opening during this year’s cycle. In the early phase of the hiring cycle, she finds a candidate who would bring net benefit \( x_1 \) to the organization if hired. She must decide between hiring this candidate immediately or waiting. Waiting means she continues searching in the late phase of the cycle, finding another candidate with benefit \( x_2 \). Waiting carries the risk that the early candidate accepts an offer from a different firm in the interim, which happens with probability \( 0 < 1 - q \leq 1 \). This situation has the payoff functions \( u_1(x_1) = x_1 \) and \( u_2(x_1, x_2) = q \cdot \max(x_1, x_2) + (1 - q)x_2 \). In the late phase, there is \( q \) probability the manager gets payoff equal to the higher of the two candidates’ qualities, and complementary probability that only the second candidate is available.

The following regularity conditions define the class of optimal-stopping problems I study.

Assumption 1 (regularity conditions). 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 exist \( x'_1, x''_1, x'_2, x''_2 \in \mathbb{R} \) so that \( u_1(x'_1) > u_2(x''_1, x'_2) \) and \( u_1(x''_1) < u_2(x'_1, x''_2) \).

(d) \( u_1, u_2 \) are continuous and \( x_2 \mapsto u_2(\bar{x}_1, x_2 + \bar{k}) \) is absolutely integrable with respect to the objective distribution of \( X_2 \) for all \( \bar{x}_1, \bar{k} \in \mathbb{R} \).

Assumption 1(a) says \( u_1, u_2 \) are strictly increasing in the draws of their respective periods. Assumption 1(b) says a higher realization of the early draw increases first-period payoff more than it changes second-period payoff. Under Assumption 1(a), Assumption 1(b) is satisfied whenever \( u_2 \) is not a function of \( x_1 \), as in optimal-stopping problems where stopping in period
$k$ gives payoff only depending on the $k$-th draw. Assumption 1(c) says there exist situations where the agent wants to stop and other situations where the agent wants to continue. The technical Assumption 1(d) ensures continuation payoffs are well-defined. These conditions are satisfied by my recurring example.

**Claim 1.** Example 1 satisfies Assumption 1 whenever the objective distribution of $X_2$ has a finite first moment.

Proofs of results in Sections 2 to 4 can be found in Appendix A.

I now define strategies and histories of the stage game.

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

Without loss I only consider pure strategies, because there always exists a payoff-maximizing pure strategy under any belief about the distribution of draws.

**Definition 2.** The history of the stage game is an element $h \in H := \mathbb{R} \times (\mathbb{R} \cup \{\emptyset\})$. If an agent decides to stop after $X_1 = x_1$, her history is $(x_1, \emptyset)$. 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 $\emptyset$ is a censoring indicator, emphasizing that the hypothetical second-period draw is unobserved when the agent does not continue into the second period. In Example 1, if the HR manager hires the first candidate, she stops her recruitment efforts early and the counterfactual second candidate that she would have found had she kept the position open remains unknown.

### 2.1.2 Feasible Models and the Objective Model

I work with a general class of distributions for the main results. Both the true data-generating process and the agents’ domain of learning can be described in terms of a pair of densities on $\mathbb{R}$ satisfying the following:

**Assumption 2** (log-concavity and symmetry). $f_1(\cdot \mid 0)$ and $f_2(\cdot \mid 0)$ are strictly positive densities on $\mathbb{R}$ with finite second moments, and they are strictly log-concave, symmetric, and mean-zero.

A leading example of strictly log-concave and symmetric distributions is the Gaussian distribution. Another example is the logistic distribution. The mean-zero condition is only a normalization, since we can shift any log-concave distribution symmetric around its mean to be centered around 0.

For $\tau_1, \tau_2 \in \mathbb{R}$, let $f_1(\cdot \mid \tau_1)$ and $f_2(\cdot \mid \tau_2)$ represent shifted versions of $f_1(\cdot \mid 0)$ and $f_2(\cdot \mid 0)$ centered around $\tau_1$ and $\tau_2$, respectively. More precisely, $f_1(x_1 \mid \tau_1) := f_1(x_1 - \tau_1 \mid 0)$ and $f_2(x_2 \mid \tau_2) := f_2(x_2 - \tau_2 \mid 0)$ for $x_1, x_2 \in \mathbb{R}$. 

10
Objectively, draws $X_1, X_2$ in the stage game are independently distributed with $X_1 \sim f_1(\cdot \mid \mu_1^*)$ and $X_2 \sim f_2(\cdot \mid \mu_2^*)$. The parameters $\mu_1^*, \mu_2^* \in \mathbb{R}$ are the true fundamentals. In Example 1, $\mu_1^*$ and $\mu_2^*$ stand for the true qualities of the two applicant pools in the early and late phases of the hiring season.

Agents are uncertain about the distribution of $(X_1, X_2)$. The next definition describes the set of distributions that a gambler’s fallacy agent deems plausible.

**Definition 3.** The set of feasible models $\{\Psi(\mu_1, \mu_2; \gamma) : (\mu_1, \mu_2) \in \mathcal{M}\}$ is a family of joint distributions of $(X_1, X_2)$ indexed by feasible fundamentals $(\mu_1, \mu_2) \in \mathcal{M} \subseteq \mathbb{R}^2$, for some bias parameter $\gamma > 0$. Here $\Psi(\mu_1, \mu_2; \gamma)$ refers to the joint distribution

$$X_1 \sim f_1(\cdot \mid \mu_1),$$

$$(X_2 \mid X_1 = x_1) \sim f_2(\cdot \mid \mu_2 - \gamma(x_1 - \mu_1)),$$

where $X_2 \mid (X_1 = x_1)$ is the conditional distribution of $X_2$ given $X_1 = x_1$.

I write $\mathbb{E}_\Psi$ and $\mathbb{P}_\Psi$ throughout for expectation and probability with respect to model $\Psi$. When $\mathbb{E}$ and $\mathbb{P}$ are used without subscripts, they refer to expectation and probability under the true model, $\Psi^* = \Psi(\mu_1^*, \mu_2^*; 0)$.

I model the gambler’s fallacy as an additive shift in the agent’s belief about $X_2$’s distribution following different $X_1$ realizations, so that $(X_2 \mid X_1 = x_1)$ increases in first-order stochastic dominance order as $x_1$ decreases. Conditional on the fundamentals, if the realization of $X_1$ is higher than expected, then the agent believes bad luck is due in the near future and the second draw is likely below average. Conversely, an exceptionally bad early draw likely portends above-average luck in the next period. This interpretation is clearer in the following equivalent formulation of $\Psi(\mu_1, \mu_2; \gamma)$: $X_1 = \mu_1 + \epsilon_1, X_2 = \mu_2 + \epsilon_2$ where $\epsilon_1 \sim f_1(\cdot \mid 0)$ and $(\epsilon_2 \mid \epsilon_1) \sim f_2(\cdot \mid -\gamma \epsilon_1)$. The mean-zero terms $\epsilon_1, \epsilon_2$ represent the idiosyncratic factors, or “luck,” that determine how $X_1$ and $X_2$’s realizations deviate from their unconditional means $\mu_1$ and $\mu_2$. The negative correlation between $\epsilon_1$ and $\epsilon_2$ conditional on $\mu_1, \mu_2$ represents a belief in reversal of luck. Larger $\gamma > 0$ implies greater magnitudes in these expected reversals and thus more bias.

It is useful to keep the Gaussian case in mind, which I will use to derive closed-form versions of some general results. I will also specialize to the Gaussian case for some of the extensions in Section 5.

---

1. Study gambler’s fallacy for continuous random variables, where the magnitude of $X_1$ affects the agent’s prediction about $X_2$. Chen, Moskowitz, and Shue (2016)’s analysis of baseball umpire data provides support for the continuous version of the statistical bias. They find that an umpire is more likely to call the current pitch a ball after having called the previous pitch a strike, controlling for the actual location of the pitch. Crucially, the effect size is larger after more obvious strikes, where “obviousness” is based on the distance of the pitch to the center of the regulated strike zone. This distance can be thought of as a continuous measure of the “quality” of each pitch.
Example 2 (the Gaussian case). Objectively, $X_1 \sim \mathcal{N}(\mu_1, \sigma^2)$ and $X_2 \sim \mathcal{N}(\mu_2, \sigma^2)$ are independent Gaussian random variables each with variance $\sigma^2 > 0$. But the agent believes $X_1, X_2$ are a pair of correlated Gaussian random variables with $X_1 \sim \mathcal{N}(\mu_1, \sigma^2)$ and $(X_2 \mid X_1 = x_1) \sim \mathcal{N}(\mu_2 - \gamma(x_1 - \mu_1), \sigma^2)$ for some $(\mu_1, \mu_2) \in \mathcal{M}$.

The set of feasible models is indexed by the set of feasible fundamentals, $\mathcal{M}$. We may think of the agents as learning about the unconditional means of $X_1$ and $X_2$, with $\mathcal{M}$ as the domain of their inference.

Remark 1. I will consider several specifications of $\mathcal{M}$ throughout this paper.

(a) $\mathcal{M} = \mathbb{R}^2$. The agent thinks all values $(\mu_1, \mu_2) \in \mathbb{R}^2$ are possible.

(b) $\mathcal{M} = \diamond$, where $\diamond$ is a bounded parallelogram in $\mathbb{R}^2$ whose left and right edges are parallel to the $y$-axis, whose top and bottom edges have slope $-\gamma$. The agent is uncertain about both $\mu_1$ and $\mu_2$, but her uncertainty has bounded support.\(^8\)

(c) $\mathcal{M} = \{\mu_1\} \times [\mu_2, \bar{\mu}_2]$. The agent has a correct, dogmatic belief about $\mu_1$, but has uncertainty about $\mu_2$ supported on a bounded interval.

(d) $\mathcal{M} = \{(\mu, \mu) : \mu \in \mathbb{R}\}$. The agent is convinced that the first-period and second-period fundamentals are the same, but is uncertain what this common parameter is.

While the agent can freely update her belief about the fundamentals on $\mathcal{M}$, she holds a dogmatic belief about $\gamma > 0$.\(^9\) This implies the set of feasible models excludes the true model, $\Psi^* = \Psi(\mu_1^*, \mu_2^*; 0)$, so Bayesian updating within the class of feasible models amounts to misspecified learning. I use misspecification as a tool to represent and study the gambler’s fallacy. This approach is motivated by field evidence on the bias’ persistence: for example, Chen, Moskowitz, and Shue (2016) show that even very experienced decision-makers exhibit a non-negligible amount of the gambler’s fallacy in high-stakes settings.

In the social-learning environment I study in Section 3, short-lived agents each observes one iteration of the stage game, so no one has a large enough dataset to identify the misspecification problem. In Online Appendix OA 7, I discuss why even agents with large datasets may never question their feasible models: the misspecification is “attentionally stable” in the sense of Gagnon-Bartsch, Rabin, and Schwartzstein (2018).

Before stating my main results, I first establish a proposition about the optimal stage-game strategy. This will motivate a slight strengthening of Assumption 1 that I need for some results. For $c \in \mathbb{R}$, write $S_c$ for the cutoff strategy such that $S_c(x_1) =$ Stop if and only if $x_1 > c$. That is, $S_c$ accepts all early draws above a cutoff threshold $c$.

---

\(^8\)Any prior belief over fundamentals $(\mu_1, \mu_2)$ supported on a bounded set in $\mathbb{R}^2$ can be arbitrarily well-approximated by a prior belief over a large enough $\diamond$.

\(^9\)Section 5.3 studies the extension where agents are uncertain about $\gamma$, but the support of their prior belief about $\gamma$ lies to the left of 0 and is bounded away from it.
Proposition 1. Under Assumption 1 and for $\gamma > 0$,

- Under each feasible model $\Psi(\mu_1, \mu_2; \gamma)$, there exists a cutoff threshold $C(\mu_1, \mu_2; \gamma) \in \mathbb{R}$ such that it is strictly optimal to continue whenever $x_1 < C(\mu_1, \mu_2; \gamma)$ and strictly optimal to stop whenever $x_1 > C(\mu_1, \mu_2; \gamma)$.

- For every $\mu_1 \in \mathbb{R}$, $\mu_2 \mapsto C(\mu_1, \mu_2; \gamma)$ is strictly increasing.

- For every $\mu_1 \in \mathbb{R}$, $\mu_2 \mapsto C(\mu_1, \mu_2; \gamma)$ is Lipschitz continuous with Lipschitz constant $1/\gamma$.

The content of this proposition is threefold.

First, it shows that the best strategy for the class of optimal-stopping problems I study takes a cutoff form. This is because a higher $x_1$ both increases the payoff to stopping and, under the gambler’s fallacy, predicts worse draws in the next period. Both forces push in the direction of stopping. The optimality of cutoff strategies leads to an endogenous, asymmetric censoring of histories, formalizing the idea that agents stop after “good enough” draws.

Second, holding fixed $\mu_1$, the cutoff threshold increases with $\mu_2$. This is because the agent can afford to be choosier in the first period when prospects in the second period improve.

The third statement about Lipschitz continuity, on the other hand, gives a bound on how quickly $\mu_2 \mapsto C(\mu_1, \mu_2; \gamma)$ increases. Suppose that one agent believes draws are generated according to $\Psi(\mu_1, \mu_2; \gamma)$, while another agent believes they are generated according to $\Psi(\mu_1, \mu_2 + 1; \gamma)$. If the first agent is indifferent between stopping and continuing after $X_1 = c$, then the second agent prefers stopping after $X_1 = c + 1/\gamma$. This is because the predicted conditional mean of $X_2$ falls by $(1/\gamma) \cdot \gamma = 1$ when $X_1$ increases by $1/\gamma$ under any feasible model, which cancels out the relative optimism of the second agent about the unconditional distribution of $X_2$.

The Lipschitz constant $1/\gamma$ is guaranteed for every optimal-stopping problem satisfying Assumption 1 and every $\gamma > 0$. But, $1/\gamma$ may not be the best Lipschitz constant. My results use the slightly stronger condition that $\mu_2 \mapsto C(\mu_1^*, \mu_2; \gamma)$ has a Lipschitz constant strictly smaller than $1/\gamma$. Instead of making an assumption on $C$ directly, I strengthen Assumption 1(b) on the stage-game primitives to imply the desired infinitesimally stronger Lipschitz continuity.

Assumption 3 ($\ell$-Lipschitz continuity).

Either: (a) There exists $0 < \ell < \frac{1}{\gamma}$ so that for every $x_1, x_2 \in \mathbb{R}$ and $d > 0$,

$$u_1(x_1 + \ell d) - u_1(x_1) \geq u_2(x_1 + \ell d, x_2 + (1 - \gamma \ell) d) - u_2(x_1, x_2)$$

Or: (b) $u_2$ is Lipschitz continuous and only a function of $x_2$, and furthermore there exists $\epsilon > 0$ so that $u'_1(x_1) > \epsilon$ for all $x_1 \in \mathbb{R}$. 

13
Assumption 3(a) is satisfied by my recurring example.

Claim 2. Example 1 satisfies Assumption 3(a) with \( \ell = \frac{1}{1+\gamma} \) for every probability of recall \( 0 \leq q < 1 \) and every bias \( \gamma > 0 \).

### 2.2 Main Results

I now state my two main results, which concern learning dynamics under the gambler’s fallacy in two different social-learning environments. Precise details of these environments will follow in Sections 3 and 4, respectively.

In the first environment, short-lived agents arrive one per round, \( t = 1, 2, 3, \ldots \). Agent in round \( t = 1 \) starts with a full-support prior density \( m_0 : \diamond \rightarrow \mathbb{R}_{>0} \), where \( \diamond \) is a bounded parallelogram in \( \mathbb{R}^2 \) as in Remark 1(b). In round \( t \), agent \( t \) adopts the final belief \( \tilde{m}_{t-1} \) of her immediate predecessor as her prior belief, then chooses a cutoff threshold \( \tilde{C}_t \) to maximize her expected payoff based on this belief. She observes what happens in her stage game and uses Bayes’ rule to update her belief from \( \tilde{m}_{t-1} \) to \( \tilde{m}_t \), which then becomes the prior belief of agent \( t + 1 \).

In this environment, the sequences of cutoffs \( (\tilde{C}_t) \) and posterior belief densities \( (\tilde{m}_t) \) are stochastic processes whose randomness derives from the randomness of draws. Draws are objectively independent, both between the two periods in the same round of the stage game and across different rounds. Write \( (\tilde{\mu}_{1,t}, \tilde{\mu}_{2,t}) \) for the random element in \( \diamond \) given by the density \( \tilde{m}_t \).

**Theorem 1.** Suppose Assumptions 1, 2, and 3 hold, and the second derivative of \( \ln (f_2(x | \mu^\bullet_2)) \) is uniformly bounded for \( x \in \mathbb{R} \). There exists a unique steady state \( \mu^\infty_2, c^\infty \in \mathbb{R} \) not dependent on \( m_0 \), so that provided \( (\mu^\bullet_1, \mu^\infty_2) \in \diamond \), almost surely \( \lim_{t \to \infty} \tilde{C}_t = c^\infty \) and \( (\tilde{\mu}_{1,t}, \tilde{\mu}_{2,t})_{t \geq 1} \) converges in \( L^1 \) to \( (\mu^\bullet_1, \mu^\infty_2) \). The steady state satisfies \( \mu^\infty_2 < \mu^\bullet_2 \) and \( c^\infty < c^\bullet \), where \( c^\bullet \) is the objectively optimal cutoff threshold.

In other words, almost surely behavior and belief converge in the society, and this steady state is independent of the prior over fundamentals (provided its support is large enough). In the steady state, agents hold overly pessimistic beliefs about the fundamentals and stop too early, relative to the objectively optimal strategy. (The additional regularity assumption that the second derivative of \( \ln (f_2(x | \mu^\bullet_2)) \) is uniformly bounded is satisfied by the Gaussian and logistic distributions.)

In the second environment, short-lived agents arrive in generations, \( t = 0, 1, 2, \ldots \), with a continuum of agents per generation. Agents’ prior belief about the fundamentals is given by a full-support density \( m_0 \) on \( \mathbb{R}^2 \), as in Remark 1(a). Each agent observes the stage-game...

\(^{10}\)I focus on learning across different iterations of the stage game and assume agents do not update beliefs within the stage game.
histories of all predecessors from all past generations to make inferences about the funda-
mentals. Due to the large generations, cutoffs and beliefs are deterministic in generations
t ≥ 1, which I denote as c[t] and μ[t] = (μ1[t], μ2[t]) respectively. The society is initialized at
an arbitrary cutoff strategy Sc[0] in the 0th generation, the initial condition.

**Theorem 2.** Suppose Assumptions 1 and 2 hold. Starting from any initial condition and any
m0, cutoffs (c[t])t≥1 and beliefs (μ2,[t])t≥1 form monotonic sequences across generations. When
Assumption 3 also holds, there exists a unique steady state μ∞,c∞ ∈ R so that c[t] → c∞
and (μ1,[t], μ2,[t]) → (μ1∗, μ2∗) monotonically, regardless of the initial condition and m0. This
steady state is the same as the one in Theorem 1.

The monotonicity of beliefs and cutoffs across generations reflects the positive feedback
between changes in beliefs and changes in behavior. Suppose generation t is more pessimistic
than generation t − 1 about the second-period fundamental, μ2,[t] < μ2,[t−1]. The monotonicity
result implies that beliefs move in the same direction again in generation t + 1, that is
μ2,[t+1] < μ2,[t]. The information of generation t + 1 differs from that of generation t only in
that agents in generation t + 1 observe all stage-game histories of generation t. This means
generation t’s stopping behavior differs from that of generation t − 1 in such a way as to
generate histories that amplify, not dampen, the initial change in beliefs from generation
t − 1 to generation t.

### 2.3 Intuition for the Main Results

In the learning environments of this paper, each agent censors her stage game history through
her stopping strategy, where the strategy choice depends on her beliefs. To build intuition
for how this censoring effect relates to the two main theorems, I first consider a biased agent
with feasible fundamentals M = R2, facing a large sample of histories all censored according
to some cutoff threshold c. I characterize her inference about fundamentals when the sample
size grows and analyze how her inference depends on the cutoff threshold c.

Suppose c ∈ R ∪ {∞} and Ψ is a model. Then H(Ψ; c) refers to the distribution of histories
when draws are generated by Ψ and histories censored according to S_c.

**Definition 4.** For c ∈ R and Ψ a model, H(Ψ; c) ∈ Δ(Δ) is the distribution of histories
given by

\[
H(Ψ; c)[E_1 \times E_2] := \mathbb{P}_Ψ[(E_1 \cap (c, \infty)) \times E_2] \quad \text{for } E_1, E_2 \in \mathcal{B}(R)
\]

\[
H(Ψ; c)[E_1 \times \emptyset] := \mathbb{P}_Ψ[(E_1 \cap (-\infty, c]) \times \mathbb{R}] \quad \text{for } E_1 \in \mathcal{B}(R),
\]

where \( \mathcal{B}(R) \) is the collection of Borel subsets of \( R \).

I abbreviate \( H(Ψ^*; c) \) as simply \( H^*(c) \), the true distribution of histories under the cutoff
threshold c. The next definition gives a measure of the difference between the distribution
of histories under the feasible model with fundamentals \((\mu_1, \mu_2)\) and the true distribution of histories, both with the same censoring threshold \(c\).

**Definition 5.** For \(\mu_1, \mu_2 \in \mathbb{R}, c \in \mathbb{R} \cup \{\infty\}\) the Kullback-Leibler (KL) divergence from \(\mathcal{H}^*(c)\) to \(\mathcal{H}(\Psi(\mu_1, \mu_2; \gamma); c)\), denoted by \(D_{KL}(\mathcal{H}^*(c) \mid \mid \mathcal{H}(\Psi(\mu_1, \mu_2; \gamma); c) )\), is

\[
\begin{align*}
&\int_c^\infty f_1(x_1 \mid \mu_1^*) \cdot \ln \left( \frac{f_1(x_1 \mid \mu_1^*)}{f_1(x_1 \mid \mu_1)} \right) dx_1 \\
&\quad + \int_{-\infty}^c \left\{ \int_{-\infty}^\infty f_1(x_1 \mid \mu_1^*) \cdot f_2(x_2 \mid \mu_2^*) \cdot \ln \left[ \frac{f_1(x_1 \mid \mu_1^*) \cdot f_2(x_2 \mid \mu_2^*)}{f_1(x_1 \mid \mu_1) \cdot f_2(x_2 \mid \mu_2 - \gamma(x_1 - \mu_1))} \right] dx_2 \right\} dx_1.
\end{align*}
\]

The minimizers of KL divergence with respect to cutoff \(c\),

\[(\mu_1^*, \mu_2^*) \in \arg \min_{\mu_1, \mu_2 \in \mathbb{R}} D_{KL}(\mathcal{H}^*(c) \mid \mid \mathcal{H}(\Psi(\mu_1, \mu_2; \gamma); c) ),\]

are called the pseudo-true fundamentals with respect to \(c\).

To interpret, the likelihood of the history \(h = (x_1, x_2)\) with \(x_1 \leq c\) is \(f_1(x_1 \mid \mu_1^*) \cdot f_2(x_2 \mid \mu_2^*)\) under the true model \(\Psi^*\), \(f_1(x_1 \mid \mu_1) \cdot f_2(x_2 \mid \mu_2 - \gamma(x_1 - \mu_1))\) under the feasible model \(\Psi(\mu_1, \mu_2; \gamma)\). The likelihood of the history \(h = (x_1, \infty)\) with \(x_1 > c\) is \(f_1(x_1 \mid \mu_1^*)\) under the true model, \(f_1(x_1 \mid \mu_1)\) under the feasible model. The likelihoods of all other histories are 0 under both models. So the KL divergence expression in Definition 5 is the expected log-likelihood ratio of the history under the true model versus under the feasible model with fundamentals \((\mu_1, \mu_2)\), where expectation over histories is taken under the true model. In general, this optimization objective depends on the cutoff threshold \(c\) and I denote the pseudo-true fundamentals as \(\mu_1^*(c), \mu_2^*(c)\) to emphasize this dependence. The pseudo-true fundamentals correspond to the biased agent’s inference about the fundamentals in large samples.

The next proposition characterizes the pseudo-true fundamentals and contains the key intuition behind the two main theorems.

**Proposition 2.** Under Assumption 2, for any \(c \in \mathbb{R} \cup \{\infty\}\), the KL divergence minimization problem in Definition 5 admit a unique solution \((\mu_1^*(c), \mu_2^*(c)) \in \mathbb{R}^2\). Furthermore:

- \(\mu_1^*(c) = \mu_1^*\) for any \(c \in \mathbb{R} \cup \{\infty\}\)
- \(\mu_2^*(c) < \mu_2^*\) for any \(c \in \mathbb{R}\) and \(\mu_2^*(\infty) = \mu_2^*\)
- \(\mu_2^*(c)\) is strictly increasing in \(c\).

In the Gaussian case, the pseudo-true fundamental \(\mu_2^*(c)\) admits a closed-form expression that readily verifies Proposition 2.
Example 2 (continued). In the Gaussian case, for $c \in \mathbb{R} \cup \{\infty\}$,

$$
\mu^*_2(c) = \mu^*_2 - \gamma (\mu^*_1 - \mathbb{E}[X_1 | X_1 \leq c]).
$$

The censoring effect is crucial for misinference: as Proposition 2 shows, the pseudo-true fundamentals are unbiased in the absence of censoring (i.e., when $c = \infty$). Here is why the directional data censoring I study leads to over-pessimism. In every feasible model of draws $\Psi(\mu_1, \mu_2; \gamma)$, the realization of $X_2$ depends on two factors: the second-period fundamental $\mu_2$, and a reversal effect based on the realization of $X_1$. The biased agent cannot end up with a correct or over-optimistic belief about $\mu_2$, else she would be systematically disappointed by realizations of $X_2$ in her dataset. This is because $X_2$ is only uncensored when $X_1$ is low enough, a contingency where the agent expects positive reversal on average. Over-pessimism can therefore be thought of as “two wrongs making a right,” as the biased agent’s pessimism about the unconditional mean of $X_2$ counteracts her false expectation of positive reversals in the dataset of censored histories.

This mechanism explains the long-run pessimism in Theorem 1 and Theorem 2. In fact, in the large-generations setting of Theorem 2, every generation $t \geq 1$ holds strictly pessimistic beliefs, so over-pessimism is also a short-run phenomenon provided there are enough predecessors per generation. The idea that asymmetric data censoring after favorable draws combined with the gambler’s fallacy leads to pessimistic inference is robust. In the Gaussian case, it continues to hold when agents are uncertain about variances (Section 5.2), when the feasible fundamentals reflect agents’ knowledge that $\mu_1 = \mu_2$ as in Remark 1(d) (Section 5.4), when the stage game has more than two periods (Online Appendix OA 2), under an alternative method-of-moments inference procedure (Online Appendix OA 5), when a fraction of agents suffer from selection neglect (Online Appendix OA 8.2), when higher draws bring worse payoffs (Online Appendix OA 8.1), and with high probability after observing a finite dataset containing just 100 censored histories (Online Appendix OA 9.1).

The severity of the biased agent’s pessimism increases with the severity of censoring. The intuition is that the agent wants to infer a lower $\mu^*_2$ to better match $X_2$’s in histories that start with bad $X_1$’s, but doing so carries the cost of a worse model fit for histories that start with intermediate $X_1$’s. More severe censoring — generated by a strategy that accepts not only very good $X_1$’s but also intermediate ones — alleviates this cost, as histories that start with intermediate $X_1$’s no longer contain their associated $X_2$’s. The optimal inference $\mu^*_2$ thus decreases.

The comparative static $\frac{d\mu^*_2}{dc} > 0$ is central to Theorem 2’s positive-feedback loop. In the large-generations model, Generation 1 observes a large dataset of histories drawn from $\mathcal{H}^*(c_{[0]})$ and chooses a cutoff $c_{[1]}$. Generation 2 then observes histories from all predecessor generations, that is histories drawn from both $\mathcal{H}^*(c_{[0]})$ and $\mathcal{H}^*(c_{[1]})$. If $c_{[1]} < c_{[0]}$, then Generation 2’s dataset features (on average) more severe censoring than Generation 1’s
dataset. Thus, Generation 2 comes to a more pessimistic inference about the second-period fundamental. By Proposition 1, this leads to a further lowering of the cutoff threshold, $c_2 < c_1$, and the pattern continues.

3 Convergence, Over-Pessimism, and Early Stopping

This section studies a social-learning environment where biased agents act one at a time and pass down beliefs to their successors. I define the steady state of the stage game, prove its existence and uniqueness, and show it features over-pessimistic beliefs and early stopping. Then, I turn to the stochastic process of beliefs and behavior in the social-learning environment, showing that this process almost surely converges to the steady state.

3.1 Steady State: Existence, Uniqueness, and Other Properties

A steady state is a triplet consisting of fundamentals $(\mu_1^n, \mu_2^n) \in \mathbb{R}^2$ and a cutoff threshold $c^n \in \mathbb{R}$ that endogenously determine each other. The cutoff strategy with acceptance threshold $c^n$ maximizes expected payoff under the feasible model $\Psi(\mu_1^n, \mu_2^n; \gamma)$, while the fundamentals are the pseudo-true fundamentals under data censoring with threshold $c^n$.

More precisely,

Definition 6. A steady state consists of $\mu_1^n, \mu_2^n, c^n \in \mathbb{R}$ such that:

1. $c^n = C(\mu_1^n, \mu_2^n; \gamma)$.
2. $\mu_1^n = \mu_1^*(c^n)$ and $\mu_2^n = \mu_2^*(c^n)$.

Steady states correspond to Esponda and Pouzo (2016)'s pure Berk-Nash equilibria for an agent whose prior is supported on the feasible models with feasible fundamentals $\mathcal{M} = \mathbb{R}^2$, under the restriction that equilibrium belief puts full confidence in a single fundamental pair. The set of steady states depends on $\gamma$, since the severity of the bias changes both the optimal cutoff thresholds under different fundamentals and inference about fundamentals from stage-game histories.

The “steady state” defined here almost surely characterizes the long-run learning outcome in the society where biased agents act one by one. This convergence does not follow from Esponda and Pouzo (2016), for their results only imply local convergence from prior beliefs sufficiently close to the equilibrium beliefs, and only in a “perturbed game” environment where learners receive idiosyncratic payoff shocks to different actions. I will show global convergence of the stochastic processes of beliefs and behavior without payoff shocks.

Like almost all examples of Berk-Nash equilibrium in Esponda and Pouzo (2016), my steady state generates data with positive KL divergence relative to the implied data distribution under the steady-state beliefs. That is, $\mathcal{H}^* (c^n) \neq \mathcal{H}(\Psi(\mu_1^n, \mu_2^n; \gamma); c^n)$, so the steady
state is not a self-confirming equilibrium. This is because for every censoring threshold \( c \) (and in particular for \( c = c^\infty \)), the KL divergences of the true history distribution to the history distributions under different feasible models is bounded away from 0.

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

**Definition 7.** For \( \gamma > 0 \), the iteration map \( \mathcal{I} : \mathbb{R} \to \mathbb{R} \) is given by

\[
\mathcal{I}(\mu_2; \gamma) := \mu_2^*(C(\mu_1^\bullet, \mu_2, \gamma))
\]

Given that \( \mu_1^*(c) = \mu_1^\bullet \) for all \( c \) from Proposition 2, it is straightforward to see that steady-state beliefs about \( \mu_2 \) are in bijection with fixed points of the iteration map \( I \). This shows steady-state belief about \( \mu_2 \) exhibits over-pessimism.

**Proposition 3.** Under Assumption 2, every steady state satisfies \( \mu_2^\infty < \mu_2^\bullet, \mu_1^\infty = \mu_1^\bullet \).

Furthermore, steady state is unique under the additional Assumption 3.

**Proposition 4.** Under Assumptions 1, 2, and 3, \( \mathcal{I} \) is a contraction mapping with contraction constant \( 0 < \ell_\gamma < 1 \). Therefore, a unique steady state exists.

The contraction mapping property of \( I \) comes from two lemmas. First, we can use the strict log-concavity assumption to show that \( \mu_2^*(c) \) is Lipschitz continuous with constant \( \gamma \).

**Lemma 1.** Under Assumption 2, \( \mu_2^*(c) \) is Lipschitz continuous with Lipschitz constant \( \gamma \).

Next, the indifference threshold is Lipschitz continuous with a Lipschitz constant strictly less than \( 1/\gamma \) once we adopt Assumption 3.

**Lemma 2.** Under Assumptions 1, 2, and 3, \( \mu_2 \mapsto C(\mu_1^\bullet, \mu_2; \gamma) \) is Lipschitz continuous with a Lipschitz constant \( \ell < 1/\gamma \).

Even under Assumptions 1 and 2 alone, the basic regularity conditions we maintain throughout, it turns out \( I \) is “almost” a contraction mapping for any \( \gamma > 0 \), in the sense that \( |I(\mu_2) - I(\mu_2^\prime)| < |\mu_2^\prime - \mu_2^\prime\prime| \) for every \( \mu_2, \mu_2^\prime, \mu_2^\prime\prime \in \mathbb{R} \). But, there is no guarantee of a uniform contraction constant strictly less than 1. The slight strengthening in Assumption 3 ensures such a uniform contraction constant exists.

I now show the steady-state stopping threshold always features stopping too early. For every \( \mu_1^\bullet, \mu_2^\bullet \in \mathbb{R} \), the objectively optimal stopping strategy takes the form of a cutoff \( c^\bullet \in

---

\(^{11}\) For example, under the history distribution \( \mathcal{H}^\bullet(c^\infty) \), \( \mathbb{E}[h_2|c^\infty - 1 \leq h_1 \leq c^\infty] = \mathbb{E}[h_2|c^\infty - 2 \leq h_1 \leq c^\infty - 1] \) since draws are objectively independent. However, under the history distribution driven by the steady-state feasible model \( \Psi(\mu_1^\infty, \mu_2^\infty; \gamma) \), we must have \( \mathbb{E}[h_2|c^\infty - 1 \leq h_1 \leq c^\infty] < \mathbb{E}[h_2|c^\infty - 2 \leq h_1 \leq c^\infty - 1] \) since \( \gamma > 0 \). This feature contrasts with Heidhues, Koszegi, and Strack (2018)’s model that results in a self-confirming learning outcome.
$\mathbb{R} \cup \{\pm \infty\}$, where $c^* = -\infty$ means always stopping and $c^* = \infty$ means never stopping.\footnote{This follows from Lemma A.2 in the Appendix, which shows even when $\gamma = 0$, the difference between stopping payoff at $x_1$ and expected continuation payoff after $x_1$ is strictly increasing and continuous in $x_1$.} I show that $c^* > c^\infty$ for every steady-state cutoff $c^\infty$. (This result only requires Assumptions 1 and 2 and does not require uniqueness of steady states.)

Early stopping does not directly follow from over-pessimism. In fact, outside of the steady state, there is an intuition that a biased agent may stop later than a rational agent, not earlier. For a concrete illustration, consider Example 1 with $q = 0$, so there is no probability of recall. Suppose the true fundamentals are $\mu_1^* \gg \mu_2^*$. If a biased agent has the correct beliefs about the fundamentals, she perceives a greater continuation value after $X_1 = \mu_2^*$ than a rational agent with the same correct beliefs, since the former holds a false expectation of positive reversals after a bad (relative to $\mu_1^*$) early draw. Even though $c^* = \mu_2^*$ and the rational agent chooses to stop, the biased agent chooses to continue and has an indifference threshold strictly above $c^*$. By continuity, the biased agent’s cutoff threshold remains strictly above $c^*$ even under slightly pessimistic beliefs about $\mu_2$.

Nevertheless, the next proposition shows that in the steady state, it is unambiguous that the biased agent stops too early relative to the objectively optimal threshold. The early-stopping result strengthens the over-pessimism result. In the steady state, agents must be sufficiently pessimistic as to overcome the opposite intuition about late stopping just discussed.

**Proposition 5.** Under Assumptions 1 and 2, every steady-state stopping threshold $c^\infty$ is strictly lower than the objectively optimal threshold, $c^*$.

To understand why, note the biased agent believes in different distributions of $X_2$ following different realizations of $X_1$, with more pessimistic beliefs after higher realizations. In a steady state $(\mu_1^\infty, \mu_2^\infty, c^\infty)$, the agent’s subjective belief about $X_2$ following $X_1 = c^\infty$ must be a leftward shift of the true distribution $f_2(\cdot | \mu_2^*)$. Else, the agent would have subjective distributions of $X_2$ that stochastically dominate the true distribution whenever $S_{c^\infty}$ prescribes continuing, so heuristically she could improve the fit of her model by lowering her belief about $\mu_2$. The biased agent’s indifference at $c^\infty$ is thus based on an overly pessimistic belief about the continuation value, so we must have $c^\infty < c^*$.

### 3.2 Social Learning with Agents Acting One by One

This section shows the steady state defined and studied earlier corresponds to the long-run learning outcome for a society of biased agents acting one at a time. I outline the convergence proof for a simpler variant of Theorem 1, where agents start off knowing $\mu_1^*$ and only entertain uncertainty over $\mu_2$. That is, the feasible fundamentals are given by Remark 1(c) rather than Remark 1(b). This simplification is without much loss: even when agents
are initially uncertain about $\mu_1$, they will almost surely learn it in the long run regardless of the stochastic process of their predecessors’ stopping strategies. Intuitively, this is because $X_1$ can never be censored, so no belief distortion in $\mu_1$ is possible.\footnote{This is similar to the intuition for why $\mu^*_1(c) = \mu^*_1$ for every $c$.} Once agents have learned $\mu^*_1$, the rest of the argument proceeds much like the case where $\mu^*_1$ is known from the start. In the next section I comment on the key steps in extending the proof to the case uncertainty over two-dimensional fundamentals $(\mu_1, \mu_2)$, but will defer the details to Online Appendix OA 3.

In the learning environment, time is discrete and partitioned into rounds $t = 1, 2, 3, \ldots$ One short-lived agent arrives per round. Agent 1 starts with a prior belief $M_0$ given by a prior density $m_0 : [\mu_2, \bar{\mu}_2] \to \mathbb{R}_{>0}$, while agent $t \geq 2$ adopts the final belief $\tilde{M}_{t-1}$ of agent $t - 1$ as her prior belief.\footnote{The same learning dynamics obtain in an environment where every agent starts with the common prior belief $M_0$ and observes the stage-game histories of all predecessors.} Next, agent $t$ chooses a cutoff threshold $\tilde{C}_t$ maximizing expected payoff based on expected utility and plays the stage game. She updates her belief from $\tilde{M}_{t-1}$ to $\tilde{M}_t$ by applying Bayes’ rule to her stage-game history, $\tilde{H}_t \in \mathbb{H}$.

The sequences $(\tilde{M}_t), (\tilde{C}_t), (\tilde{H}_t)$ are stochastic processes whose randomness stem from randomness of the stage-game draws realizations in different rounds. The convergence theorem is about the almost sure convergence of processes $(\tilde{M}_t)$ and $(\tilde{C}_t)$. To define the probability space formally, consider the $\mathbb{R}^2$-valued stochastic process $(X_t)_{t \geq 1} = (X_{1,t}, X_{2,t})_{t \geq 1}$, where $X_t$ and $X_{t'}$ are independent for $t \neq t'$. Within each $t$, $X_{1,t} \sim f_1(\cdot \mid \mu^*_1)$, $X_{2,t} \sim f_2(\cdot \mid \mu^*_2)$ are also independent. Interpret $X_t$ as the pair of potential draws in the $t$-th round of the stage game. Clearly, there exists a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, with sample space $\Omega = (\mathbb{R}^2)^\infty$ interpreted as paths of the process just described, $\mathcal{A}$ the Borel $\sigma$-algebra on $\Omega$, and $\mathbb{P}$ the measure on sample paths so that the process $X_t(\omega) = \omega_t$ has the desired distribution. The term “almost surely” means “with probability 1 with respect to the realization of the infinite sequence of all (potential) draws”, i.e. $\mathbb{P}$-almost surely. The processes $(\tilde{M}_t), (\tilde{C}_t), (\tilde{H}_t)$ are defined on this probability space and adapted to the filtration $(\mathcal{F}_t)_{t \geq 1}$, where $\mathcal{F}_t$ is the sub-$\sigma$-algebra generated by draws up to round $t$, $\mathcal{F}_t = \sigma((X_s)_{s=1}^t)$.

Under Assumptions 1, 2, and 3, by Proposition 4 there exists a unique steady state $(\mu^*_1, \mu^*_2, c^\infty)$. Theorem 1’ shows that, provided the support of $m_0$ contains $\mu^*_2$, $m'_2$ is continuous, and the second derivative of $\ln(f_2(\cdot \mid \mu^*_2))$ is uniformly bounded, the stochastic processes $(\tilde{C}_t)$ and $(\tilde{M}_t)$ almost surely converge to the steady state. This is a global convergence result since the bounded interval $[\mu_2, \bar{\mu}_2]$ can be arbitrarily large and the prior density $m_0$ can assign arbitrarily small probability to neighborhoods around $\mu^*_2$.\footnote{This is similar to the intuition for why $\mu^*_1(c) = \mu^*_1$ for every $c$.}
surely, \( \lim_{t \to \infty} \tilde{C}_t = c^\infty \) and \( \lim_{t \to \infty} \mathbb{E}_{\mu_2 \sim \tilde{N}_t} |\mu_2 - \mu_2^\infty| = 0 \), where \( c^\infty \) is the unique steady-state cutoff threshold.

I will now discuss the obstacles to proving convergence and provide an outline of my argument. In each round \( t \), the cutoff choice of the \( t \)-th agent determines how history \( \tilde{H}_t \) will be censored. We can think of each \( c \in \mathbb{R} \) as generating a different “type” of data. As we saw in Proposition 2, different “types” of data (in large samples) lead to different inferences about the fundamentals for biased agents, so the cutoff \( \tilde{C}_t \) influences the belief that will be passed on to agent \( t + 1 \). Yet \( \tilde{C}_t \) is an endogenous, ex-ante random object that depends on the belief of the \( t \)-th agent, meaning that belief and behavior co-evolve to complicate the analysis of learning dynamics.

To be more precise, the log-likelihood of all \( X_2 \) data up to the end of round \( t \) under fundamental \( \mu_2 \in [\mu_2^c, \tilde{\mu}_2] \) is the random variable

\[
\sum_{s=1}^{t} \ln(f_2(X_{2,s}; \mu_2) - \gamma(X_{1,s} - \mu_1^\ast)) \cdot 1\{X_{1,s} \leq \tilde{C}_s\}.
\]

The \( s \)-th summand contains the indicator \( 1\{X_{1,s} \leq C_s\} \), referring to the fact that \( X_{2,s} \) would be censored if \( X_{1,s} \) exceeds the cutoff \( \tilde{C}_s \). The cutoff \( \tilde{C}_s \) depends on histories in periods 1, 2, ..., \( s - 1 \), hence indirectly on \( (X_k)_{k<s} \). This makes the summands non-exchangeable: they are correlated and non-identically distributed. So the usual law of large numbers does not apply.

A first step to gaining traction on this problem is use a statistical tool from Heidhues, Koszegi, and Strack (2018), a version of law of large numbers for martingales whose quadratic variation grows linearly.

**Proposition 10** from Heidhues, Koszegi, and Strack (2018): Let \((y_t)_t\) be a martingale that satisfies a.s. \([y_t] \leq vt\) for some constant \( v \geq 0 \). We have that a.s. \( \lim_{t \to \infty} \frac{y_t}{t} = 0 \).

After simplifying the problem with this result, I establish a pair of mutual bounds on asymptotic behavior and asymptotic beliefs. If we know cutoff thresholds are asymptotically bounded between \( c^l \) and \( c^h \), \( c^l < c^h \), then beliefs about \( \mu_2 \) must be asymptotically supported on the interval \([\mu_2^*(c^l), \mu_2^*(c^h)]\). Conversely, if belief is asymptotically supported on the subinterval \([\mu_2^c, \tilde{\mu}_2] \subseteq [\mu_2^c, \mu_2^c] \), then cutoff thresholds must be asymptotically bounded between \( C(\mu_1^c, \mu_2^c; \gamma) \) and \( C(\mu_1^c, \tilde{\mu}_2; \gamma) \).

**Lemma A.19.** For \( c^l \geq C(\mu_1^c, \mu_2; \gamma) \), if almost surely \( \liminf_{t \to \infty} \tilde{C}_t \geq c^l \), then almost surely

\[
\lim_{t \to \infty} \tilde{M}_t([\mu_2^c, \mu_2^*(c^l)]) = 0.
\]

Also, for \( c^h \leq C(\mu_1^c, \tilde{\mu}_2; \gamma) \), if almost surely \( \limsup_{t \to \infty} \tilde{C}_t \leq c^h \), then almost surely

\[
\lim_{t \to \infty} \tilde{M}_t([\mu_2^*(c^h), \tilde{\mu}_2]) = 0.
\]
Lemma A.20. For $\mu_2 \leq \mu_1^l < \mu_2^h \leq \bar{\mu}_2$, if $\lim_{t \to \infty} \bar{M}_t([\mu_1^l, \mu_2^h]) = 1$ almost surely, then $\lim \inf_{t \to \infty} \bar{C}_t \geq C(\mu_1^l, \mu_2^h; \gamma)$ and $\lim \sup_{t \to \infty} \bar{C}_t \leq C(\mu_1^*, \mu_2^h; \gamma)$ almost surely.

Applying this pair of lemmas to $\text{supp}(M_0) = [\mu_2, \bar{\mu}_2]$, we conclude that asymptotically $\bar{M}_t$ must be supported on the subinterval $[I(\mu_2), I(\bar{\mu}_2)]$, where $I$ is the iteration map from Definition 7 first used in analyzing the existence and uniqueness of steady states. Under Assumptions 1, 2, and 3, Proposition 4 implies that $I$ is a contraction mapping whose iterates converge to $\mu_2^\infty$. Therefore by repeatedly applying the pair of Lemmas A.19 and A.20, we can refine the bound on asymptotic beliefs down to the singleton $\{\mu_2^\infty\}$, showing the almost-sure convergence of beliefs there. The almost-sure convergence of behavior follows easily from Lemma A.20.

### 3.3 Uncertainty About $\mu_1$

The hypotheses of Theorem 1 differ from those of Theorem 1’ in that agents start off with uncertainty about $\mu_1$. I now comment on the key step to proving almost-sure convergence of beliefs and behavior in the environment with two-dimensional uncertainty about fundamentals.

The structure of the inference problem in my setting is such that I can separately bound the agents’ asymptotic beliefs in two “directions,” thus reducing the task of proving a two-dimensional belief bound into a pair of tasks involving one-dimensional belief bounds. To understand why, consider a pair of fundamentals, $(\mu_1, \mu_2)$ and $(\mu_1', \mu_2') = (\mu_1 + d, \mu_2 - \gamma d)$ for some $d > 0$, satisfying $\mu_1, \mu_1' \leq \mu_1^\star$. That is, $(\mu_1, \mu_2)$ and $(\mu_1', \mu_2')$ lie on the same line with slope $-\gamma$. For any uncensored history $(x_1, x_2) \in \mathbb{R}^2$, the likelihood of second-period draw $x_2$ is the same under both pairs of fundamentals,

$$f_2(x_2 \mid \mu_2 - \gamma(x_1 - \mu_1)) = f_2(x_2 \mid \mu_2' - \gamma(x_1 - \mu_1')).$$

This is because the feasible model $\Psi(\mu_1, \mu_2; \gamma)$ has a lower first-period mean but also a higher second-period unconditional mean, compared to the model $\Psi(\mu_1', \mu_2'; \gamma)$. An agent who believes in the first model feels less disappointed by the draw $x_1$, since she evaluates it against a lower expectation. This leads a weaker anticipation of positive reversal under the gambler’s fallacy, compared to another agent who believes in the second model. But, this difference is canceled out by the more optimistic belief about the unconditional distribution of second-period draw, $\mu_2 > \mu_2'$.

This argument shows that both pairs of fundamentals $(\mu_1, \mu_2)$ and $(\mu_1', \mu_2')$ explain $X_2$ data equally well in all uncensored histories. This is important as it shows regardless of agent $t$’s strategy, she would always find that $(\mu_1, \mu_2)$ and $(\mu_1', \mu_2')$ lead to the same likelihood of second-period data in her history $\tilde{H}_t$. At the same time, $(\mu_1', \mu_2')$ provides a strictly better
fit for $X_1$ data on average than $(\mu_1, \mu_2)$, since $\mu_1 < \mu'_1 \leq \mu_1^*$, This means in the long run, fundamentals $(\mu_1, \mu_2)$ should receive much less posterior probability than $(\mu'_1, \mu'_2)$, as the latter better rationalize the data overall.

This heuristic comparison of the asymptotic goodness-of-fit for two feasible models is formalized by computing the directional derivative for data log-likelihood along the vector \( \begin{pmatrix} 1 \\ -\gamma \end{pmatrix} \) in the space of fundamentals. I establish an (almost-sure) positive lowerbound on this directional derivative to the left of $\mu_1^*$, and an analogous negative upperbound to the right of $\mu_1^*$. This allows me to show the region colored in red receives 0 posterior probability asymptotically, by comparing each point in red with a corresponding point in blue along a line of slope $-\gamma$. By repeating this argument (and applying the symmetric bound to the right of $\mu_1^*$), I show that belief is asymptotically concentrated along an $\epsilon$-width vertical strip containing the steady state beliefs, $(\mu_1^*, \mu_2^\infty)$.

![Diagram](image)

Having restricted the long-run belief to a small vertical strip, we have completed one “direction” of the belief bounds and effectively reduced the dimensionality of uncertainty back to one. The rest of the argument proceeds similarly to the case where agents know $\mu_1^*$ discussed before, iteratively restricting agents’ asymptotic behavior and asymptotic belief about $\mu_2$. These restrictions amount to “vertical” belief refinements within the $\epsilon$-strip, so eventually belief is restricted to the single point $(\mu_1^*, \mu_2^\infty)$, the unique steady-state belief.

### 4 The Positive-Feedback Cycle

In this section, I consider a social-learning environment where agents arrive in large generations and all agents in the same generation act simultaneously. I will prove Theorem 2, fully characterizing the learning dynamics in this environment. I will also discuss the positive-feedback loop between distorted beliefs about fundamentals and distorted stopping behavior.
4.1 Social Learning in Large Generations

There is an infinite sequence of generations, \( t \in \{0, 1, 2, \ldots \} \). Each generation is “large” and will be modeled as a continuum of agents, \( n \in [0, 1] \). In the search problem of Example 1, for instance, different generations refer to cohorts of HR managers working in different hiring cycles. Each agent lives for one generation, so agent \( n \) from generation 1 is unrelated to agent \( n \) from generation 2. The realizations of draws \( X_1, X_2 \) are independent across all stage games, including those from the same generation. Generation 0 agents play some strategy \( S_{c[0]} \), where \( c[0] \in \mathbb{R} \) is the initial condition of social learning.

Write \( h_{\tau,n} \in \mathbb{H} \) for the stage-game history of agent \( n \) from generation \( \tau \). Before playing her own stage game, each agent in generation \( t \geq 1 \) observes an infinite dataset consisting of all histories \( (h_{\tau,n})_{n \in [0,1]} \) from each predecessor generation, \( 0 \leq \tau \leq t-1 \). If all \(^{15}\) generation \( \tau \) predecessors used the stopping strategy \( S_{c_{\tau}} \) for some \( c_{\tau} \in \mathbb{R} \), then the sub-dataset \( (h_{\tau,n})_{n \in [0,1]} \) has the distribution \( \mathcal{H}^*(c_{\tau}) \). Agents are told the stopping strategies of their predecessors from all past generations\(^ {16} \) and use the entire dataset of histories to infer fundamentals. The space of feasible fundamentals is \( \mathcal{M} = \mathbb{R}^2 \) as in Remark 1(a), so agents can flexibly estimate the unconditional means of draws from different periods, subject to their dogmatic belief in reversals.

Agents only infer from predecessors’ histories, not from their behavior. This is rational as information sets are nested across generations. For \( t_2 > t_1 \), generation \( t_2 \) observes all the social information that generation \( t_1 \) saw. In addition, generation \( t_2 \)’s dataset contains a complete record of everything that happened in generation \( t_1 \)’s stage games. Since generation \( t_1 \) has no private information that is unobserved by generation \( t_2 \), the behavior of these predecessors is uninformative about the fundamentals beyond what generation \( t_2 \) can learn from the dataset of histories.

In the large-generations model, generation \( t \) agents infer fundamentals \( (\mu_1[t], \mu_2[t]) \) that minimize the sum of the KL divergences between the implied history distribution under the feasible model \( \Psi(\mu_1[t], \mu_2[t]; \gamma) \) on the one hand, and the \( t \) observed history distributions in generations \( 0 \leq \tau \leq t-1 \) on the other hand. Then, these agents use the stopping strategy optimal for the inferred feasible model. I formally define generation \( t \)’s minimization objective below.

**Definition 8.** The large-generations pseudo-true fundamentals with respect to cutoff thresholds \( (c_{\tau})_{\tau=0}^{t-1} \) solve

\[
\min_{\mu_1, \mu_2 \in \mathbb{R}} \sum_{\tau=0}^{t-1} D_{KL}( \mathcal{H}^*(c_{\tau}) \| \mathcal{H}(\Psi(\mu_1, \mu_2; \gamma); c_{\tau}) ),
\]

where \( D_{KL} \) is KL divergence from Definition 5. Denote the minimizers as \( \mu_1^*(c_0, \ldots, c_{t-1}) \) and \( \mu_2^*(c_0, \ldots, c_{t-1}) \). All generation \( \tau \) predecessors had the same information about the fundamentals, so all of them would have found the same stopping strategy subjectively optimal.

\(^{15}\) These stopping rules can also be exactly inferred from the infinite dataset.
\( \mu^*(c_0, \ldots, c_{t-1}) \).

I interpret the continuum of agents in each generation as an idealized, tractable modeling device representing a large but finite number of agents. Appendix OA 4 provides a finite-population foundation for inference and behavior in the continuum-population model. There, for the Gaussian case, I show that when an agent observe \( t \) finite sub-datasets of histories drawn from distributions \( H^*_{\tau}(c_r) \) for \( 0 \leq \tau \leq t-1 \), as these datasets grow large her inference and behavior almost surely converge to the infinite-population analogs.

### 4.2 Learning Dynamics in Large Generations

Now I develop the proof of Theorem 2.

**Theorem 2.** Suppose Assumptions 1 and 2 hold. Starting from any initial condition and any \( m_0 \), cutoffs \( (c_{[t]} \rangle_{t \geq 1} \) and beliefs \( (\mu_{[t]} \rangle_{t \geq 1} \) form monotonic sequences across generations. When Assumption 3 also holds, there exists a unique steady state \( \mu_{\infty}^2, c_{\infty} \in \mathbb{R} \) so that \( c_{[t]} \to c_{\infty} \) and \( (\mu_{1,[t]}, \mu_{2,[t]} \to (\mu_{\infty}^1, \mu_{\infty}^2) \) monotonically, regardless of the initial condition and \( m_0 \). This steady state is the same as the one in Theorem 1.

Towards a proof, first consider learning dynamics in a related auxiliary environment. The auxiliary environment is identical to the large-generations environment just described, except that agents in each generation \( t \geq 1 \) only infer from the histories of the immediate predecessor generation, \( t - 1 \). Write \( \mu_{[t]}^A \) and \( c_{[t]}^A \) for the inference and cutoff threshold in generation \( t \) of the auxiliary environment, where the superscript “A” distinguishes them from the corresponding processes of the baseline large-generations environment.

We have \( \mu_{1,[t]}^A = \mu_{1}^* \) for every \( t \geq 1 \), from from Proposition 2. Also, it is easy to see that \( (\mu_{2,[t]}^A)_{t \geq 1} \) are iterates of the \( \mathcal{I} \) map from Definition 7, and that they must be monotonic since the pair of comparative statics \( \frac{\partial C}{\partial \mu_2} > 0 \) and \( \frac{d \mu_2^*}{dc} > 0 \) share the same sign. Changes in beliefs across successive generations are amplified, not dampened, by the corresponding changes in cutoff thresholds.

The monotonicity and convergence results of Theorem 2 follow from comparing the learning dynamics of the baseline large-generations environment to the dynamics of the auxiliary environment. Roughly speaking, for agents in late enough generations of the baseline large-generations environment, most of the histories in their dataset are censored according to stopping thresholds very similar to the limit threshold. So, observing sub-datasets of histories from all past generations induces very similar inferences as observing just one dataset of histories from the immediate predecessor generation.

While the large-generations environment I set out to study has the same long-run learning outcome as the auxiliary environment, the two environments may differ in their short-run welfare. For example, in settings where learning leads generations further and further astray from the objectively optimal strategy, the auxiliary environment speeds up this harmful
Figure 1: The dynamics of beliefs about $\mu_2$ in the first four generations for the Gaussian case (with $\sigma^2 = 1$). The stage game is search (without recall), with true fundamentals are $\mu_1^* = \mu_2^* = 0$, bias parameter $\gamma > 0$, and initial condition is $c_{[0]} = 0$. In both the baseline large-generations environment and the auxiliary environment, beliefs are monotonic across generations, an illustration of Theorem 2. Beliefs in both environments converge to the same steady-state beliefs, though the rate of convergence is faster in the auxiliary environment.

learning. This is because the less-censored histories from the earlier generations no longer moderate the society’s descent into pessimism when agents only infer from the immediate predecessor generation. In Figure 1, I plot the dynamics of beliefs in the first four generations for a society playing the stage game from Example 1 with $q = 0$. Consider the Gaussian case with $\mu_1^* = \mu_2^* = 0$, $\gamma = -0.5$, $\sigma^2 = 1$, and the society starts at the objectively optimal cutoff threshold, $c_{[0]} = 0$. Society mislearns monotonically in both the baseline large-generations environment and the auxiliary environment. This mislearning is more exaggerated in the in the auxiliary environment, but both environments lead to the same long-run outcome.

The map $\mathcal{I}(\cdot; \gamma)$ connects the environment where large generations of agents act simultaneously to the environment where agents act one by one. We can think of $\mathcal{I}(\cdot; \gamma)$ as the one-generation-forward belief map in the auxiliary society, whose belief dynamics are closely related to the belief dynamics of the baseline large-generations environment. There are no large generations at all in the environment where agents act one by one, but there $\mathcal{I}$ still plays a critical role in establishing the long-run convergence of beliefs and behavior. Intuitively, in the learning environment of Section 3, a belief based on the histories of one predecessor from each of many past generations replaces a belief based on a large dataset of histories from many agents all belonging to the same past generation.

I combine the asymptotic early-stopping result of Theorem 1 with the monotonic learning dynamics of Theorem 2 to deduce:

**Corollary 1.** Suppose Assumptions 1, 2, and 3 hold. In the large-generations environment,
if society starts at the objectively optimal initial condition $c_0 = c^*$, then expected payoff strictly decreases across all successive generations.

This stark “monotonic” mislearning result relies crucially on the endogenous-data setting. Each generation uses a lower acceptance threshold relative to their predecessors, a change with the side effect of changing censoring threshold of their successors’ data. The new “type” of censored data causes the next generation to become more pessimism about the fundamentals than any past generation.

5 Extensions

In this section I explore a number of alternative model specifications to examine the robustness of my main results. The Online Appendix contains the proofs of results in this section and additional extensions.

5.1 Comparative Statics

In the first extension, I consider how learning dynamics react to changes in stage-game parameters. In general, when agents learn from exogenous data, their decision problem does not influence learning outcomes. This observation holds independently of whether agents are misspecified. On the other hand, correctly specified agents in my setting always learn correctly in the long run, so the stage game is again irrelevant. With misspecified learners in an endogenous-data setting, however, changes in the stage game carry long-run consequences on agents’ beliefs about the fundamentals.

**Definition 9.** Given a pair of second-period payoff functions $u^H, u^L$, say $u^H$ payoff dominates $u^L$ (abbreviated $u^H \succ u^L$) if for every $x_1 \in \mathbb{R}$, $u^H(x_1, x_2) \geq u^L(x_1, x_2)$ for every $x_2 \in \mathbb{R}$, and also $u^H(x_1, x_2) > u^L(x_1, x_2)$ for a positive-measure set of $x_2$ in $\mathbb{R}$.

For instance, in Example 1, increasing $q$ (the probability of recall) creates a new optimal-stopping problem that payoff dominates the old one. More generally, starting from any stage game with payoff functions $u_1$ and $u_2$, we can impose an extra waiting cost $\kappa_{\text{wait}} > 0$ for continuing into the second period. This generates a new stage game with payoff functions $u_1$ and $u_2^L$ with $u_2^L = u_2 - \kappa_{\text{wait}}$. The modified stage game is payoff dominated by the unmodified one.

When $u^H \succ u^L$, a society facing the problem $(u_1, u^H)$ always uses a higher stopping threshold than a society facing the problem $(u_1, u^L)$, given the same beliefs about fundamentals. To state this formally, let $C_{u_1, u_2}$ be the optimal cutoff threshold function for the stage game $(u_1, u_2)$.

**Lemma 3.** Suppose Assumption 1 holds for stage games $(u_1, u^H)$ and $(u_1, u^L)$, and $u^H \succ u^L$. For all $\mu_1, \mu_2 \in \mathbb{R}$, $\gamma > 0$, $C_{u_1, u^H}(\mu_1, \mu_2; \gamma) > C_{u_1, u^L}(\mu_1, \mu_2; \gamma)$. 

28
The next proposition shows that when one stage game payoff dominated another in terms of second-period payoffs, the dominated stage game leads to more pessimistic beliefs and a lower cutoff threshold in the steady state.

**Proposition 6.** Suppose both \((u_1, u^H_2)\) and \((u_1, u^L_2)\) satisfy Assumptions 1 and 3, and that \(u^H_2 \succ u^L_2\). Under Assumption 2, the steady state of \((u_1, u^H_2)\) features strictly more optimistic belief about the second-period fundamental and a strictly higher cutoff threshold than the steady state of \((u_1, u^L_2)\).

Combined with my main results on learning dynamics (Theorems 1 and 2), Proposition 6 illustrates how changing the stage-game payoff structure affects long-run inference. Consider two societies of gambler’s fallacy agents with the same bias parameter \(\gamma > 0\), facing stage games \((u_1, u^H_2)\) and \((u_1, u^L_2)\) respectively, where \(u^H_2\) payoff dominates \(u^L_2\). Even if the latter society starts with a much more optimistic belief about \(\mu_2\), in the long run the second society will end up with strictly more pessimistic beliefs and will use strictly lower cutoff thresholds. Since steady-state beliefs are too pessimistic in both societies, the second society’s long-run beliefs are more distorted.

This comparative statics result provides novel predictions about how the economic environment affects biased agents’ inference. Applied to the hiring context from Example 1, this result says when managers are more impatient (i.e., suffer a greater waiting cost) or when they have a lower chance of recalling previous applicants, then they will end up with more pessimistic beliefs about the labor pool. The direction of the comparative statics is another expression of the positive-feedback cycle between stopping threshold and inference. When managers become more impatient, for instance, they use a lower acceptance threshold as they wish to finish recruiting earlier. The lower cutoff intensifies the censoring effect on histories, leading to more pessimistic inference about the fundamentals. The extra pessimism, in turn, leads future managers to further lower their acceptance threshold, amplifying the initial change in behavior that came from the increase in waiting cost.

From a policy perspective, subsidizing longer search (i.e., decreasing \(\kappa_{\text{wait}}\)) unambiguously improves asymptotic learning accuracy for biased agents. So, even a policymaker who is ignorant of \((\mu^*_1, \mu^*_2)\) can partially correct society’s long-run beliefs. We can also think of this policy as a test of misspecification, as it alters steady-state beliefs only when agents are misspecified. The test can be conducted without knowledge of the true data-generating process.

### 5.2 Fictitious Variation and Censored Datasets

For this and subsequent extensions, I specialize to the Gaussian case.

So far, I have assumed agents hold dogmatic and correct beliefs about the variance of \(X_1\) and the conditional variance of \(X_2 | (X_1 = x_1)\). In this extension, I expand the set of
feasible models and consider agents who are uncertain about the variances of the draws and jointly estimate means and variances. I show that agents exaggerate variances in a way that depends on the severity of data censoring, and study how this belief in fictitious variation strengthens the positive-feedback cycle between beliefs and behavior.

For \( \mu_1, \mu_2 \in \mathbb{R}, \sigma_1^2, \sigma_2^2 \geq 0 \), and \( \gamma \geq 0 \), let \( \Psi(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \gamma) \) refer to the joint distribution

\[
X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2) ~\quad (X_2 \mid X_1 = x_1) \sim \mathcal{N}(\mu_2 - \gamma(x_1 - \mu_1), \sigma_2^2).
\]

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

\[
\{ \Psi(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \gamma) : \mu_1, \mu_2 \in \mathbb{R}, \sigma_1^2, \sigma_2^2 \geq 0 \}
\]

for a fixed bias parameter \( \gamma > 0 \). For this extension, “fundamentals” refer to the four parameters \( \mu_1, \mu_2, \sigma_1^2, \sigma_2^2 \).

Following Definition 5, write \( D_{KL}(\mathcal{H}^*(c) \parallel \mathcal{H}(\Psi(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \gamma); c)) \) to denote the KL divergence between the true distribution of histories with censoring threshold \( c \) and the implied history distribution under the fundamentals \( \mu_1, \mu_2, \sigma_1^2, \sigma_2^2 \). This divergence is given by

\[
\int_c^\infty \phi(x_1; \mu_1^*, (\sigma^*)^2) \cdot \ln \left( \frac{\phi(x_1; \mu_1^*, (\sigma^*)^2)}{\phi(x_1; \mu_1, \sigma_1^2)} \right) dx_1
\]

\[
+ \int_0^c \left\{ \int_c^\infty \phi(x_1; \mu_1^*, (\sigma^*)^2) \cdot \phi(x_2; \mu_2^*, (\sigma^*)^2) \cdot \ln \left[ \frac{\phi(x_1; \mu_1^*, (\sigma^*)^2) \cdot \phi(x_2; \mu_2^*, (\sigma^*)^2)}{\phi(x_1; \mu_1, \sigma_1^2) \cdot \phi(x_2; \mu_2 - \gamma(x_1 - \mu_1), \sigma_2^2)} \right] dx_2 \right\} dx_1,
\]

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

The next Proposition characterizes the pseudo-true fundamentals \( \mu_1^*, \mu_2^*, (\sigma_1^*)^2, (\sigma_2^*)^2 \) that minimize Equation (2) in closed-form expressions.

**Proposition 7.** The solutions of

\[
\min_{\mu_1, \mu_2 \in \mathbb{R}, \sigma_1^2, \sigma_2^2 \geq 0} D_{KL}(\mathcal{H}^*(c) \parallel \mathcal{H}(\Psi(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \gamma); c))
\]

are \( \mu_1^* = \mu_1^*, \mu_2^* = \mu_2^* - \gamma (\mu_1^* - \mathbb{E}[X_1 \mid X_1 \leq c]), (\sigma_1^*)^2 = (\sigma^*)^2 \), and

\[
(\sigma_2^*)^2 = (\sigma^*)^2 + \gamma^2 \text{Var}[X_1 \mid X_1 \leq c].
\]

Comparing Proposition 7 with the expressions for \( \mu_1^*, \mu_2^* \) in Example 2 shows that the agent makes the same misinference about the means regardless of whether she knows the
variances. This shows the robustness of the over-pessimism prediction in an environment where agents jointly estimate both means and variances.

Biased agents correctly estimate the first-period variance, \((\sigma^*_1)^2 = (\sigma^*)^2\), but overestimate second-period variance. The magnitude of this distortion increases in the severity of the gambler’s fallacy but decreases with the severity of the censoring, as \(\text{Var}[X_1 | X_1 \leq c]\) increases in \(c\) for \(X_1\) Gaussian.

Here is the intuition. Whereas the objective conditional distribution of \(X_2|(X_1 = x_1)\) is independent of \(x_1\), the agent entertains different beliefs about this distribution for different \(x_1\). The agent’s inference about \(\mu^*_2\) ensures her belief about \(X_2|X_1 = x_1\) fits the data well following “typical” realizations of \(x_1\) under the censoring restriction \(X_1 \leq c\). But the agent continues to be surprised by streaks of bad draws in the data. To better account for these surprising observations, the agent increases estimated conditional variance of \(X_2|(X_1 = x_1)\) and attributes these unexpected realizations of \(X_2\) to “noise.” More “noise” is needed when \(\text{Var}[X_1|X_1 \leq c]\) larger, for the frequency of these surprising observations depends on how much \(X_1\) tends to deviate from its typical value of \(\mathbb{E}[X_1|X_1 \leq c]\) under the restriction \(X_1 \leq c\).

An equivalent formulation of this result helps interpret the distorted \((\sigma^*_2)^2\). We may write the feasible model \(\Psi(\mu_1, \mu_2, \sigma^2_1, \sigma^2_2; \gamma)\) with \(\sigma^2_2 \geq \sigma^2_1\) as

\[
X_1 = \mu_1 + \epsilon_1 \\
X_2 = \mu_2 + \zeta + \epsilon_2
\]

where \(\epsilon_1 \sim \mathcal{N}(0, \sigma^2_1)\), \(\epsilon_2|\epsilon_1 \sim \mathcal{N}(-\gamma \epsilon_1, \sigma^2_1)\), and \(\zeta \sim \mathcal{N}(0, \sigma^2_\zeta)\), 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 shock to the average quality of the second-period applicant. A positive \(\sigma^2_\zeta\) 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 7 says that in an environment where all jobs are objectively homogeneous with respect to the fit of the late applicants, biased managers who find it possible that jobs are heterogeneous in this dimension will indeed estimate a positive amount of this heterogeneity, \(\sigma^2_\zeta > 0\), 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. That is, the biased managers reason that the realization of the vacancy-specific fixed effect, \(\zeta\), must have been far from 0.

This phenomenon relates to findings in Rabin (2002) and Rabin and Vayanos (2010), who refer to exaggeration of variance under the gambler’s fallacy as fictitious variation. The key innovation of Proposition 7 is to show, in an endogenous-data setting, how the degree of fictitious variation depends on the severity of censoring. To highlight this point, I now derive two results focusing on the interplay between fictitious variation and endogenous censoring.
For simplicity, I derive these results using the auxiliary large-generations environment defined in Section 4.1, where agents arrive in large generations and only infer from the histories of the immediate predecessor generation.

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 4 continues to obtain — cutoffs, beliefs about fundamentals, and beliefs about variances form monotonic sequences across generations. This weak convexity includes the case of search with recall in Example 1 for any recall probability $0 \leq q < 1$ and highlights a new channel for amplifying changes in behavior across generations — inference about variance.

**Definition 10.** 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 subset of $\mathbb{R}$. 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 subset of $\mathbb{R}$.

**Proposition 8.** Suppose the optimal-stopping problem is convex. Suppose agents start with a full-support prior over $\{\Psi(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2; \gamma) : \mu_1, \mu_2 \in \mathbb{R}, \sigma_1^2, \sigma_2^2 \geq 0\}$ and society starts at the initial condition $c[0] \in \mathbb{R}$. For $t \geq 1$, denote the beliefs of generation $t$ as $(\mu_1[t], \mu_2[t], \sigma_1^2[t], \sigma_2^2[t])$ and their cutoff threshold as $c[t]$. Then $\mu_1[t] = \mu_1^*$, $(\sigma_1^2[t])^2 = (\sigma^*)^2$ for all $t$, while $(\mu_2[t])_{t \geq 1}$, $(\sigma_2^2[t])_{t \geq 1}$, and $(c[t])_{t \geq 1}$ are monotonic sequences.

The intuition is straightforward. Suppose generation $t$ uses a more relaxed acceptance threshold $c[t] < c[t-1]$ than generation $t-1$, resulting in a more severely censored dataset. By the usual censoring effect without variance uncertainty, generation $t+1$ becomes more pessimistic about second-period mean than generation $t$. In addition, by Proposition 7 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 means, for a larger variance in $X_2(X_1 = x_1)$ improves the expected payoff when continuing due to the convexity of $u_2$ in $x_2$. Combining these two forces, we deduce $c[t+1] < c[t]$.

The intuition just discussed shows that uncertainty about variance strengthens the monotonicity result. To be more precise, suppose $c[t] < c[t-1]$. Consider a hypothetical generation $t+1$ agent who dogmatically adopts generation $t$’s beliefs about variances, $\sigma_1^2[t]$ and $\sigma_2^2[t]$, and infers from the class of models $\{\Psi(\mu_1, \mu_2, \sigma_1^2[t], \sigma_2^2[t]; \gamma) : \mu_1, \mu_2 \in \mathbb{R}\}$. Based on generation $t$’s histories, this hypothetical agent makes inferences about means and chooses a cutoff threshold, $\tilde{\mu}_1[t+1], \tilde{\mu}_2[t+1], \tilde{c}[t+1]$. By comparing Proposition 7 and Example 2, $\tilde{\mu}_1[t+1] = \mu_1[t+1]$, $\tilde{\mu}_2[t+1] = \mu_2[t+1]$, but $c[t+1] < \tilde{c}[t+1] < c[t]$. That is, while the cutoff threshold of this hypothetical agent follows the monotonicity pattern in the previous two generations, $\tilde{c}[t+1] < c[t] < c[t-1]$, the cutoff adjusts downwards by an even greater amount, $c[t+1] < \tilde{c}[t+1]$, when agents are uncertain about variances.
The second result compares the learning dynamics of two societies facing the same optimal-stopping problem and starting at the same initial condition. 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 9 shows that in Generation 1, the two societies hold the same beliefs about the means of the distributions. But in all later generations \(t \geq 2\), the society with uncertainty about variances ends up with a more pessimistic/optimistic belief about the second-period mean 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 7 implies that the two societies make the same inferences about the means when given the same data. 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 \(\{\Psi(\mu_1, \mu_2, (\sigma^*)^2, (\sigma^*^2) ; \gamma) : \mu_1, \mu_2 \in \mathbb{R}\}\). Agents in society B start with a full-support prior over \(\{\Psi(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2 ; \gamma) : \mu_1, \mu_2 \in \mathbb{R}, \sigma_1^2, \sigma_2^2 \geq 0\}\). Fix the same Generation 0 initial condition \(c_{[0]} \in \mathbb{R}\) for both societies. For \(t \geq 1\), denote the beliefs of Generation \(t\) in society \(k \in \{A, B\}\) as \((\mu_{1,[k,t]}, \mu_{2,[k,t]}, (\sigma_{1,[k,t]}^2, (\sigma_{2,[k,t]}^2)^2)\) and their cutoff threshold as \(c_{[k,t]}\).

**Proposition 9.** 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, then \(\mu_{2,[B,t]} < \mu_{2,[A,t]}\) and \(c_{[B,t]} < c_{[A,t]}\) for every \(t \geq 2\).

### 5.3 Objectively Correlated \((X_1, X_2)\) and Uncertainty About \(\gamma\)

So far I have assumed that draws \((X_1, X_2)\) within the stage game are objectively independent, and that agents have a dogmatic \(\gamma > 0\), interpreted as the severity of the gambler’s fallacy bias. This extension considers a joint relaxation of these two assumptions.

Suppose the true model is \((X_1, X_2) \sim \Psi(\mu_1, \mu_2, \gamma^*)\), where possibly \(\gamma^* \neq 0\). Agents jointly estimate \((\mu_1, \mu_2, \gamma) \in \mathbb{R}^3\), with a prior supported on \(\mathbb{R} \times \mathbb{R} \times [\gamma, \bar{\gamma}]\) where \([\gamma, \bar{\gamma}]\) is a finite interval. The next result generalizes the pseudo-true fundamentals in Example 2. It shows that when \(\gamma^* \notin [\gamma, \bar{\gamma}]\), the agent infers \(\gamma^*\) equal to the boundary point of the interval that is closer to \(\gamma^*\). Given the estimated pseudo-true parameter \(\gamma^*\), the estimates of the first- and second-period fundamentals take similar forms to those in Example 2.

**Proposition 10.** Suppose \(\gamma^* \notin [\gamma, \bar{\gamma}]\). Let \(\tilde{\gamma} = \bar{\gamma}\) if \(\gamma^* > \tilde{\gamma}\), otherwise \(\tilde{\gamma} = \gamma\) when \(\gamma^* < \gamma\).

The solution of the KL-divergence minimization problem

\[
\min_{\mu_1, \mu_2 \in \mathbb{R}, \gamma \in [\gamma, \bar{\gamma}]} D_{KL}(\mathcal{H}(\Psi(\mu_1, \mu_2, \gamma^*) ; c) \mid \mid \mathcal{H}(\Psi(\mu_1, \mu_2, \gamma) ; c))
\]

is given by \(\mu_1^*(c) = \mu_1^*, \mu_2^*(c) = \mu_2^* + (\gamma^* - \tilde{\gamma}) \cdot (\mu_1^* - \mathbb{E}_{\Psi^*}[X_1 | X_1 \leq c]), \gamma^*(c) = \tilde{\gamma}\). 

33
Intuitively, we may expect the closest distance (in the KL divergence sense) from the set of feasible models \(\{\Psi(\mu_1, \mu_2; \gamma) : \mu_1, \mu_2 \in \mathbb{R}\}\) to the objective distribution \(\Psi(\mu_1^*, \mu_2^*; \gamma^*)\) to decrease in \(|\gamma - \gamma^*|\). Proposition 10 confirms this intuition, showing that the pseudo-true model from the set \(\{\Psi(\mu_1, \mu_2; \gamma) : \mu_1, \mu_2 \in \mathbb{R}, \gamma \in [\bar{\gamma}, \tilde{\gamma}]\}\) lies in the subset \(\{\Psi(\mu_1, \mu_2; \gamma) : \mu_1, \mu_2 \in \mathbb{R}, \gamma = \tilde{\gamma}\}\), where \(\tilde{\gamma}\) is the closest point (in the Euclidean sense) to \(\gamma^*\) in the interval [\(\bar{\gamma}, \tilde{\gamma}\)].

When \(\gamma^* = 0\) and \(\tilde{\gamma} < 0\), this result shows that over-pessimism in inference is robust to agents learning the correlation of \(X_1\) and \(X_2\), provided the support of their uncertainty about correlation lies to the left of 0 and excludes 0. In this case, it is also easy to see that the learning dynamics in the large-generations auxiliary environment are the same as when agents start with a dogmatic belief in \(\gamma = \tilde{\gamma}\).

### 5.4 Inference under the Constraint \(\mu_1 = \mu_2\)

I now consider the special case where the true fundamentals are time-invariant, \(\mu_1^* = \mu_2^* = \mu^* \in \mathbb{R}\). If agents’ feasible fundamentals are \(\mathcal{M} = \mathbb{R}^2\) as in Remark 1(a), then Proposition 2 continues to apply. But now suppose agents know the fundamentals are time-invariant and only have uncertainty over this common value, so the set of feasible fundamentals is the diagonal \(\mathcal{M} = \{(x, x) : x \in \mathbb{R}\}\), as in Remark 1(d). I investigate inference in this setting when agents’ prior belief over feasible models is supported on \(\{\Psi(\mu, \mu; \gamma) : \mu \in \mathbb{R}\}\).

Let \(\mu_1^*(c) \in \mathbb{R}\) stand for the common fundamental that minimizes the KL divergence relative to the history distribution \(\mathcal{H}^*(c)\), that is

\[
\mu_1^*(c) := \arg\min_{\mu \in \mathbb{R}} D_{KL}(\mathcal{H}^*(c) \parallel \mathcal{H}(\Psi(\mu, \mu; \gamma); c))
\]

The next result characterizes \(\mu_1^*(c)\).

**Proposition 11.** \(\mu_1^*(c)\) equals \(\frac{1}{1+\mathbb{P}[X_1 \leq \bar{c}-(1+\gamma)^2]} \mu_1^1(c) + \frac{\mathbb{P}[X_1 \leq \bar{c}-(1+\gamma)^2]}{1+\mathbb{P}[X_1 \leq \bar{c}-(1+\gamma)^2]} \mu_1^2(c)\), where \(\mu_1^1(c) = \mu^*\) and \(\mu_1^2(c) = \mu^* - \frac{\gamma}{1+\gamma} (\mu^* - \mathbb{E}[X_1 \mid X_1 \leq \bar{c}])\).

Agents face two kinds of data about the common fundamental: observations of first-period draws and observations of second-period draws. Feasible models \(\Psi(\mu_1^*(c), \mu_2^*(c); \gamma)\) and \(\Psi(\mu_2^*(c), \mu_2^*(c); \gamma)\) minimize the KL divergence of these two kinds of data, respectively.\(^{17}\)

The overall KL-divergence minimizing estimator is a certain convex combination between these two points. Through the term \(\mathbb{P}[X_1 \leq \bar{c}]\), the relative weight given to \(\mu_2^2(c)\) increases as the cutoff \(c\) increases, because the second-period data is observed more often if the dataset of histories is censored with a higher cutoff in the first period.

\(^{17}\)Note that \(\mu_2^2(c)\) differs from the pseudo-true fundamental \(\mu_2^*(c)\) from Example 2. The estimator \(\mu_2^2(c)\) 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, \(\mu_1^*\).
For any censoring threshold $c$ generating the history distribution, agents underestimate the common fundamental. We have $\mu_2^c(c) < \mu^*$ while $\mu_1^c(c) = \mu^*$. This shows the robustness of the over-pessimism result from the setting with $\mathcal{M} = \mathbb{R}^2$. However, the extent of over-pessimism about $\mu_2$ is dampened relative to agents who can flexibly estimate different $\mu_1$ and $\mu_2$ for the two periods. Compared with the unconstrained pseudo-true fundamentals from Example 2, we have $\mu_2^c(c) > \mu_2^*(c)$ since $\frac{\gamma}{1+\gamma} < \gamma$, hence $\mu_2^*(c) > \mu_2^c(c)$. This makes intuitive sense: when unconstrained, agents come to two different beliefs about $\mu_1$ and $\mu_2$, even though they are objectively the same. They hold correct beliefs about $\mu_1$ but pessimistic beliefs about $\mu_2$. When constrained to a common inference across two fundamentals, agents distort their belief about $\mu_1$ downwards and their belief about $\mu_2$ upwards, relative to the unconstrained environment.

6 Concluding Remarks

This paper studies endogenous learning dynamics of misspecified agents. I focus on the gambler’s fallacy, a non-self-confirming misspecification where no feasible beliefs of the biased agents can exactly match the data. In natural optimal-stopping problems, agents tend to stop after “good enough” early draws, where the threshold for “good enough” evolves as agents update their beliefs about the underlying distributions. Stopping decisions thus impose an endogenous censoring effect on histories, which in turn affects the beliefs of subsequent agents. The statistical bias interacts with data censoring, generating over-pessimism about the fundamentals and resulting in stopping too early in the long run. These asymptotic mistakes are driven by a positive-feedback loop between distorted beliefs and distorted behavior.

I have studied a particular behavioral bias (the gambler’s fallacy) in a natural environment where censoring happens (histories in optimal-stopping problems). The key mechanism I highlight, the interaction between data censoring and bias, applies more broadly and delivers different predictions in different contexts. For example, the same mechanism would lead to an over-estimation of $\mu_2$ if the agents believe in some $\gamma < 0$. I am leaving open the interaction of other kinds of behavioral learning with other censoring mechanisms to future work.

References


Appendix

A Proofs of Results in Sections 2, 3, and 4

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. If \( x_1 > 0 \) and \( x_2 < 0 \), then \( u_2(x_1, x_2) = q \cdot x_1 + (1 - q)x_2 < x_1 = u_1(x_1) \), and conversely \( x_1 < 0, x_2 > 0 \) imply \( u_2(x_1, x_2) > u_1(x_1) \). This shows Assumption 1(c) holds. It is clear that \( u_1, u_2 \) are continuous.

Also,

\[
|u_2(\bar{x}_1, x_2 + \bar{k})| \leq q(|\bar{x}_1| + |x_2 + \bar{k}|) + (1 - q)|x_2 + \bar{k}| \leq q|\bar{x}_1| + |\bar{k}| + |x_2|.
\]

Since the objective distribution satisfies \( \mathbb{E}(|X_2|) < \infty \), we have \( \mathbb{E}(|u_2(\bar{x}_1, x_2 + \bar{k})|) \leq q|\bar{x}_1| + |\bar{k}| + \mathbb{E}(|X_2|) < \infty \). This shows Assumption 1(d) holds.

A.2 Proofs of Proposition 1 and Lemma 2

The argument behind Proposition 1 consists of three lemmas (A.1, A.3, and A.4) that correspond to the three statements in the proposition. Along the way, I will also prove Lemma 2.

A.2.1 The Optimal Strategy Has a Cutoff Form

In the first part, I establish lemma A.1.

Lemma A.1. Under Assumption 1 and the feasible model \( \Psi(\mu_1, \mu_2; \gamma) \) for any \( \gamma > 0 \), there exists a cutoff \( C(\mu_1, \mu_2; \gamma) \), such that: (i) the agent strictly prefers stopping after any \( x_1 > C(\mu_1, \mu_2; \gamma) \); (ii) the agent is indifferent between continuing and stopping after \( x_1 = C(\mu_1, \mu_2; \gamma) \); (iii) the agent strictly prefers continuing after any \( x_1 < C(\mu_1, \mu_2; \gamma) \).

Suppose \( X_1 = x_1 \). Consider the payoff difference between accepting it and continuing under the feasible model \( \Psi(\mu_1, \mu_2; \gamma) \) for \( \gamma \geq 0 \):

\[
D(x_1; \mu_1, \mu_2, \gamma) := u_1(x_1) - \mathbb{E}_{\Psi(\mu_1, \mu_2; \gamma)}[u_2(x_1, X_2)|X_1 = x_1].
\]

I abbreviate this as \( D(x_1) \) when \( \Psi \) is fixed. Lemma A.2 summarizes some properties of \( D \).

(The proofs of some technical results stated in the Appendix, like Lemma A.2, appear in the Online Appendix.)

Lemma A.2. \( D \) is strictly increasing and continuous. If \( \gamma > 0 \), then there are \( x_1' < x_1'' \) so that \( D(x_1') < 0 < D(x_1'') \).
Lemma A.1 follows readily from Lemma A.2.

Proof. Applying Lemma A.2 and using the fact that $\gamma > 0$, $D$ changes sign and is strictly increasing and continuous. So, there exists a unique $c^* \in \mathbb{R}$ satisfying $D(c^*) = 0$. It is clear that the best stopping strategy under $\Psi$ is the cutoff strategy that stops after every $x_1 > c^*$ and continues after every $x_1 < c^*$. This establishes property (ii) of the optimal strategy. Properties (i) and (iii) follow from the fact that $D$ is strictly increasing. \hfill \Box

A.2.2 Cutoff Threshold Increasing in $\mu_2$

In the second part, I prove the lemma:

**Lemma A.3.** Under Assumption 1, for any $\mu_1 \in \mathbb{R}$ and $\gamma > 0$, the indifference threshold $C(\mu_1, \mu_2; \gamma)$ is strictly increasing in $\mu_2$.

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

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

$$u_1(C(\hat{\mu}_1, \hat{\mu}_2; \gamma)) = \mathbb{E}_{\tilde{X}_2 \sim f_2(\cdot | \hat{\mu}_2 - \gamma C(C(\hat{\mu}_1, \hat{\mu}_2; \gamma) - \hat{\mu}_1))] [u_2(C(\hat{\mu}_1, \hat{\mu}_2; \gamma), \tilde{X}_2)]$$

But if agent were to instead believe $(\hat{\mu}_1, \hat{\mu}_2)$ where $\hat{\mu}_2 > \hat{\mu}_2$, then the conditional distribution of $X_2$ given $X_1 = C(\hat{\mu}_1, \hat{\mu}_2; \gamma)$ would be $f_2(\cdot | \hat{\mu}_2 - \gamma (C(\hat{\mu}_1, \hat{\mu}_2; \gamma) - \hat{\mu}_1))$. We have

$$u_1(C(\hat{\mu}_1, \hat{\mu}_2; \gamma)) < \mathbb{E}_{\tilde{X}_2 \sim f_2(\cdot | \hat{\mu}_2 - \gamma C(C(\hat{\mu}_1, \hat{\mu}_2; \gamma) - \hat{\mu}_1))] [u_2(C(\hat{\mu}_1, \hat{\mu}_2; \gamma), \tilde{X}_2)]$$

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

A.2.3 Proof of Lemma 2

Proof. In fact, this lemma holds for any $\mu_1 \in \mathbb{R}$. I first prove this for Assumption 3(a).

For $\mu_2'' > \mu_2'$, write the corresponding optimal cutoffs as $c'' := C(\mu_1, \mu_2''; \gamma)$ and $c' := C(\mu_1, \mu_2'; \gamma)$. I show that $|c'' - c'| < \ell |\mu_2'' - \mu_2'|$.

Under the model $\Psi(\mu_1, \mu_2''; \gamma)$, the expected continuation payoff after $X_1 = c' + \ell (\mu_2'' - \mu_1')$ is

$$\mathbb{E}_{\tilde{X}_2 \sim f_2(\cdot | \mu_2'' - \gamma (c' + \ell (\mu_2'' - \mu_1'), \tilde{X}_2)] [u_2(c' + \ell (\mu_2'' - \mu_1'), \tilde{X}_2)]$$

$$= \mathbb{E}_{\tilde{X}_2 \sim f_2(\cdot | \mu_2' - (c' - \mu_1')) [u_2(c' + \ell (\mu_2' - \mu_1'), \tilde{X}_2 + (\mu_2'' - \mu_1') - \gamma \ell (\mu_2'' - \mu_1')]]$$

$$= \mathbb{E}_{\tilde{X}_2 \sim f_2(\cdot | \mu_2') [u_2(c' + \ell d, \tilde{X}_2 + (1 - \gamma \ell) d)]}$$

39
where we put \( d = |\mu''_2 - \mu'_2| > 0 \). From Assumption 3(a), for every \( x_2 \in \mathbb{R} \), \( u_2(c' + \ell d, x_2 + (1 - \gamma \ell)d) - u_2(c', x_2) < u_1(c' + \ell d) - u_1(c') \). This means

\[
\mathbb{E}_{\tilde{X}_2 \sim f_{2}(\mu'_2 - \gamma(c' - \mu_1))}[u_2(c' + \ell d, \tilde{X}_2 + (1 - \gamma \ell)d) - u_2(c', \tilde{X}_2)] < u_1(c' + \ell d) - u_1(c')
\]

\[
\mathbb{E}_{\tilde{X}_2 \sim f_{2}(\mu'_2 - \gamma(c' - \mu_1))}[u_2(c' + \ell d, \tilde{X}_2 + (1 - \gamma \ell)d)] - u_1(c' + \ell d) < \mathbb{E}_{\tilde{X}_2 \sim f_{2}(\mu'_2 - \gamma(c' - \mu_1))}[u_2(c', \tilde{X}_2)] - u_1(c').
\]

The cutoff \( c' \) satisfies the indifference condition, \( u_1(c') = \mathbb{E}_{\tilde{X}_2 \sim f_{2}(\mu'_2 - \gamma(c' - \mu_1))}[u_2(c', \tilde{X}_2)] \), so RHS is 0. But LHS is the difference between expected continuation payoff and stopping payoff at \( X_1 = c' + \ell (\mu'' - \mu'_1) \) under the model \( \Psi(\mu_1, \mu''; \gamma) \), which shows the agent strictly prefers stopping. This means \( c'' < c' + \ell (\mu'' - \mu'_1) \). But \( \mu_2 \mapsto C(\mu_1, \mu_2; \gamma) \) is increasing by Lemma A.3, which means \( c'' > c' \). Together, these two inequalities imply \( |c'' - c'| < \ell (\mu'' - \mu'_1) \).

Now, replace Assumption 3(a) with Assumption 3(b). By Lipschitz continuity of \( u_2 \), suppose \( |u_2(x'_2) - u_2(x''_2)| < L \cdot |x'_2 - x''_2| \) for some \( L > 0 \) and all \( x'_2, x''_2 \in \mathbb{R} \). Let \( \beta = \min(\frac{\epsilon/\gamma}{L \gamma + \epsilon}, \frac{1}{2 \gamma}) \) and put \( \ell = \frac{1}{\gamma} - \beta \), so \( 0 < \ell < \frac{1}{\gamma} \). Let any \( \Delta > 0 \) be given. Let \( c = C(\mu_1, \mu_2, \gamma) \). I show that \( C(\mu_1, \mu_2 + \Delta, \gamma) < c + \ell \Delta \).

We have \( u_1(c + \ell \Delta) - u_1(c) > (\frac{1}{\gamma} - \beta) \epsilon \Delta > (\frac{1}{\gamma} - \frac{\epsilon/\gamma}{L \gamma + \epsilon}) \epsilon \Delta \) and

\[
\mathbb{E}_{\Psi(\mu_1, \mu_2 + \Delta; \gamma)}[u_2(X_2) \mid X_1 = c + \ell \Delta] - \mathbb{E}_{\Psi(\mu_1, \mu_2; \gamma)}[u_2(X_2) \mid X_1 = c] \\
\leq L \cdot (\Delta - \ell \Delta \gamma) = \Delta L \gamma \beta \leq \Delta L \gamma \cdot \frac{\epsilon/\gamma}{L \gamma + \epsilon}
\]

By simple algebra, \((\frac{1}{\gamma} - \frac{\epsilon/\gamma}{L \gamma + \epsilon}) \epsilon \Delta = \Delta L \gamma \cdot \frac{\epsilon/\gamma}{L \gamma + \epsilon}\). Since \( u_1(c) = \mathbb{E}_{\Psi(\mu_1, \mu_2; \gamma)}[u_2(X_2) \mid X_1 = c] \), we conclude \( u_1(c + \ell \Delta) > \mathbb{E}_{\Psi(\mu_1, \mu_2 + \Delta; \gamma)}[u_2(X_2) \mid X_1 = c + \ell \Delta] \). By Lemma A.1, this implies \( c + \ell \Delta > (\mu_1, \mu_2 + \Delta, \gamma) \).

**A.2.4 Lipschitz Continuity with Constant 1/\( \gamma \)**

Now I prove the lemma:

**Lemma A.4.** Under Assumption 1, for every \( \gamma > 0 \) and \( \mu_1 \in \mathbb{R} \), \( \mu_2 \mapsto C(\mu_1, \mu_2; \gamma) \) is Lipschitz continuous with Lipschitz constant 1/\( \gamma \).

**Proof.** The proof of Lemma 2 also applies when \( \ell = \frac{1}{\gamma} \), which implies that when the inequality in Assumption 3(a) is satisfied with \( \ell = \frac{1}{\gamma} \), \( \mu_2 \mapsto C(\mu_1, \mu_2; \gamma) \) is Lipschitz continuous with Lipschitz constant 1/\( \gamma \). But this reduces the inequality to \( u_1(x_1 + \frac{1}{\gamma} d) - u_1(x_1) \geq u_2(x_1 + \frac{1}{\gamma} d, x_2) - u_2(x_1, x_2) \), which is true by Assumption 1(b).

**A.3 Proof of Claim 2**

**Proof.** For \( d > 0 \),

\[
u_1(x_1 + \frac{1}{1 + \gamma} d) - u_1(x_1) = \frac{1}{1 + \gamma} d
\]
while
\[ u_2(x_1 + \frac{1}{1 + \gamma} d, x_2 + (1 - \frac{\gamma}{1 + \gamma})d) - u_2(x_1, x_2) \]
\[ = u_2(x_1 + \frac{1}{1 + \gamma} d, x_2 + \frac{1}{1 + \gamma}d) - u_2(x_1, x_2) \]
\[ = q \max(x_1 + \frac{1}{1 + \gamma} d, x_2 + \frac{1}{1 + \gamma}d) + (1 - q)(x_2 + \frac{1}{1 + \gamma}d) - q \max(x_1, x_2) - (1 - q)x_2 \]
\[ = q - \frac{1}{1 + \gamma} d + (1 - q) \frac{1}{1 + \gamma} d = \frac{1}{1 + \gamma} d. \]

This shows that when \( \ell = \frac{1}{1 + \gamma} \), we have \( u_1(x_1 + \ell d) - u_1(x_1) = u_2(x_1 + \ell d, x_2 + (1 - \gamma \ell)d) - u_2(x_1, x_2) \) for every \( x_1, x_2 \in \mathbb{R}, d > 0 \). \( \square \)

**A.4 Proof of Proposition 2**

**A.4.1 Preliminary Definitions and Lemmas**

I first require some preliminary definitions and lemmas.

The first result says for any censoring threshold \( c \in \mathbb{R} \cup \{\infty\} \), KL divergence cannot be minimized at \((\mu_1, \mu_2)\) where \( \mu_1 \neq \mu_1^\bullet \).

**Lemma A.5.** For every \( \gamma > 0, c \in \mathbb{R} \cup \{\infty\}, \mu_1, \mu_2 \in \mathbb{R} \) with \( \mu_1 \neq \mu_1^\bullet \), we have
\[ D_{KL}(\mathcal{H}(c)\|\mathcal{H}(\Psi(\mu_1, \mu_2; \gamma); c)) > D_{KL}(\mathcal{H}^\bullet(c)\|\mathcal{H}(\Psi(\mu_1^\bullet, \mu_2 - \gamma(\mu_1^\bullet - \mu_1); \gamma); c)) \]

Lemma A.5 shows that solutions to the KL divergence minimization problem (if any exist) can only take the form \((\mu_1^\bullet, \mu_2)\) for some \( \mu_2 \in \mathbb{R} \). Thus motivated, we define
\[ L(\mu_2 \mid x_1) := \int_{-\infty}^{\infty} f_2(x_2 \mid \mu_2^\bullet) \ln[f_2(x_2 \mid \mu_2 - \gamma(x_1 - \mu_1^\bullet))]dx_2, \]
the expected log-likelihood of second-period data under the fundamentals \((\mu_1^\bullet, \mu_2)\) and after the realization \( X_1 = x_1 \). Also, put
\[ \bar{L}(\mu_2 \mid c) := \int_{-\infty}^{c} f_1(x_1 \mid \mu_1^\bullet) \cdot L(\mu_2 \mid x_1)dx_1 \]
for \( c \in \mathbb{R} \cup \{\infty\} \). Note \( \bar{L}(\mu_2 \mid c) \) is \( -D_{KL}(\mathcal{H}(c)\|\mathcal{H}(\Psi(\mu_1^\bullet, \mu_2; \gamma); c)) \) up to a constant not depending on \( \mu_2 \).

I establish some properties of \( L \) and \( \bar{L} \) that will be used in the remainder of the proof.

**Lemma A.6.** For all \( x_1, \mu_2 \in \mathbb{R} \), \( L(\mu_2 \mid x_1) = L(\mu_2^\bullet \mid x_1 - \frac{1}{\gamma}(\mu_2 - \mu_2^\bullet)) \), so \( \frac{\partial}{\partial \mu_2} L(\mu_2 \mid x_1) = -\frac{1}{\gamma} \frac{\partial}{\partial x_1} L(\mu_2^\bullet \mid x_1 - \frac{1}{\gamma}(\mu_2 - \mu_2^\bullet)) \).

**Proof.** Follows easily from the definition of \( L(\mu_2 \mid x_1) \). \( \square \)
Lemma A.7. For every \( \mu_2 \in \mathbb{R} \), \( L(\mu_2 \mid \cdot) \) is strictly concave. For every \( x_1 \in \mathbb{R} \), \( L(\cdot \mid x_1) \) is strictly concave. For every \( c \in \mathbb{R} \cup \{\infty\} \), \( \bar{L}(\cdot \mid c) \) is strictly concave. Finally, \( \frac{\partial^2}{\partial x_1 \partial \mu_2} L(\mu_2 \mid x_1) > 0 \).

Finally, I note a convenient property of strict log-concavity.

Lemma A.8. If \( f(x) > 0 \) is strictly log concave, then for any \( K > 0 \), \( x \mapsto \frac{f(x+K)}{f(x)} \) is strictly decreasing.

### A.4.2 Existence and Uniqueness of KL Divergence Minimizers

If \( \mu_2^* \in \mathbb{R} \) satisfies the FOC \( \frac{\partial}{\partial \mu_2} \bar{L}(\mu_2^* \mid c) = 0 \), then \( (\mu_1^*, \mu_2^*) \) is the unique KL divergence minimizer across all \( \mathbb{R}^2 \). This is because \( \mu_2^* \) satisfies the FOC in minimizing \( D_{KL}(\mathcal{H}(c) \| \mathcal{H}(\Psi(\mu_1^*, \mu_2^*; \gamma); c)) \) across \( \mu_2 \in \mathbb{R} \), a strictly convex objective function by the third statement in Lemma A.7. Furthermore, \( D_{KL}(\mathcal{H}(c) \| \mathcal{H}(\Psi(\mu_1^*, \mu_2^*; \gamma); c)) < D_{KL}(\mathcal{H}(c) \| \mathcal{H}(\Psi(\mu_1, \mu_2; \gamma); c)) \) for any \( \mu_1 \neq \mu_1^*, \mu_2 \in \mathbb{R} \) by Lemma A.5.

The next Lemma shows the FOC has a solution at \( \mu_2 = \mu_2^* \) when \( c = \infty \).

Lemma A.9. \( \frac{\partial}{\partial \mu_2} \bar{L}(\mu_2^* \mid \infty) = 0 \).

**Proof.** I first show \( L(\mu_2^* \mid x_1) \) is symmetric around \( x_1 = \mu_1^* \). Suppose \( x_1^h - \mu_1^* = \mu_1^* - x_1^i > 0 \). Then, \( L(\mu_2^* \mid x_1^i) \) is:

\[
\int_{-\infty}^\infty f_2(x_2 \mid \mu_2^*) \ln[f_2(x_2 \mid \mu_2^* - \gamma(x_1^h - \mu_1^*))]dx_2
\]

\[
= \int_{-\infty}^{\mu_2^*} f_2(x_2 \mid \mu_2^*) \ln[f_2(x_2 \mid \mu_2^* - \gamma(x_1^h - \mu_1^*))]dx_2 + \int_{\mu_2^*}^{\infty} f_2(x_2 \mid \mu_2^*) \ln[f_2(x_2 \mid \mu_2^* - \gamma(x_1^h - \mu_1^*))]dx_2
\]

\[
= \int_{-\infty}^{\mu_2^*} f_2(x_2 \mid \mu_2^*) \ln[f_2(x_2 + \gamma(x_1^h - \mu_1^*)) \mid \mu_2^*]dx_2 + \int_{\mu_2^*}^{\infty} f_2(x_2 \mid \mu_2^*) \ln[f_2(x_2 + \gamma(x_1^h - \mu_1^*)) \mid \mu_2^*]dx_2
\]

Using the symmetry of \( f_2(\cdot \mid \mu_2^*) \) around \( \mu_2^* \), let \( \tilde{g}_2(y) = f_2(\mu_2^* + y \mid \mu_2^*) = f_2(\mu_2^* - y \mid \mu_2^*) \) for \( y \geq 0 \). Let \( y_2 = \mu_2^* - x_2 \) in the first integral and \( y_2 = x_2 - \mu_2^* \) in the second integral in the sum. We then get

\[
L(\mu_2^* \mid x_1^h) = \int_0^\infty \tilde{g}_2(y_2) \left( \ln[f_2(\mu_2^* - y_2 + \gamma(x_1^h - \mu_1^*)) \mid \mu_2^*] + \ln[f_2(\mu_2^* + y_2 + \gamma(x_1^h - \mu_1^*)) \mid \mu_2^*] \right) dy_2.
\]

Analogous argument shows

\[
L(\mu_2^* \mid x_1^i) = \int_0^\infty \tilde{g}_2(y_2) \left( \ln[f_2(\mu_2^* - y_2 + \gamma(x_1^i - \mu_1^*)) \mid \mu_2^*] + \ln[f_2(\mu_2^* + y_2 + \gamma(x_1^i - \mu_1^*)) \mid \mu_2^*] \right) dy_2.
\]

For every \( y_2 \geq 0 \), we have \( \|\mu_2^* - y_2 + \gamma(x_1^h - \mu_1^*)\| - \|\mu_2^*\| = \|\mu_2^* + y_2 + \gamma(x_1^h - \mu_1^*) - \|\mu_2^*\| \) since \( x_1^i - \mu_1^* = -(x_1^h - \mu_1^*) \). As \( f_2(\cdot \mid \mu_2^*) \) is symmetric about \( \mu_2^* \), this shows

\[
\ln[f_2(\mu_2^* - y_2 + \gamma(x_1^i - \mu_1^*)) \mid \mu_2^*] = \ln[f_2(\mu_2^* + y_2 + \gamma(x_1^h - \mu_1^*)) \mid \mu_2^*].
\]
A similar symmetry argument shows that\( \ln[f_2(\mu_2^* + y_2 + \gamma(x_1^* - \mu_1^*)) | \mu_2^*] = \ln[f_2(\mu_2^* - y_2 + \gamma(x_1^* - \mu_1^*)) | \mu_2^*] \) for all \( y_2 \geq 0 \). Hence we conclude \( \bar{L}(\mu_2^* | x_1^*) = \bar{L}(\mu_2^* | x_1^*) \).

To finish the argument, apply the second statement in Lemma A.6 to get:

\[
\frac{\partial}{\partial \mu_2} \bar{L}(\mu_2^* | \infty) = \int_{-\infty}^{\infty} f_1(x_1 | \mu_1^*) \cdot (-\frac{1}{\gamma}) \cdot \frac{\partial L(\mu_2^* | x_1 - \frac{1}{\gamma}(\mu_2^* - \mu_2^*))}{\partial x_1} dx_1 \\
= -\frac{1}{\gamma} \int_{-\infty}^{\infty} f_1(x_1 | \mu_1^*) \frac{\partial L(\mu_2^* | x_1)}{\partial x_1} dx_1 \\
= -\frac{1}{\gamma} \left( \int_{-\infty}^{\mu_1^*} f_1(x_1 | \mu_1^*) \frac{\partial L(\mu_2^* | x_1)}{\partial x_1} dx_1 + \int_{\mu_1^*}^{\infty} f_1(x_1 | \mu_1^*) \frac{\partial L(\mu_2^* | x_1)}{\partial x_1} dx_1 \right)
\]

By symmetry of \( x \mapsto L(\mu_2^* | x_1) \) around \( x_1 = \mu_1^* \) just established, \( \frac{\partial L}{\partial x_1}(\mu_2^* | \mu_1^* - y) = -\frac{\partial L}{\partial x_1}(\mu_2^* | \mu_1^* + y) \) for all \( y \geq 0 \). At the same time, \( f_1(\mu_1^* - y | \mu_1^*) = f_1(\mu_1^* + y | \mu_1^*) \). Therefore the sum of the two integrals is 0.

In fact, the FOC also has a solution for any \( c \in \mathbb{R} \), as the next lemma shows.

**Lemma A.10.** For any \( \bar{c} \in \mathbb{R} \), there exists some \( \mu_2^* \in \mathbb{R} \) so that \( \frac{\partial}{\partial \mu_2} \bar{L}(\mu_2^* | \bar{c}) = 0 \).

### A.4.3 Monotonicity of \( \mu_2^*(c) \) in \( c \)

So far, I have shown that \((\mu_1^*(c), \mu_2^*(c)) \in \mathbb{R}^2 \) are well-defined for all \( c \in \mathbb{R} \cup \{ \infty \} \) and characterize the unique solution pair to the KL divergence minimization problem, with \( \mu_1^*(c) = \mu_1^* \) and \( \mu_2^*(\infty) = \mu_2^* \). To finish proving Proposition 2, it remains to show that \( \mu_2^*(c) \) is strictly increasing over \((-\infty, \infty] \).

**Lemma A.11.** Let \( c_1, c, c_2 \in \mathbb{R} \cup \{ \infty \} \) with \( c_1 < c < c_2 \). Then \( \frac{\partial}{\partial c} L(\mu_2^*(c) | c_1) < 0 \) and \( \frac{\partial}{\partial c} L(\mu_2^*(c) | c_2) > 0 \). As a result, whenever \( c', c'' \in (-\infty, \infty] \) with \( c' < c'' \), we have \( \mu_2^*(c') < \mu_2^*(c'') \).

**Proof.** First-order condition for \( \mu_2^*(c) \) implies that

\[
\frac{\partial}{\partial \mu_2} \bar{L}(\mu_2^*(c) | c) = 0 \Rightarrow -\frac{1}{\gamma} \int_{-\infty}^{c} f_1(x_1 | \mu_1^*) \cdot \frac{\partial}{\partial x_1} L(0 | x_1 - \frac{1}{\gamma}(\mu_2^*(c))) dx_1 = 0.
\]

From Lemma A.7, \( x_1 \mapsto \frac{\partial}{\partial x_1} L(0 | x_1 - \frac{1}{\gamma}(\mu_2^*(c))) \) is strictly decreasing. If at the rightmost point on integration interval, we have \( \frac{\partial}{\partial x_1} L(0 | c - \frac{1}{\gamma}(\mu_2^*(c))) \geq 0 \), then \( \frac{\partial}{\partial x_1} L(0 | x_1 - \frac{1}{\gamma}(\mu_2^*(c))) > 0 \) for all \( x_1 < c \). This would lead to \( \frac{\partial}{\partial \mu_2} \bar{L}(\mu_2^*(c) | c) \neq 0 \), a contradiction. Therefore \( \frac{\partial}{\partial x_1} L(0 | c - \frac{1}{\gamma}(\mu_2^*(c))) < 0 \).
For \( c_h > c \), we have that

\[
\frac{\partial}{\partial \mu_2} \tilde{L}(\mu_2^*(c) \mid c_h) = -\frac{1}{\gamma} \int_{-\infty}^{c_h} f_1(x_1 \mid \mu_1^*) \cdot \frac{\partial}{\partial x_1} L(0 \mid x_1 - \frac{1}{\gamma} \mu_2^*(c)) dx_1
\]

\[
= -\frac{1}{\gamma} \int_{-\infty}^{c} f_1(x_1 \mid \mu_1^*) \cdot \frac{\partial}{\partial x_1} L(0 \mid x_1 - \frac{1}{\gamma} \mu_2^*(c)) dx_1
\]

\[
+ (-\frac{1}{\gamma}) \int_{c}^{c_h} f_1(x_1 \mid \mu_1^*) \cdot \frac{\partial}{\partial x_1} L(0 \mid x_1 - \frac{1}{\gamma} \mu_2^*(c)) dx_1
\]

\[
= 0 + (-\frac{1}{\gamma}) \int_{c}^{c_h} f_1(x_1 \mid \mu_1^*) \cdot \frac{\partial}{\partial x_1} L(0 \mid x_1 - \frac{1}{\gamma} \mu_2^*(c)) dx_1.
\]

Since \( \frac{\partial}{\partial x_1} L(0 \mid c - \frac{1}{\gamma} \mu_2^*(c)) < 0 \), we also get \( \frac{\partial}{\partial x_1} L(0 \mid x_1 - \frac{1}{\gamma} \mu_2^*(c)) < 0 \) for any \( x_1 > c \) since \( L(0 \mid \cdot) \) is strictly concave by Lemma A.7. Therefore \( \frac{\partial}{\partial \mu_2} \tilde{L}(\mu_2^*(c) \mid c_h) > 0 \) since the density \( f_1(x_1 \mid \mu_1^*) \) is strictly positive.

For \( c_l < c \), we have that

\[
\frac{\partial}{\partial \mu_2} \tilde{L}(\mu_2^*(c) \mid c_l) = -\frac{1}{\gamma} \int_{-\infty}^{c_l} f_1(x_1 \mid \mu_1^*) \cdot \frac{\partial}{\partial x_1} L(0 \mid x_1 - \frac{1}{\gamma} \mu_2^*(c)) dx_1.
\]

If \( \frac{\partial}{\partial x_1} L(0 \mid x_1 - \frac{1}{\gamma} \mu_2^*(c)) \geq 0 \) for all \( x_1 \leq c_l \), then clearly this gives \( \frac{\partial}{\partial \mu_2} \tilde{L}(\mu_2^*(c) \mid c_l) < 0 \) as desired. Otherwise, write the integral as

\[
-\frac{1}{\gamma} \left[ \int_{-\infty}^{c} f_1(x_1 \mid \mu_1^*) \cdot \frac{\partial}{\partial x_1} L(0 \mid x_1 - \frac{1}{\gamma} \mu_2^*(c)) dx_1 - \int_{c_l}^{c} f_1(x_1 \mid \mu_1^*) \cdot \frac{\partial}{\partial x_1} L(0 \mid x_1 - \frac{1}{\gamma} \mu_2^*(c)) dx_1 \right],
\]

which simplifies to \( \frac{1}{\gamma} \int_{c_l}^{c} f_1(x_1 \mid \mu_1^*) \cdot \frac{\partial}{\partial x_1} L(0 \mid x_1 - \frac{1}{\gamma} \mu_2^*(c)) dx_1 \). If \( \frac{\partial}{\partial x_1} L(0 \mid x_1 - \frac{1}{\gamma} \mu_2^*(c)) \geq 0 \) for any \( x_1 \in [c_l, c] \), then we must also get \( \frac{\partial}{\partial x_1} L(0 \mid x_1 - \frac{1}{\gamma} \mu_2^*(c)) \geq 0 \) for all \( x_1 \leq c_l \), but this returns us to the case we have already considered. Thus \( \frac{\partial}{\partial x_1} L(0 \mid x_1 - \frac{1}{\gamma} \mu_2^*(c)) < 0 \) for all \( x_1 \in [c_l, c] \), and the integral is strictly negative, showing \( \frac{\partial}{\partial \mu_2} \tilde{L}(\mu_2^*(c) \mid c_l) < 0 \).

Finally, consider \( c', c'' \in (-\infty, \infty) \) with \( c' < c'' \). We must have \( \frac{\partial}{\partial \mu_2} \tilde{L}(\mu_2^*(c') \mid c'') > 0 \) by what we have established. But \( \tilde{L}(\cdot \mid c'') \) is strictly concave from Lemma A.7 and its FOC has a solution by Lemma A.10. So \( \mu_2^*(c'') > \mu_2^*(c') \). \( \square \)
A.5 Proof of the Expression for $\mu_2^*(c)$ in Example 2

Proof. Rewrite Definition 5 as

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

which is:

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

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:

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

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

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

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

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

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

\[
= (\mu_1 - \mu_1^*) + \gamma \frac{\partial \xi}{\partial \mu_2}
\]
By the first order conditions, at the minimum \((\mu_1^*, \mu_2^*)\), we must have \(\frac{\partial \ell}{\partial \mu_2}(\mu_1^*, \mu_2^*) = \frac{\partial \ell}{\partial \mu_1}(\mu_1^*, \mu_2^*) = 0 \Rightarrow \mu_1^* = \mu_1^\bullet\). So \(\mu_2^*\) satisfies \(\frac{\partial \ell}{\partial \mu_2}(\mu_1^*, \mu_2^*) = 0\), which by straightforward algebra shows \(\mu_2^*(c) = \mu_2^* - \gamma (\mu_1^* - \mathbb{E}[X_1 | X_1 \leq c])\). \(\square\)

### A.6 Proof of Lemma 1

**Proof.** Let \(c \in \mathbb{R}, K > 0\) be given. I show that \(|\mu_2^*(c+K) - \mu_2^*(c)| < \gamma K\). We have

\[
\bar{L}(\mu_2 | c) = \int_{-\infty}^{c} f_1(x_1 | \mu_1^\bullet)L(0 | x_1 - \frac{1}{\gamma \mu_2})dx_1
\]

and so

\[
\bar{L}(\mu_2 + \gamma K | c + K) = \int_{-\infty}^{c+K} f_1(x_1 | \mu_1^\bullet)L(0 | (x_1 - K) - \frac{1}{\gamma \mu_2})dx_1
\]

\[
= \int_{-\infty}^{c} f_1(x_1 + K | \mu_1^\bullet)L(0 | x_1 - \frac{1}{\gamma \mu_2})dx_1
\]

This implies

\[
\frac{\partial}{\partial \mu_2} \bar{L}(\mu_2 + \gamma K | c + K) = -\frac{1}{\gamma} \int_{-\infty}^{c} f_1(x_1 + K | \mu_1^\bullet) \frac{\partial}{\partial x_1} L(0 | x_1 - \frac{1}{\gamma \mu_2})dx_1.
\]

When \(\mu_2 = \mu_2^*(c)\), first-order condition implies that \(\frac{\partial}{\partial \mu_2} \bar{L}(\mu_2^*(c) | c) = 0\), that is

\[
\int_{-\infty}^{c} f_1(x_1 | \mu_1^\bullet) \frac{\partial}{\partial x_1} L(0 | x_1 - \frac{1}{\gamma \mu_2^*(c)})dx_1 = 0.
\]

We may write \(\frac{\partial}{\partial \mu_2} \bar{L}(\mu_2^*(c) + \gamma K | c + K)\) as

\[
-\frac{1}{\gamma} \int_{-\infty}^{c} \left( \frac{f_1(x_1 + K | \mu_1^\bullet)}{f_1(x_1 | \mu_1^\bullet)} \right) f_1(x_1 | \mu_1^\bullet) \frac{\partial}{\partial x_1} L(0 | x_1 - \frac{1}{\gamma \mu_2^*(c)})dx_1.
\]

The term \(x_1 \mapsto \frac{\partial}{\partial x_1} L(0 | x_1 - \frac{1}{\gamma \mu_2^*(c)})\) is positive for low values of \(x_1\) and negative for high values of \(x_1\), due to strict concavity of \(L(0 | \cdot)\) by Lemma A.7. Let \(r\) be such that \(\frac{\partial}{\partial x_1} L(0 | r - \frac{1}{\gamma \mu_2^*(c)}) = 0\). Then,

\[
\frac{\partial}{\partial \mu_2} \bar{L}(\mu_2^*(c) + \gamma K | c + K) < -\frac{1}{\gamma} \cdot \int_{-\infty}^{r} \left( \frac{f_1(x_1 + K | \mu_1^\bullet)}{f_1(x_1 | \mu_1^\bullet)} \right) f_1(x_1 | \mu_1^\bullet) \frac{\partial}{\partial x_1} L(0 | x_1 - \frac{1}{\gamma \mu_2^*(c)})dx_1
\]

\[
- \frac{1}{\gamma} \cdot \int_{r}^{c} \left( \frac{f_1(x_1 + K | \mu_1^\bullet)}{f_1(x_1 | \mu_1^\bullet)} \right) f_1(x_1 | \mu_1^\bullet) \frac{\partial}{\partial x_1} L(0 | x_1 - \frac{1}{\gamma \mu_2^*(c)})dx_1.
\]

By Lemma A.8, \(x \mapsto \frac{\partial}{\partial x_1} f_1(x_1 | \mu_1^\bullet)\) is strictly decreasing. So the weight \((\frac{f_1(x_1 + K | \mu_1^\bullet)}{f_1(x_1 | \mu_1^\bullet)})\) under-weighs the integrand on the interval \((-\infty, r)\), while the same weight over-weighs the integrand on
This amounts to an under-weighting of the positive part of the integrand and an over-weighting of the negative part, thus under-estimating the integral value. Accounting for the term \(-\frac{1}{\gamma}\) gives the inequality above.

FOC \(\frac{\partial}{\partial \mu_2} \tilde{L}(\mu^*_2(c) \mid c) = 0\) then implies RHS must be 0. So, \(\frac{\partial}{\partial \mu_2} \tilde{L}(\mu^*_2(c) + \gamma K \mid c + K) < 0\).

Since \(\tilde{L}(\cdot \mid c + K)\) is strictly concave by Lemma A.7, this implies \(\mu^*_2(c + K) < \mu^*_2(c) + \gamma K\).

Given that we must have \(\mu^*_2(c + K) > \mu^*_2(c)\) from Proposition 2, this shows \(|\mu^*_2(c + K) - \mu^*_2(c)| < \gamma K\).

\[\square\]

A.7 Proof of Proposition 4

Proof. Consider the map \(I\) as discussed in the text, \(I(\mu_2) := \mu_2(C(\mu_1^*, \mu_2; \gamma))\). If \(\hat{\mu}_2\) is a fixed point of \(I\), then there is a steady state with \(\mu^*_2 = \mu_1^*, \mu^*_2 = \hat{\mu}_2, c^* = C(\mu_1^*, \hat{\mu}_2; \gamma)\). So, existence of steady states follows from existence of fixed points of \(I\).

Conversely, suppose \((\mu^*_1, \mu^*_2, c^*)\) is a steady state. From Proposition 2, \(\mu^*_1 = \mu_1^*(c^*) = \mu_1^*\). From the definition of a steady state, \(\mu^*_2 = \mu_2(c^*) = c^* = C(\mu_1^*, \mu_2^*; \gamma)\). That is to say, \(\mu_2^* = \mu_2(C(\mu_1^*, \mu_2^*; \gamma))\), so \(\mu_2^*\) is a fixed point of \(I\). So, uniqueness of steady states follows from uniqueness of fixed points of \(I\).

Since \(\mu_2 \mapsto C(\mu_1^*, \mu_2; \gamma)\) is a contraction mapping with Lipschitz constant \(\ell < 1/\gamma\) by Lemma 2 and \(\mu_2^*(c)\) is a contraction mapping with Lipschitz constant \(\gamma\) by Lemma 1, their composition \(I\) is a contraction mapping with Lipschitz constant \(\ell \gamma < 1\). This proposition follows from properties of contraction mappings.

\[\square\]

A.8 Proof of Proposition 5

I will use the following lemma.

Lemma A.12. For any \(c \in \mathbb{R}\), \(\mu_2^*(c) - \gamma(c - \mu_1^*) < \mu_2^*\).

Here is the proof of Proposition 5.

Proof. Suppose \((\mu_1^*, \mu_2^*, c^*)\) is a steady state. If \(c^* = \infty\), then \(c^* < c^*\) trivially as \(c^* \in \mathbb{R}\).

Now suppose \(c^* \neq \infty\). By Proposition 1, agent is indifferent between stopping and continuing after \(X_1 = c^*\) under the feasible model \(\Psi(\mu_1^*, \mu_2^*; \gamma)\). This implies

\[
\begin{align*}
  u_1(c^*) &= \mathbb{E}_{\Psi(\mu_1^*, \mu_2^*; \gamma)}[u_2(c^*, X_2) | X_1 = c^*] \\
  &= \mathbb{E}_{X_2 \sim f_2(\cdot \mid \mu_2^* - \gamma(c^* - \mu_1^*))}[u_2(c^*, \tilde{X}_2)]
\end{align*}
\]

By the definition of steady state, \(\mu_2^* = \mu_2^*(c^*)\). By Lemma A.12, \(\mu_2^*(c^*) - \gamma(c^* - \mu_1^*) < \mu_2^*\).

Therefore, \(f_2(\cdot \mid \mu_2^* - \gamma(c^* - \mu_1^*))\) is first-order stochastically dominated by \(f_2(\cdot \mid \mu_2^*)\).

Since \(u_2\) is strictly increasing in its second argument by Assumption 1(a), we therefore have \(u_1(c^*) < \mathbb{E}_{\tilde{X}_2 \sim f_2(\cdot \mid \mu_2^*)}[u_2(c^*, \tilde{X}_2)]\). The LHS is the objective payoff of stopping at \(c^*\) while the
RHS is the objective expected payoff of continuing at \( c^\infty \). Since the best stopping strategy under the objective model \( \Psi^\bullet \) has the cutoff form, we must have \( c^\infty < c^\bullet \). \( \square \)

### A.9 Proof of Theorem 1′

The hypotheses of Theorem 1′ will be maintained throughout this section. I also also abbreviate \( f_1(· | \mu^1_1) =: g_1(·) \) and \( f_2(· | \mu^2_2) =: g_2(·) \). Finally, let \( \kappa_{g_2} \in \mathbb{R}_{>0} \) be such that \( \left| \frac{d^2}{dx^2} \ln(g_2(x)) \right| < \kappa_{g_2} \) for all \( x \in \mathbb{R} \).

#### A.9.1 Optimality of Cutoff Strategies

I first develop an extension of Lemma A.1. I show that for an agent who knows \( \mu^1_1 \) and has some belief over \( \mu_2 \) with supported bounded by \([\mu_2, \bar{\mu}_2]\), there exists a cutoff strategy that uniquely maximizes payoff across all cutoff strategies, so the “myopically optimal” cutoff strategy is well defined. Furthermore, this myopically optimal cutoff strategy also achieves weakly larger expected payoff compared to any arbitrary stopping strategy. So, restriction to cutoff strategies is without loss.

**Lemma A.13.** For an agent who knows \( \mu^1_1 \) and who holds some belief \( \nu \in \Delta([\mu_2, \bar{\mu}_2]) \) about second-period fundamental, there exists \( c^* \in \mathbb{R} \) such that: (i) the cutoff strategy \( S_{c^*} \) achieves weakly higher expected payoff than any other (not necessarily cutoff-based) stopping strategy \( S : \mathbb{R} \to \{\text{Stop, Continue}\} \); (ii) for any other \( c' \neq c^* \), \( S_{c^*} \) achieves strictly higher expected payoff than \( S_{c'} \).

#### A.9.2 The Log Likelihood Process

Next, I define the processes of data log likelihood (for a given fundamental). For each \( \mu_2 \in [\mu_2, \bar{\mu}_2] \), let \( \ell_t(\mu_2)(\omega) \) be the log likelihood that the true second-period fundamental is \( \mu_2 \) and histories \((\tilde{H}_s)_{s \leq t}(\omega)\) are generated by the end of round \( t \). It is given by

\[
\ell_t(\mu_2)(\omega) := \ln(m_0(\mu_2)) + \sum_{s=1}^{t} \ln(\text{lik}(\tilde{H}_s(\omega); \mu_2))
\]

where \( \text{lik}(x_1, \varnothing; \mu_2) := g_1(x_1) \) and \( \text{lik}(x_1, x_2; \mu_2) := g_1(x_1) \cdot f_2(x_2 | \mu_2 - \gamma(x_1 - \mu^1_1)) \).

I record a useful decomposition of \( \ell_t(\mu_2) \), the derivative of the log-likelihood process. Let

\[
\lambda(z) := \frac{d}{dz} \ln(g_2(z)) = \frac{g_2(z)}{g_2(\bar{\mu}_1)}.
\]

Define two stochastic processes:

\[
\varphi_s(\mu_2) := -\lambda(X_{2,s} - \mu_2 + \mu^*_2 + \gamma(X_{1,s} - \mu^*_1)) \cdot 1\{X_{1,s} \leq \tilde{C}_s\}
\]

\[
\kappa_{g_2} \in \mathbb{R}_{>0} \text{ is such that } \left| \frac{d^2}{dx^2} \ln(g_2(x)) \right| < \kappa_{g_2} \text{ for all } x \in \mathbb{R}.
\]

\text{In particular this implies if there exists at least one steady state, then } c^\bullet \neq -\infty.

\text{One can construct other stopping strategies with the same expected payoff by, for example, modifying the stopping decision of the optimal cutoff strategy at finitely many } x_1.
\[ \tilde{\varphi}_s(\mu_2) := \frac{\partial}{\partial \mu_2} \tilde{L}(\mu_2 \mid \tilde{C}_s) \]

Note that \( \tilde{\varphi}_s(\mu_2) \) is measurable with respect to \( \mathcal{F}_{s-1} \), since \((C_t)\) is a predictable process. Write \( \xi_s(\mu_2) := \varphi_s(\mu_2) - \tilde{\varphi}_s(\mu_2) \) and \( y_t(\mu_2) := \sum_{s=1}^t \xi_s(\mu_2) \). Write \( z_t(\mu_2) := \sum_{s=1}^t \tilde{\varphi}_s(\mu_2) \).

Lemma A.14. \( \ell_t'(\mu_2) = \frac{m_t'(\mu_2)}{m_0(\mu_2)} + y_t(\mu_2) + z_t(\mu_2) \)

Proof. We may expand \( \ell_t'(\mu_2) \) as

\[ \ln(m_0(\mu_2)) + \sum_{s=1}^t \ln(g_1(X_{1,s})) + \sum_{s=1}^t \ln(f_2(X_{2,s} \mid \mu_2 - \gamma(X_{1,s} - \mu_1^*)) \cdot 1\{X_{1,s} \leq \tilde{C}_s\}. \]

The derivative of the first term is \( \frac{m_t'(\mu_2)}{m_0(\mu_2)} \). The second term does not depend on \( \mu_2 \). In the third term, we use the fact that \( f_2(\cdot \mid \tau) \) are translations of each other and that \( g_2(\cdot) = f_2(\cdot \mid \mu_2^*) \) to write:

\[ f_2(X_{2,s} \mid \mu_2 - \gamma(X_{1,s} - \mu_1^*)) = g_2(X_{2,s} - \mu_2 + \mu_2^* + \gamma(X_{1,s} - \mu_1^*)). \]

This shows that the derivative of each summand in the third term with respect to \( \mu_2 \) is

\[ \frac{g_2'(X_{2,s} - \mu_2 + \mu_2^* + \gamma(X_{1,s} - \mu_1^*))}{g_2(X_{2,s} - \mu_2 + \mu_2^* + \gamma(X_{1,s} - \mu_1^*))} \cdot 1\{X_{1,s} \leq \tilde{C}_s\} = \varphi_s(\mu_2). \]

So in sum, \( \ell_t'(\mu_2) = \frac{m_t'(\mu_2)}{m_0(\mu_2)} + \sum_{s=1}^t \varphi_s(\mu_2) \). The lemma then follows from simple rearrangements. \( \square \)

Now I derive two results about the \( \xi_t(\mu_2) \) processes for different values of \( \mu_2 \).

Lemma A.15. There exists \( \kappa_\xi < \infty \) so that for every \( \mu_2 \in [\mu_2, \tilde{\mu_2}] \) and for every \( t \geq 1, \omega \in \Omega, \mathbb{E}[\xi_t^2(\mu_2) \mid \mathcal{F}_{t-1}](\omega) \leq \kappa_\xi. \)

The proof can be found in the Online Appendix.

Lemma A.16. For every \( t \geq 1, \mu_2 \in [\mu_2, \tilde{\mu_2}] \) and \( \omega \in \Omega, |\xi_t'(\mu_2)(\omega)| \leq 2\kappa_{g_2}. \)

Proof. In the proof of Lemma A.15, we established \( \mathbb{E}[\varphi_t(\mu_2) \mid \mathcal{F}_{t-1}] = \tilde{\varphi}_t(\mu_2) \). So, we have \( |\xi_t'(\mu_2)(\omega)| \leq |\varphi_t'(\mu_2)(\omega)| + |\mathbb{E}[\varphi_t'(\mu_2) \mid \mathcal{F}_{t-1}](\omega)| \). We have

\[ \varphi_t'(\mu_2) = \lambda'(X_{2,s} - \mu_2 + \mu_2^* + \gamma(X_{1,s} - \mu_1^*)) \cdot 1\{X_{1,s} \leq \tilde{C}_s\}, \]

with \( |\lambda'(z)| \leq \kappa_{g_2} \) for all \( z \in \mathbb{R} \). This shows \( |\varphi_t'(\mu_2)(\omega)| \leq \kappa_{g_2} \) for all \( \omega \), and similarly \( |\mathbb{E}[\varphi_t'(\mu_2) \mid \mathcal{F}_{t-1}](\omega)| \leq \kappa_{g_2} \) for all \( \omega \). \( \square \)
A.9.3 Heidhues, Koszegi, and Strack (2018)’s Law of Large Numbers

I use a statistical result from Heidhues, Koszegi, and Strack (2018) to show that the $y_t$ term in the decomposition of $\ell'_t$, almost surely converges to 0 in the long run, and furthermore this convergence is uniform on $[\mu_2, \bar{\mu}_2]$. This lets me focus on summands of the form $\bar{\varphi}_s(\mu_2)$, which can be interpreted as the expected contribution to the log likelihood derivative from round $s$ data. This lends tractability to the problem as $\bar{\varphi}_s(\mu_2)$ only depends on $\bar{C}_s$, but not on $X_{1,s}$ or $X_{2,s}$.

**Lemma A.17.** For every $\mu_2 \in [\mu_2, \bar{\mu}_2]$, $\lim_{t \to \infty} \left| \frac{y_t(\mu_2)}{t} \right| = 0$ almost surely.

**Proof.** Heidhues, Koszegi, and Strack (2018)’s Proposition 10 shows that if $(y_t)$ is a martingale such that there exists some constant $v \geq 0$ satisfying $|y_t| \leq vt$ almost surely, where $|y_t|$ is the quadratic variation of $(y_t)$, then almost surely $\lim_{t \to \infty} \frac{y_t}{t} = 0$.

Consider the process $y_t(\mu_2)$ for a fixed $\mu_2 \in [\mu_2, \bar{\mu}_2]$. By definition $y_t = \sum_{s=1}^{t} \varphi_s(\mu_2) - \bar{\varphi}_s(\mu_2)$. As established in the proof of Lemma A.15, for every $s$, $\bar{\varphi}_s(\mu_2) = E[\varphi_s(\mu_2)|F_{s-1}]$. So for $t' < t$,

$$E[y_t(\mu_2)|F_{t'}] = \sum_{s=1}^{t'} \varphi_s(\mu_2) - \bar{\varphi}_s(\mu_2) + E\left[ \sum_{s=t'+1}^{t} \varphi_s(\mu_2) - \bar{\varphi}_s(\mu_2) | F_{t'} \right]$$

$$= \sum_{s=1}^{t'} \varphi_s(\mu_2) - \bar{\varphi}_s(\mu_2) + \sum_{s=t'+1}^{t} E[E[\varphi_s(\mu_2) - \bar{\varphi}_s(\mu_2)|F_{s-1}] | F_{t'}]$$

$$= \sum_{s=1}^{t'} \varphi_s(\mu_2) - \bar{\varphi}_s(\mu_2) + 0 = y_{t'}(\mu_2).$$

This shows $(y_t(\mu_2))$ is a martingale. Also,

$$[y(\mu_2)]_t = \sum_{s=1}^{t-1} E[(y_s(\mu_2) - y_{s-1}(\mu_2))^2 | F_{s-1}]$$

$$= \sum_{s=1}^{t-1} E[\xi_s^2(\mu_2)|F_{s-1}] \leq \kappa \xi \cdot t$$


**Lemma A.18.** $\lim_{t \to \infty} \sup_{\mu_2 \in [\mu_2, \bar{\mu}_2]} \left| \frac{y_t(\mu_2)}{t} \right| = 0$ almost surely.

**Proof.** From the proof of Lemma 11 in Heidhues, Koszegi, and Strack (2018), it suffices to find a sequence of random variables $B_t$ such that $\sup_{\mu_2 \in [\mu_2, \bar{\mu}_2]} |\xi'_t(\mu_2)| \leq B_t$ almost surely, $\sup_{t \geq 1} \frac{1}{t} \sum_{s=1}^{t} E[B_s] < \infty$, and $\lim_{t \to \infty} \frac{1}{t} \sum_{s=1}^{t} (B_s - E[B_s]) = 0$. But Lemma A.16 establishes the constant random variable $B_t = 2\kappa g_2$ as a bound on $\xi'_t(\mu_2)$ for every $t, \mu_2, \omega$, which satisfies these requirements. □
A.9.4 Bounds on Asymptotic Beliefs and Asymptotic Cutoffs

For each $t$, let $\tilde{M}_t$ be the (random) posterior belief induced by the (random) posterior density $\tilde{m}_t$ after updating prior $m_0$ using $t$ rounds of histories.

**Lemma A.19.** For $c' \geq C(\mu^*_1, \mu^*_2; \gamma)$, if almost surely $\lim_{t \to \infty} \tilde{C}_t \geq c'$, then almost surely $\lim_{t \to \infty} \tilde{M}_t( [\mu_2, \mu^*_2(c')] ) = 0$. Also, for $c^h \leq C(\mu^*_1, \tilde{\mu}_2; \gamma)$, if almost surely $\limsup_{t \to \infty} \tilde{C}_t \leq c^h$, then almost surely $\lim_{t \to \infty} \tilde{M}_t( (\mu^*_2(c^h), \tilde{\mu}_2] ) = 0$.

**Proof.** I first show that for all $\epsilon > 0$, there exists $\delta > 0$ such that almost surely,

$$\liminf_{t \to \infty} \inf_{\mu_2 \in [\mu_2, \mu^*_2(c')] - \epsilon} \frac{\ell_t(\mu_2)}{t} \geq \delta.$$  

From Lemma A.14, we may rewrite LHS as

$$\liminf_{t \to \infty} \inf_{\mu_2 \in [\mu_2, \mu^*_2(c')] - \epsilon} \left[ \frac{1}{t} \frac{m'_0(\mu_2)}{m_0(\mu_2)} + \frac{y_t(\mu_2)}{t} + \frac{z_t(\mu_2)}{t} \right],$$

which is no smaller than taking the inf separately across the three terms in the bracket,

$$\liminf_{t \to \infty} \inf_{\mu_2 \in [\mu_2, \mu^*_2(c')] - \epsilon} \frac{1}{t} \frac{m'_0(\mu_2)}{m_0(\mu_2)} + \liminf_{t \to \infty} \inf_{\mu_2 \in [\mu_2, \mu^*_2(c')] - \epsilon} \frac{y_t(\mu_2)}{t} + \liminf_{t \to \infty} \inf_{\mu_2 \in [\mu_2, \mu^*_2(c')] - \epsilon} \frac{z_t(\mu_2)}{t}.$$

Since $m'_0$ is continuous and $m_0$ is strictly positive (and continuous) on $[\mu_2, \tilde{\mu}_2]$ by the hypotheses of Theorem 1', $m'_0/m_0$ is bounded on $[\mu_2, \tilde{\mu}_2]$, so we in fact have

$$\lim_{t \to \infty} \inf_{\mu_2 \in [\mu_2, \mu^*_2(c')] - \epsilon} \frac{1}{t} \frac{m'_0(\mu_2)}{m_0(\mu_2)} = 0.$$

To deal with the second term,

$$\liminf_{t \to \infty} \inf_{\mu_2 \in [\mu_2, \mu^*_2(c')] - \epsilon} \frac{y_t(\mu_2)}{t} \geq \liminf_{t \to \infty} \inf_{\mu_2 \in [\mu_2, \tilde{\mu}_2]} \frac{y_t(\mu_2)}{t} = -\liminf_{t \to \infty} \sup_{\mu_2 \in [\mu_2, \tilde{\mu}_2]} -\frac{y_t(\mu_2)}{t}.$$

Lemma A.18 gives $\lim_{t \to \infty} \sup_{\mu_2 \in [\mu_2, \tilde{\mu}_2]} -\frac{y_t(\mu_2)}{t} = 0$ almost surely, so this second term is non-negative almost surely.

It suffices then to find $\delta > 0$ and show $\liminf_{t \to \infty} \inf_{\mu_2 \in [\mu_2, \mu^*_2(c')] - \epsilon} \frac{z_t(\mu_2)}{t} \geq \delta$ almost surely. Put $\delta := \frac{\partial}{\partial \mu_2} \tilde{L}(\mu^*_2(c') - \epsilon | c')$ and I will show $\tilde{\varphi}_s(\mu_2)(\omega) \geq \delta$ whenever $\tilde{C}_s(\omega) \geq c'$ and $\mu_2 \leq \mu^*_2(c' - \epsilon)$. To see this, note that when $\tilde{C}_s(\omega) = c \in \mathbb{R}$, $\tilde{\varphi}_s(\mu_2)(\omega) = \frac{\partial}{\partial \mu_2} \tilde{L}(\mu_2 | c)$ and $\tilde{L}(\cdot | c)$ is strictly concave in its first argument by Lemma A.7. Therefore, if $\tilde{\varphi}_s(\mu^*_2(c') - \epsilon)(\omega) \geq \delta$, then we also get $\tilde{\varphi}_s(\mu_2)(\omega) \geq \delta$ for any $\mu_2 \leq \mu^*_2(c' - \epsilon)$. So it suffices to show $\frac{\partial}{\partial \mu_2} \tilde{L}(\mu^*_2(c') - \epsilon | c) \geq \delta$ whenever $c \geq c'$. 

51
We have
\[
\frac{\partial}{\partial \mu_2} \tilde{L}(\mu_2 \mid c^l) = \int_{-\infty}^{c^l} g_1(x_1) \cdot \int_{-\infty}^{\infty} (-1) \cdot g_2(x_2) \cdot \lambda(x_2 - \mu_2 + \mu_2^\ast + \gamma(x_1 - \mu_1)) dx_2 dx_1.
\]
First-order condition implies that \( \frac{\partial}{\partial \mu_2} \tilde{L}(\mu_2^ast(c^l) \mid c^l) = 0 \). Since \( \lambda \) is strictly decreasing, this implies \( \delta = \frac{\partial}{\partial \mu_2} \tilde{L}(\mu_2^ast(c^l) - \epsilon \mid c^l) > 0 \). Also, again using \( \lambda \) strictly decreasing, the inner integrand is strictly increasing in \( x_1 \). Thus, \( \frac{\partial}{\partial \mu_2} \tilde{L}(\mu_2^ast(c^l) - \epsilon \mid c^l) > 0 \) implies
\[
\int_{-\infty}^{\infty} (-1) \cdot g_2(x_2) \cdot \lambda(x_2 - (\mu_2^ast(c^l) - \epsilon) + \mu_2^\ast + \gamma(c - \mu_1^ast)) dx_2 > 0
\]
for all \( c \geq c^l \). This then shows \( \frac{\partial}{\partial \mu_2} \tilde{L}(\mu_2^ast(c^l) - \epsilon \mid c) > \frac{\partial}{\partial \mu_2} \tilde{L}(\mu_2^ast(c^l) - \epsilon \mid c^l) \) for any \( c > c^l \).

Having shown that \( \bar{\varphi}_s(\mu_2)(\omega) \geq \delta \) for all \( \mu_2 \in [\mu_2^l, \mu_2^h(c^l) - \epsilon] \) whenever \( \bar{C}_s(\omega) \geq c^l \), this shows along any \( \omega \) such that \( \liminf_{t \to \infty} \bar{C}_t \geq c^l \), we also have \( \liminf_{t \to \infty} \inf_{\mu_2 \in [\mu_2^l, \mu_2^h(c^l) - \epsilon]} \bar{\varphi}_s(\mu_2) \geq \delta \), and thus
\[
\liminf_{t \to \infty} \inf_{\mu_2 \in [\mu_2^l, \mu_2^h(c^l) - \epsilon]} \frac{z_t(\mu_2)}{t} = \liminf_{t \to \infty} \inf_{\mu_2 \in [\mu_2^l, \mu_2^h(c^l) - \epsilon]} \frac{1}{t} \sum_{s=1}^{t} \bar{\varphi}_s(\mu_2) \geq \delta.
\]
From here, it is a standard exercise to establish that \( \lim_{t \to \infty} \bar{M}_t(\mu_2^l, \mu_2^h(c^l) - \epsilon) = 0 \) almost surely. Since the choice of \( \epsilon > 0 \) is arbitrary, this establishes the first part of the lemma.

The proof of the second part of the statement is exactly symmetric, except we will show that \( \limsup_{t \to \infty} \sup_{\mu_2 \in [\mu_2^l, \mu_2^h(c^h) + \epsilon, \mu_2]} \frac{z_t(\mu_2)}{t} \leq -\delta \). where
\[
-\delta = \max \left( \frac{\partial}{\partial \mu_2} \tilde{L}(\mu_2^ast(c^h) + \epsilon \mid c^h), \frac{\partial}{\partial \mu_2} \tilde{L}(\mu_2^ast(c^h) + \epsilon \mid C(\mu^ast_1, \mu^h_2; \gamma)) \right) < 0.
\]

\textbf{Lemma A.20.} For \( \mu_2^l \leq \mu_2^l < \mu_2^h \leq \tilde{\mu}_2 \), if \( \lim_{t \to \infty} \bar{M}_t(\mu_2^l, \mu_2^h) = 1 \) almost surely, then \( \liminf_{t \to \infty} \bar{C}_t \geq C(\mu_1^ast, \mu_2^l; \gamma) \) and \( \limsup_{t \to \infty} \bar{C}_t \leq C(\mu_1^ast, \mu_2^h; \gamma) \) almost surely.

\textbf{Proof.} I show \( \liminf_{t \to \infty} \bar{C}_t \geq C(\mu_1^ast, \mu_2^l; \gamma) \) almost surely . The argument establishing \( \limsup_{t \to \infty} \bar{C}_t \leq C(\mu_1^ast, \mu_2^h; \gamma) \) is symmetric.

Let \( c^l = C(\mu_1^ast, \mu_2^l; \gamma), c = C(\mu_1^ast, \mu_2^ast; \gamma), c = C(\mu_1^ast, \tilde{\mu}_2; \gamma) \). Fix some \( \epsilon > 0 \). Since \( c^l \mapsto U(c; \mu_1^ast, \mu_2^ast) \) is single peaked for every \( \mu_2 \), and since \( c^l \leq C(\mu_1^ast, \mu_2; \gamma) \) for all \( \mu_2 \in [\mu_2^l, \mu_2^h] \), we get \( U(c^l; \mu_1^ast, \mu_2^ast) - U(c^l - \epsilon; \mu_1^ast, \mu_2^ast) > 0 \) for every \( \mu_2 \in [\mu_2^l, \mu_2^h] \). As \( \mu_2 \mapsto \left( U(c^l; \mu_1^ast, \mu_2^ast) - U(c^l - \epsilon; \mu_1^ast, \mu_2^ast) \right) \) is continuous, there exists some \( \kappa^* > 0 \) so that \( U(c^l; \mu_1^ast, \mu_2^ast) - U(c^l - \epsilon; \mu_1^ast, \mu_2^ast) > \kappa^* \) for all \( \mu_2 \in [\mu_2^l, \mu_2^h] \). In particular, if \( \nu \in \Delta((\mu_2^l, \mu_2^h)) \) is a belief over second-period fundamental supported on \( [\mu_2^l, \mu_2^h] \), then \( \int U(c^l; \mu_1^ast, \mu_2^ast) - U(c^l - \epsilon; \mu_1^ast, \mu_2^ast) d\nu(\mu_2) > \kappa^* \).

Now , let \( \kappa := \sup_{c \in [c^l, c^h]} \sup_{\mu_2 \in [\mu_2^l, \mu_2]} U(c; \mu_1^ast, \mu_2), \kappa := \inf_{c \in [c^l, c^h]} \inf_{\mu_2 \in [\mu_2^l, \mu_2]} U(c; \mu_1^ast, \mu_2). \)
Find $p \in (0, 1)$ so that $p \kappa^* - (1-p)(\bar{\kappa} - \kappa) = 0$. At any belief $\hat{\nu} \in \Delta([\mu_2, \bar{\mu}_2])$ that assigns more than probability $p$ to the subinterval $[\mu_1', \mu_2']$, the optimal cutoff is larger than $\mu_1' - \epsilon$. To see this, take any $c \leq \mu_1' - \epsilon$ and I will show $c$ is suboptimal. If $\hat{\nu} < c$, then it is suboptimal after any belief on $[\mu_2, \bar{\mu}_2]$. If $c \leq \hat{\nu} \leq \mu_1' - \epsilon$, I show that $\int U(c; \mu_1', \mu_2) - U(\hat{\nu}; \mu_1', \mu_2) d\hat{\nu}(\mu_2) > 0$.

To see this, we may decompose $\hat{\nu}$ as the mixture of a probability measure $\nu$ on $[\mu_1', \mu_2']$ and another probability measure $\nu^c$ on $[\mu_2, \bar{\mu}_2] \setminus [\mu_1', \mu_2']$. Let $\hat{p} > p$ be the probability that $\nu$ assigns to $[\mu_1', \mu_2']$. The above integral is equal to:

$$\hat{p} \int_{\mu_2 \in [\mu_1', \mu_2']} U(c; \mu_1', \mu_2) - U(c; \mu_1', \mu_2) d\nu(\mu_2) + (1 - \hat{p}) \int_{\mu_2 \in [\mu_2, \bar{\mu}_2] \setminus [\mu_1', \mu_2']} U(c; \mu_1', \mu_2) - U(\hat{\nu}; \mu_1', \mu_2) d\nu^c(\mu_2)$$

Since $c$ is to the left of the optimal cutoff for all $\mu_2 \in [\mu_1', \mu_2']$ and $c \leq \mu_1' - \epsilon$, then $U(c; \mu_1', \mu_2) < U(c - \epsilon; \mu_1', \mu_2)$ for all $\mu_2 \in [\mu_1', \mu_2']$. The first summand is no less than

$$\hat{p} \int_{\mu_2 \in [\mu_1', \mu_2']} U(c; \mu_1', \mu_2) - U(c - \epsilon; \mu_1', \mu_2) d\nu(\mu_2) > \hat{p} \kappa^*.$$

Also, the integrand in the second summand is no smaller than $-(\bar{\kappa} - \kappa)$, therefore $\int U(c - \epsilon; \mu_1', \mu_2) - U(\hat{\nu}; \mu_1', \mu_2) d\nu^c(\mu_2) \geq \hat{p} \kappa^* - (1 - \hat{p})(\bar{\kappa} - \kappa)$. Since $\hat{p} > p$, we get $\hat{p} \kappa^* - (1 - \hat{p})(\bar{\kappa} - \kappa) > 0$ as desired.

Along any sample path $\omega$ where $\lim_{t \to \infty} \bar{M}_t([\mu_1', \mu_2'])[\omega] = 1$, eventually $\bar{M}_t([\mu_1', \mu_2'])[\omega] > p$ for all large enough $t$, meaning $\lim_{t \to \infty} \bar{C}_t(\omega) \geq \mu_1' - \epsilon$. This shows $\lim_{t \to \infty} \bar{C}_t \geq C(\mu_1', \mu_2'; \gamma) - \epsilon$ almost surely. Since the choice of $\epsilon > 0$ was arbitrary, we in fact conclude $\lim_{t \to \infty} \bar{C}_t \geq C(\mu_1', \mu_2'; \gamma)$ almost surely.

\section{A.9.5 The Contraction Map}

I now combine the results established so far to prove Theorem 1'.

Proof. Let $\mu_{2,[1]} := \mu_2$, $\mu_{2,[1]} := \bar{\mu}_2$. For $k = 2, 3, \ldots$, iteratively define $\mu_{2,[k]} := \mathcal{I}(\mu_2, [\mu_{2,[k-1]}]; \gamma)$ and $\mu_{2,[k]} := \mathcal{I}(\mu_{2,[k-1]}; \gamma)$.

From Lemma A.20, if $\lim_{t \to \infty} \bar{M}_t([\mu_{2,[k]}, \mu_{2,[k]}]) = 1$ almost surely, then $\lim_{t \to \infty} \bar{C}_t \geq C(\mu_1', \mu_2'; \gamma)$ and $\lim_{t \to \infty} \bar{C}_t \leq C(\mu_{2,[k]}, \mu_{2,[k]}; \gamma)$ almost surely. But using these conclusions in Lemma A.19, we further deduce that $\lim_{t \to \infty} \bar{M}_t([\mu_{2,[k]}(C(\mu_1', \mu_{2,[k]}; \gamma)), \mu_{2}(C(\mu_1', \mu_{2,[k]}; \gamma))]) = 1$ almost surely, that is to say $\lim_{t \to \infty} \bar{M}_t([\mu_{2,[k+1]}, \mu_{2,[k+1]}]) = 1$ almost surely.

As shown in the proof of Proposition 4, under Assumptions 1, 2, and 3, $\mu_2 \mapsto \mathcal{I}(\mu_2; \gamma)$ is a contraction mapping. Since $\mu_2 < \mu_2^\infty$ and $\bar{\mu}_2 > \mu_2^\infty$, $(\mu_{2,[k]})_{k \geq 1}$ is a sequence whose limit is $\mu_2^\infty$, and $(\mu_{2,[k]})_{k \geq 1}$ is a sequence whose limit is $\mu_2^\infty$. Thus, agent’s posterior converges in $L^1$ to $\mu_2^\infty$ almost surely (since the support of the prior is bounded).

In addition, $\mu_2 \mapsto C(\mu_1', \mu_{2,[k]}; \gamma) = \kappa^\infty$ and $\lim_{k \to \infty} C(\mu_1', \mu_{2,[k]}; \gamma) = \kappa^\infty$. This
means $\lim_{t \to \infty} \tilde{C}_t = c^\infty$ almost surely.

### A.10 Proof of Theorem 2

I require a lemma that shows beliefs and cutoffs are monotonic in the auxiliary environment.

**Lemma A.21.** Suppose Assumptions 1 and 2 hold. Starting from any initial condition and any $m_0$, cutoffs $(c^A_{[t]})_{t \geq 1}$ and beliefs $(\mu^A_{[t]})_{t \geq 1}$ in the auxiliary environment form monotonic sequences across generations. Also, $\lim_{t \to \infty} \mu^A_{[t]} = \mu^\infty_2$ where $\mu^\infty_2$ is the unique fixed point of $\mathcal{L}(\cdot; \gamma)$ and $\lim_{t \to \infty} c^A_{[t]} = C(\mu^*_1, \mu^\infty_2; \gamma)$.

Now I turn to the proof of Theorem 2.

**Proof.** For the first step of the proof, suppose Assumptions 1 and 2 hold.

**Step 1:** If $c_{[1]} > c_{[0]}$, then $(\mu^A_{[2]}, c_{[t]})_{t \geq 1}$ and $(c_{[t]})_{t \geq 0}$ are two increasing sequences, whereas $c_{[1]} \leq c_{[0]}$ implies $(\mu^A_{[2]})_{t \geq 1}$ and $(c_{[t]})_{t \geq 0}$ are two decreasing sequences.

By simple algebra, the problem of generation $t + 1$ amounts to maximizing the sum of $\tilde{L}(\cdot | c_{[0]}), \ldots, \tilde{L}(\cdot | c_{[t]})$. For $c_0, \ldots, c_t \in \mathbb{R}$, denote $\mu^A_{[t]}(c_0, \ldots, c_t) := \min_{\mu^A_{[t]} \in \mathbb{R}} \sum_{s=0}^t \tilde{L}(\mu^A_{[t]} | c_s)$.

Suppose $c_{[1]} > c_{[0]}$. Then $\mu^A_{[1]} = \mu^*_2(c_{[0]})$, but by Lemma A.11, $\frac{\partial}{\partial \mu^A_{[t]}} \tilde{L}(\mu^A_{[t]} | c_{[1]}) > 0$. Then, since $\tilde{L}(\cdot | c_{[0]}) + \tilde{L}(\cdot | c_{[1]})$ is strictly concave and since $\frac{\partial}{\partial \mu^A_{[t]}} \tilde{L}(\mu^A_{[t]} | c_{[0]}) + \frac{\partial}{\partial \mu^A_{[t]}} \tilde{L}(\mu^A_{[t]} | c_{[1]}) = \frac{\partial}{\partial \mu^A_{[t]}} \tilde{L}(\mu^A_{[t]} | c_{[1]}) > 0$, we must have $\mu^A_{[2]} = \mu^*_2(c_{[0]}, c_{[1]}) > \mu^A_{[1]}$. This also shows that, since $C$ is strictly increasing, $c_{[2]} > c_{[1]}$.

Assume we have established that $c_{[0]} < c_{[1]} < \ldots < c_{[t]}$ and $\mu^A_{[2]} < \ldots < \mu^A_{[t]}$ for some $t \geq 2$. By FOC of inference in generation $t$, $\sum_{s=0}^{t-1} \frac{\partial}{\partial \mu^A_{[t]}} \tilde{L}(\mu^A_{[t]} | c_{[s]}) = 0$. If we had $\frac{\partial}{\partial \mu^A_{[t]}} \tilde{L}(\mu^A_{[t]} | c_{[r-1]}) < 0$, then by single-peaked nature of $\tilde{L}(\cdot | c_{[r-1]})$, $\mu^A_{[t]} > \mu^*_2(c_{[r-1]})$. Since $c_{[0]} < c_{[1]} < \ldots < c_{[t-1]}$ implies $\mu^*_2(c_{[0]}) < \ldots < \mu^*_2(c_{[t-1]})$ by Proposition 2, we must also have $\mu^A_{[t]} > \mu^*_2(c_{[s]})$ for all $0 \leq s \leq t - 2$, that is to say $\frac{\partial}{\partial \mu^A_{[t]}} \tilde{L}(\mu^A_{[t]} | c_{[s]}) < 0$ for all $0 \leq s \leq t - 1$. This contradicts the FOC. So, $\frac{\partial}{\partial \mu^A_{[t]}} \tilde{L}(\mu^A_{[t]} | c_{[t-1]}) \geq 0$, which implies $\frac{\partial}{\partial \mu^A_{[t]}} \tilde{L}(\mu^A_{[t]} | c_{[t]}) > 0$ as $c_{[t]} > c_{[t-1]}$ from the inductive hypothesis. Hence we see that $\sum_{s=0}^{t} \frac{\partial}{\partial \mu^A_{[t]}} \tilde{L}(\mu^A_{[t]} | c_{[s]}) > 0$. This shows $\mu^A_{[t+1]} = \mu^*_2(c_{[0]}, \ldots, c_{[t]}) > \mu^A_{[t]}$ by the strict concavity of generation $t$’s objective. Also, $c_{[t+1]} > c_{[t]}$ follows.

So by induction, we have shown Step 1. (The other case of $c_{[1]} < c_{[0]}$ is symmetric.)

For the rest of this proof, suppose Assumption 3 also holds.

**Step 2:** $(\mu^A_{[t]})_{t \geq 1}$ is bounded and converges.

I first show that for every $t$, $(\mu^A_{[t]})_{t \geq 1}$ is bounded between $\mu^A_{[1]}$ and $\mu^\infty_2$. Combined with the fact that $(\mu^A_{[t]})_{t \geq 1}$ is monotonic from Step 1, the sequence must then converge.

Consider the case of $c_{[1]} > c_{[0]}$ (so $\mu^A_{[2]} > \mu^A_{[1]}$). **Step 1** implies that $(\mu^A_{[t]})_{t \geq 1}$ forms an increasing sequence. We have $\mu^A_{[1]} = \mu^*_2(c_{[0]}), \mu^A_{[1]} = c_{[1]}$. We have $\mu^A_{[2]} = \mu^*_2(c_{[1]}), \mu^A_{[2]} = \mu^*_2(c_{[1]}), \mu^A_{[2]} = \mu^*_2(c_{[1]}), \frac{\partial}{\partial \mu^A_{[2]}} \tilde{L}(\mu^A_{[1]} | c_{[1]}) = 0$ and $c_{[1]} > c_{[0]}$. This shows $\mu^A_{[2]} < \mu^A_{[2]},$ hence $c_{[2]} < c_{[2]}$. By
Step 2

Assumptions 1, 2, and 3 hold), meaning by induction. But from the proof of Lemma A.21, arguments from 

Proof.

\( \frac{\partial}{\partial \mu_2} \tilde{L}(\mu_2, t) \mid \mu_t^A \geq \tilde{c}[t] \) given that \( \mu_t^A > \mu_t \). By strict concavity of \( \tilde{L} \cdot \mid \mu_t^A \) from Lemma A.7, this shows \( \mu_2^A(\mu_t^A) > \mu_2, t \), hence also \( \mu_t^2 > \mu_t^A \). So we have established that \( \mu_2, t \leq \mu_2^A \) by induction. But from the proof of Lemma A.21, \( (\mu_2^A, \mu_t^A) \) converge upwards to \( \mu_2^\infty \) in this case (given that they are iterates of \( \mathcal{I} \)), which is a contraction map by Proposition 4 when Assumptions 1, 2, and 3 hold), meaning \( \mu_2, t \) is bounded between \( \mu_2, t \) and \( \mu_2^\infty \). 

The case of \( \mu_t^A < \mu_0 \) is symmetric (and if \( \mu_t^A = \mu_0 \) then \( \mu_2, t = \mu_2^\infty \)). We have proven Step 2. Denote \( \bar{\mu}_t = \lim_{t \to \infty} \mu_2, t \) and observe that since \( C \) is continuous in its second argument, \( \mu_t \to \bar{c} = C(\mu_t^A, \bar{\mu}_t) \).

Step 3: \( \bar{\mu}_t \) is a fixed point of \( \mathcal{I}(\cdot, \bar{c}) \), so in particular \( \bar{\mu}_t = \mu_2^\infty \) and \( \bar{c} = c^\infty \) since \( \mathcal{I}(\cdot, \bar{c}) \) has a unique fixed point by Proposition 4.

Consider the case of \( \mu_t^A > \mu_0 \), for the other case is symmetric. From the proof of Step 2, \( \mu_2, t \) is bounded above by \( \bar{\mu}_2^\infty \), so if \( \mu_2 \neq \mu_2^\infty \) by way of contradiction, then \( \mu_2 < \mu_2^\infty \). Since the iterates of \( \mathcal{I}(\cdot, \bar{c}) \) are monotonic, this implies \( \mathcal{I}(\mu_2^\infty; \bar{c}) = \mu_2^\infty > \mu_2^\infty \), that is \( \mu_2^\infty(\bar{c}) > \mu_2^\infty \). As \( \tilde{L}(\cdot, \bar{c}) \) is strictly concave, this implies \( \int_{-\infty}^{\bar{c}} \frac{\partial}{\partial \mu_2} \tilde{L}(\mu_2, x) \mid x = 0 > 0 \). Using the fact that \( \frac{\partial}{\partial \mu_2} \tilde{L}(\cdot, \bar{c}) \) is decreasing, there must exist \( \epsilon > 0 \) so that \( \int_{-\infty}^{\bar{c}} \frac{\partial}{\partial \mu_2} \tilde{L}(\mu_2, x) \mid x = 0 \geq \epsilon \) whenever \( c \in [\bar{c} - \epsilon, \bar{c}] \) and \( \mu_2 \leq \mu_2^\infty \). Since \( \mu_t^A > \bar{c} \), find large enough \( T \) so that \( \mu_t^A \geq \bar{c} - \epsilon \) whenever \( t \geq T \). Also, let \( B = \max_{\mu_2 \in [\mu_2^A, \mu_2^\infty]} \max_{c \in [c^\infty, \bar{c}]} \int_{-\infty}^{\bar{c}} \frac{\partial}{\partial \mu_2} \tilde{L}(\mu_2, x) \mid x = 0 \). So for \( t \geq T + 1 \), \( \sum_{s=0}^{t-1} \frac{\partial}{\partial \mu_2} \tilde{L}(\mu_2, s) \mid c^\infty \) \( \geq \int_{-\infty}^{\bar{c}} \frac{\partial}{\partial \mu_2} \tilde{L}(\mu_2, x) \mid x = 0 > \epsilon (T - T) \). This quantity must be strictly positive for large enough \( t \), a contradiction that says FOC is not satisfied for large \( t \). Thus, we must have \( \mu_2 = \mu_2^\infty \), hence \( \bar{c} = C(\mu_2^A, \mu_2^\infty; \bar{c}) \).

A.11 Proof of Corollary 1

Proof. Suppose \( c_1 \geq c_0 \). Since \( \mu_2^A(c) \) is increasing, we have \( \mu_2^A(c) = \mu_2^A(c_1), \mu_0^A = \mu_0^A(c_0) = \mu_2^A \). So we get \( c_2 \geq c_1 \). By Theorem 2, we deduce \( c_2(c_1) \) is an increasing sequence, so in particular \( c^\infty \geq c^\bullet \). But again by 2, \( c^\infty \) is the same as the steady-state cutoff in Theorem 1. This is a contradiction because Theorem 1 implies \( c^\infty < c^\bullet \).

This shows \( c_1 < c_0 \) and similar arguments show \( c_2(c_1) \) is a strictly decreasing sequence. Since \( c^\bullet \) is the objectively optimal cutoff threshold under the true model \( \Psi^\bullet \), and since expected payoff under the true model is a single-peaked function in acceptance threshold by Lemma A.2, this shows expected payoff is strictly decreasing across generations.