Dynamic Decision-making and Inference Distortions in Behavioral Economics: A Critical Analysis of Over- and Underreaction

By
Giorgi Machavariani

Submitted to
Central European University
Department of Undergraduate Studies

In partial fulfillment of the requirements for the Bachelor of Arts and Bachelor of Science in Data Science and Society
Major: Data Science and Economics

Supervisor: Mats Köster

Vienna, Austria

2025

Copyright Notice

Author's Declaration

I, the undersigned, **Giorgi Machavariani**, candidate for the Bachelor's degree in Data Science and Society declare herewith that the present thesis is exclusively my own work, based on my research and only such external information as properly credited in notes and bibliography. I declare that no unidentified and illegitimate use was made of the work of others, and no part of the thesis infringes on any person's or institution's copyright. I also declare that no part of the thesis has been submitted in this form to any other institution of higher education for an academic degree.

Vienna, 26 May 2025

Giorgi Machavariani

Abstract

This thesis studies systematic deviations from Bayesian updating, focusing on the tendency to overreact to weak signals and underreact to strong ones. I build on the model by Augenblick, Lazarus, and Thaler (Augenblick et al., 2025), presenting a simplified version where individuals form noisy estimates of signal strength and anchor them toward a fixed default. This generates predictable distortions: people overestimate weak signals and underestimate strong ones.

The model is theoretically straightforward and shows that asymmetric updating emerges naturally from averaging noisy estimates with an internal anchor. Empirically, the model is validated in three domains: lab experiments, a basketball prediction task, and real-world betting and financial markets. Across all settings, belief updates exceed what is warranted when signals are weak and fall short when signals are strong.

These results suggest that inference errors stem from limited precision in perceiving signal strength. This mechanism offers a unified explanation for various belief anomalies and contributes to our understanding of how people respond to uncertainty.

Acknowledgements

I would like to express my deepest gratitude to my thesis supervisor, Mats Köster, for their invaluable guidance, support, and constructive feedback throughout the entire process of this thesis. Their expertise and encouragement have been crucial in shaping this work.

I am also thankful to the faculty and staff at Central European University, with whom I've had the pleasure of working and interacting.

Special thanks to my colleagues and friends, an immensely talented set of people, for their helpful discussions, encouragement, and collaboration, and for making this journey more enjoyable.

Contents

Chapter 1

Introduction

Accurately updating beliefs in response to new information is a foundational concept across economics, psychology, and data-driven disciplines. While Bayesian updating, formalized in the 18th century, remains the normative gold standard for belief revision, a wide body of empirical work suggests that human behavior systematically deviates from this ideal. These deviations are especially pronounced when individuals encounter signals of varying strength.

In his seminal work *Thinking, Fast and Slow*, Kahneman (2011) illustrates how people's natural affinity for causal stories often undermines their ability to engage in statistical reasoning. He references the observation by statistician David Freedman that in any legal trial, the side required to explain regression to the mean is likely to lose, precisely because the average person struggles to interpret statistical regularities. Kahneman's point resonates across many probabilistic settings: humans tend to prioritize narrative coherence over statistical inference, leading to predictable cognitive distortions in belief updating. These challenges raise not only theoretical questions about rationality but also practical questions about the interpretation of uncertainty in real-world contexts.

This thesis builds upon recent developments in behavioral economics and decision theory, especially the model proposed by Augenblick et al. (2025), which was in turn a continuation of the authors' earlier work (Augenblick and Rabin, 2021). In their framework, individuals generally update beliefs in the correct direction but lack precision in assessing the strength of the signals they receive. This miscalibration results in a systematic pattern of overreaction to weak signals and underreaction to strong signals, referred to as limited strength recognition. The model provides a unified explanation for several well-documented belief-updating irregularities, such as base-rate neglect, conservatism, and asymmetric reactions to positive versus negative information.

The present study aims to replicate, simplify and extend the findings of (Augenblick et al., 2025), both in experimental and applied domains. The central hypothesis is that deviations from Bayesian updating are primarily driven by a psychological bias in strength recognition, specifically, people's inability to accurately assess the informativeness of signals. By focusing on this mechanism, the thesis engages with the broader debate over the normative and descriptive adequacy of Bayesian inference.

While Bayes' rule is theoretically optimal, meaning it uniquely satisfies dynamic consistency and consequentialism, it may be too demanding under realistic cognitive constraints. A large and growing literature documents systematic departures from Bayesian behavior. Ortoleva (2024), for instance, surveys models that incorporate biases like motivated reasoning, limited attention, and computational costs. These models argue that non-Bayesian behavior, far from being irrational, may be adaptive in environments with limited cognitive resources.

Recent experimental work lends empirical support to the limited strength recognition hypothesis. Kim et al. (2020) demonstrate that Bayesian guidance improves belief updating for weak signals, but individuals still underweight strong signals, perhaps due to difficulties in interpreting sample size. Ambuehl and Li (2018) similarly show that participants undervalue improvements in signal precision and tend to prefer information that might yield certainty, even if it is diagnostically weaker, which suggests a distorted internal scale for signal informativeness.

This compression of perceived signal strength has significant implications for real-world behavior. In financial markets, Barron (2021) finds that belief updating is more accurate when outcomes are salient and stakes are high, yet even under such conditions, individual heterogeneity persists. (Vansteenberghe, 2024) proposes a formal diagnostic of signal strength based on the alignment between the median and skewness of subjective distributions, revealing that inconsistency in these features correlates with strength misjudgment.

This psychological pattern also aligns with findings in high-dimensional statistics. Shi and Qu (2017) show that weak but nonzero signals often escape detection in model selection. Their results emphasize that accounting for these subtle signals enhances inference and prediction, mirroring the thesis's central idea that small but real cues are frequently overemphasized due to cognitive misjudgment.

Bringing together these strands of literature, this thesis posits that belief updating errors arise not from irrationality, but from bounded statistical intuition. Specifically, individuals systematically misperceive signal strength, leading to overreacting to weak cues and underreacting to strong ones. This lens offers a behavioral framework for understanding a wide range of empirical anomalies across both laboratory and real-world environments.

The remainder of the thesis is organized as follows: the next chapter reviews the theoretical and empirical literature, contrasting Bayesian and non-Bayesian models of belief formation and sharpening the focus on the limited strength recognition hypothesis. Subsequent chapters outline the empirical methods and present experimental evidence replicating and extending (Augenblick et al., 2025)'s findings in a stylized decision-making environment, which is later bridged into the discussion and conclusion.

Chapter 2

Literature Review

2.1 Behavioral Deviations from Bayesian Updating in Belief Formation

A growing body of research finds that individuals systematically deviate from Bayesian updating due to bounded rationality and cognitive imprecision. Rather than interpreting signals with full accuracy, people process information in a noisy, "imprecise" manner. (Enke and Graeber, 2023), for example, introduces the concept of cognitive uncertainty, a subjective uncertainty about one's optimal decision, and show that it leads to weakened belief updating. Cognitively uncertain individuals tend to underreact to new evidence, generating conservative belief revisions that gravitate toward an intermediate prior. This framework unifies several well-known deviations from Bayesian behavior, including conservatism, base-rate neglect, and nonlinear probability weighting, all as consequences of noisy internal processing. Experimental data support this theory: participants reporting higher cognitive uncertainty show weaker responses to new information, suggesting misperception of signal strength Enke and Graeber (2023).

Building on similar ideas, (Ba et al., 2024) proposes a model in which agents face limited attention and cognitive bandwidth, leading them to form simplified mental representations. Individuals apply a representativeness heuristic to focus on salient states while processing signals with noise. This dual constraint generates systematic biases: the model predicts overreaction in complex or noisy environments and underreaction in simpler, high-precision ones. The intuition is that surprising or ambiguous signals are overweighted, while predictable information is underweighted. Their experimental findings support these predictions and help reconcile a key puzzle: underreaction in laboratory tasks versus overreaction in noisy real-world domains like financial markets. Bounded rationality, in the form of limited attention and noisy signal processing, thus produces unforeseen biases in belief updating.

A complementary strand of research draws on psychophysical theory to formalize how signal strength is perceived. Khaw et al. (2021) propose that agents perceive informativeness through a distorted "Fechnerian" lens, leading to compression in perceived signal strength, which was also discussed in an earlier work by one of the authors Woodford (2020). In their meta-Bayesian model, signal strength is transformed concavely, weakening sensitivity to strong evidence and exaggerating responses to weak evidence. Khaw et al. (2021) apply this to risk behavior, but the logic extends to belief formation: weak signals are overperceived, strong ones diluted. This misperception reproduces the empirical pattern of overinference from weak signals and underinference from strong ones, and has been proposed as a unifying framework for belief distortions rooted in cognitive imprecision.

2.2 Misestimation of Signal Strength: Heuristics and Biases

Beyond general cognitive imprecision, several heuristic-driven biases contribute to systematic misestimation of signal strength. One influential concept is the law of small numbers (LSN), the intuitive, though incorrect, belief that small samples should closely reflect population frequencies. Rabin and Vayanos (2010) formalize this by modeling agents who expect random sequences to "self-correct," producing the gambler's fallacy. After a short streak, LSN believers expect reversals, inferring too much from weak patterns. The same model also predicts the opposite, hot-hand style persistence, after longer streaks, reconciling why people sometimes expect reversals and other times continuations. In Bayesian terms, this results in overreaction to weak signals and underreaction to strong ones. The authors apply this to asset markets, showing how it explains momentum, underreaction to news, and long-run reversals, overreaction to noise, Rabin and Vayanos (2010).

Complementing LSN, Benjamin et al. (2016) propose the concept of nonbelief in the law of large numbers (NBLLN), where individuals underestimate the informativeness of large samples Benjamin et al. (2016). Their model suggests that even representative datasets are treated as unreliable, leading to excessive reliance on priors. While a Bayesian grows more confident as sample size increases, NBLLN subjects remain uncertain, discounting strong evidence. The authors demonstrate how this bias contributes to systematic underinference in the presence of clear signals.

To test for LSN and NBLLN jointly, Benjamin et al. (2017) run incentivized experiments using coin-flip sequences. Participants overexpect balance in small samples and underappreciate the convergence of large samples. The result is a distinctive pattern: overinference from weak signals and underinference from strong ones. The design confirms both biases can occur together, leading to inconsistent belief updates.

Another persistent error is base-rate neglect, where individuals underweight prior probabilities. This causes exaggerated belief shifts in response to weak signals. Benjamin et al. (2019) provide a theoretical foundation, showing that cognitively constrained Bayesians may fail to anchor their reasoning on priors Ambuehl and Li (2018). In real-world settings like medicine or law, this leads to undue confidence in unlikely outcomes based on salient, low-probability evidence. Closely related is the representativeness heuristic: judging outcomes by how typical or vivid they appear, rather than their statistical weight. Bordalo et al. (2025) embed this in a diagnostic expectations model, arguing that people overreact to surprising signals and underreact to expected ones. This implies non-linear updates: unexpected weak signals cause excessive belief movement, while strong but expected signals are discounted.

Together, these mechanisms, LSN, NBLLN, base-rate neglect, and representativeness, show how people misestimate signal strength. Across domains, such biases lead to overreaction to noisy cues and underreaction to diagnostic evidence.

2.3 Experimental Evidence: Overinference from Weak Signals vs. Underinference from Strong Signals

A growing body of controlled experimental research provides direct evidence for asymmetries in belief updating, specifically, the tendency to overreact to weak signals and underreact to strong ones. Augenblick et al. (2025) conduct a comprehensive laboratory experiment in which participants perform a Bayesian inference task involving signals about a binary state with known signal strength. By systematically manipulating signal strength, the authors measure deviations from Bayesian updating.

The results show a non-monotonic response. For highly informative signals—expected to shift posteriors toward 0.8—participants underinfer, adjusting less than Bayes' rule prescribes. For moderate signals, responses align with Bayesian predictions. But for weak signals, participants over-update, treating minimal evidence as if it were nearly twice as diagnostic as it is. This convex update pattern, under-sensitivity to strong signals, over-sensitivity to weak ones, matches predictions from cognitive imprecision models.

The authors interpret this through a psychological and physical lens: individuals perceive signal strength in a compressed, nonlinear way. A simple model with a perception exponent $\beta \approx 0.76$ fits the pattern well, showing that people perceptually flatten informativeness toward a mid-range. This misperception carries over to valuation: participants overpaid for weak signals and underpaid for strong ones, revealing distorted willingness to pay based on perceived rather than actual signal strength (Augenblick et al., 2025).

Ambuehl and Li (2018)offer converging evidence through incentive-compatible experiments on valuation of signals with varying precision. They find that participants underreact to increased precision, undervaluing high-quality signals, while simultaneously preferring signals that might yield certainty, even when less reliable. This suggests a psychological bias favoring potentially conclusive evidence over accuracy.

These tendencies appear stable across individuals. Ambuehl and Li (2018) identify a personal "responsiveness to information" parameter that predicts both updating strength and willingness to pay. Crucially, this variation is uncorrelated with math skills, implying it reflects cognitive structure rather than ability. Their findings reinforce those of (Augenblick et al., 2025), pointing to internal processing constraints as the root of belief distortion.

Earlier work by Phillips and Edwards (1966) were among the first to document conservatism in belief updating: the tendency to underweight new evidence. In classic urn experiments, participants adjusted beliefs less than a Bayesian would. This underreaction has been robustly replicated, and recent studies like Ba et al. (2024) suggest it may coexist with overreaction in noisy or dynamic settings. These results indicate that belief updating biases are persistent and context sensitive.

Overall, this experimental literature shows that updating errors are systematic, not stochastic. Individuals apply an internal "mental filter" that compresses perceived informativeness, leading to muted responses to strong evidence and exaggerated shifts in response to flimsy signals.

2.4 Empirical Evidence and Applications

The asymmetric belief-updating biases observed in laboratory settings have clear parallels in real-world environments, particularly financial markets. One persistent empirical puzzle is that markets exhibit both underreaction and overreaction to news, seemingly contradictory behaviors that a signal strength perspective helps explain. When firms release clear, positive news, such as earnings announcements that exceed expectations, prices often adjust only partially at first, with further movement occurring gradually. This phenomenon, known as post-earnings announcement drift, reflects underinference from strong signals: investors fail to fully incorporate informative news immediately.

Conversely, markets also display excess volatility, where price movements are too large to be justified by fundamentals. This may stem from overreaction to weak, noisy signals, such as rumors, sentiment shifts, or ambiguous indicators, that traders overweight. Rabin and Vayanos (2010) connect their model of the law of small numbers to these patterns, showing how gambler's fallacy reasoning can generate short-run momentum and long-term reversals in asset returns. Underreaction to meaningful news may give way to overreaction as noise is misinterpreted as trend, followed by correction once randomness is recognized.

Supporting this interpretation, Bordalo et al. (2025) document that professional macroeconomic forecasters often overreact to recent surprises, placing excessive weight on salient data Bordalo et al. (2025). Forecast errors tend to mean-revert. More broadly, the authors, along with some extra collaborators, develop the diagnostic expectations framework to explain credit cycles, macroeconomic booms and busts, and asset price bubbles. These events are modeled as belief shifts driven by salient but noisy signals that initially displace expectations from fundamentals before eventual correction, which is further discussed with complementary references in (Bordalo et al., 2025).

In contrast, survey-based studies such as Coibion and Gorodnichenko (2012) find information rigidity in settings like inflation expectations. Here, agents underreact to strong public signals, updating forecasts too modestly in response to macroeconomic data. This mirrors the underinference observed in controlled experiments and suggests that even high-quality information does not eliminate cognitive conservatism.

Recent work connects these patterns to bounded cognition. Enke and Graeber (2023) find that cognitive uncertainty among consumers and managers dampens responses to new information (Enke and Graeber, 2023). Meanwhile, modelslike (Woodford, 2020) propose that memory and perception constraints lead agents to optimally compress past signals, producing either sticky expectations or overshooting depending on signal characteristics. These models illustrate how cognitive resource allocation shapes observable deviations from Bayesian rationality.

Individual heterogeneity also matters. While experts or individuals with high cognitive reflection may exhibit smaller biases, even professionals are not immune. Under complexity or noise, even incentivized actors display conservatism or overreaction, reinforcing the view that these biases stem from core cognitive mechanisms rather than inexperience.

In sum, empirical evidence supports the view that belief updating is shaped by bounded rationality. Rather than uniformly overreacting or underreacting, individuals apply internal filters that dampen extreme updates and exaggerate minor ones. Theoretical models grounded in cognitive noise, signal compression, or heuristic-driven expectations provide coherent explanations for these behaviors. Alongside experimental validation, this literature presents a compelling account of belief formation at both the individual and market level. A consistent theme emerges: misestimation of signal strength—overinference from weak signals and underinference from strong ones—is a robust and consequential feature of human judgment.

2.5 The Current Research Gap and Contributions

As the previous sections have shown, the literature on belief updating has moved beyond normative Bayesian models toward more descriptive behavioral frameworks. While many approaches capture deviations such as underreaction, overreaction, or distorted beliefs, they often address specific contexts or mechanisms in isolation. Recent work, however, seeks more unified explanations. One such framework is the limited strength recognition hypothesis, which this thesis directly builds upon.

Augenblick et al. (2025) propose a comprehensive model that accounts for both overinference and underinference within a single structure. Their central idea is that people form noisy estimates of signal strength, even when signal direction is interpreted correctly. This cognitive imprecision leads to belief updates that "shrink" toward average signal strength, producing overreaction to weak signals and underreaction to strong ones. Their predictions hold across experimental and real-world settings, including financial markets and sports betting.

Despite its promise, this model has not yet been widely tested beyond the authors' original experiments. Additionally, many classic updating tasks, such as the urn problem, are limited in their ability to separate motivational from cognitive biases, leaving a gap in the empirical validation of the model.

This thesis helps fill that gap by replicating and formalizing the core predictions of (Augenblick et al., 2025). It uses stylized Bayesian environments with controlled signal-to-noise ratios to test whether individuals exhibit the predicted pattern of asymmetric updating. The design minimizes motivational influences, focusing on abstract probabilistic reasoning. This isolates the cognitive mechanism from confounds such as wishful thinking or self-serving biases, often present in high-stakes belief updating tasks Barron (2021).

Methodologically, this thesis adapts tools from (Augenblick et al., 2025) such as shrinkage estimation and posterior log-likelihood comparisons, to a simplified binary decision task. These tasks are structured for clearer measurement and easier interpretation. While earlier studies have documented belief updating errors in naturalistic settings, this work prioritizes experimental clarity to directly test the model's core logic.

Chapter 3

Methodology

3.1 Theoretical Foundations: Overinference from Weak and Underinference from Strong Signals

This section introduces the core theoretical framework developed by Augenblick, Lazarus, and Thaler in (Augenblick et al., 2025) to explain a puzzling pattern in human belief updating: people tend to overreact to weak signals and underreact to strong ones. Unlike classical models that assume uniform overreaction or underreaction, their theory proposes that the nature of the inference bias critically depends on the signal's strength. This framework blends standard Bayesian logic with cognitive constraints, resulting in a psychologically plausible model of conservative yet biased updating.

3.1.1 Signal Structure and Bayesian Baseline

The theoretical framework begins with a simple binary-state environment in which the true state of the world is denoted by $\theta \in \{0,1\}$. The decision-maker (DM) receives a signal s that provides probabilistic information about θ. This signal has two components: a direction, which suggests which state is more likely, and a magnitude, which captures the strength or informativeness of the signal. The signal's direction is denoted by $s_d \in \{+1,-1\}$, where $+1$ indicates that the signal favors state $\theta = 1$, and -1 indicates it favors $\theta = 0$. The direction determines the sign of the update in log-odds space.

To formalize this, suppose that the signal s is drawn from a distribution that depends on the true state θ. Specifically, let $p(s \mid \theta)$ denote the likelihood of observing signal s conditional on the state θ. If one considers two possible distributions, one under $\theta = 1$ and another under $\theta = 0$, then the informativeness of a given signal can be captured by how much more likely it is under one state than the other.

This leads to the definition of the (log-)likelihood ratio:

$$S(s) = \left| \log \left(\frac{p(s \mid \theta = 1)}{p(s \mid \theta = 0)} \right) \right|.$$

This expression measures how strongly the signal s favors one state over the other. The greater the value of $S(s)$, the more informative the signal is about the underlying state. Note that $S(s)$ is defined as a non-negative quantity. The direction of the signal (captured by s_d) determines whether the log-odds are increased or decreased. The quantity $S(s)$ is therefore referred to as the true signal strength.

Next, let's describe how a Bayesian decision-maker updates their belief about the true state depending on the signal strength $S(s)$. Denote the agent's prior that the state is high by $\pi_0 := p(\theta = 1)$ and it's posterior upon observing the signal s by $\pi_1 := p(\theta = 1|s)$. Moreover, for any probability $\pi \in (0,1)$, let's define the logit function as $\text{logit}(\pi) := \log\left(\frac{\pi}{1-\pi}\right)$, which maps probabilities to the real line. As we show in more detail below, Bayes' rule then implies that

$$\text{logit}(\pi_1) = \text{logit}(\pi_0) \pm S(s),$$

where the sign of the update depends on the direction of the signal. This transformation reflects standard Bayesian updating in log-odds space, where the informativeness of the signal linearly shifts the prior log-odds.

Now I derive the log-odds updating formula step by step. By Bayes' rule, the posterior belief after observing signal s is given by:

$$\pi_1 = p(\theta = 1 \mid s) = \frac{p(\theta = 1) \cdot p(s \mid \theta = 1)}{p(s)} = \frac{\pi_0 \cdot p(s \mid \theta = 1)}{\pi_0 \cdot p(s \mid \theta = 1) + (1 - \pi_0) \cdot p(s \mid \theta = 0)}.$$

Using the identity $p(\theta = 0 \mid s) = 1 - \pi_1$, resulting is the odds form:

$$\frac{\pi_1}{1-\pi_1} = \frac{\pi_0}{1-\pi_0} \cdot \frac{p(s \mid \theta = 1)}{p(s \mid \theta = 0)}.$$

Taking natural logarithms of both sides yields the log-odds updating rule:

$$\log\left(\frac{\pi_1}{1-\pi_1}\right) = \log\left(\frac{\pi_0}{1-\pi_0}\right) + \log\left(\frac{p(s \mid \theta = 1)}{p(s \mid \theta = 0)}\right).$$

This can be compactly written using the logit function as:

$$\operatorname{logit}(\pi_1) = \operatorname{logit}(\pi_0) + \log\left(\frac{p(s \mid \theta = 1)}{p(s \mid \theta = 0)}\right).$$

In cases where the signal favors $\theta = 0$, the log-likelihood ratio is negative, and the updating formula becomes:

$$\operatorname{logit}(\pi_1) = \operatorname{logit}(\pi_0) \pm S(s),$$

where $S(s)$ denotes the log-likelihood ratio, as defined earlier:

$$S(s) = \left|\log\left(\frac{p(s \mid \theta = 1)}{p(s \mid \theta = 0)}\right)\right|.$$

3.2 Simplification and Contribution: A Tractable Behavioral Model

To make the theoretical structure more analytically manageable while preserving its core logic, I develop a simplified version of the model presented by Augenblick et al. (2025). This version retains the central insight, which is that belief updating involves an averaging process between a noisy internal estimate and a default strength, but simplifies the informational environment and behavioral assumptions. This not only facilitates clearer exposition but also enables straightforward derivation of key results, including a formal version of Proposition 1 under minimal conditions.

3.2.1 Simplified Environment

I consider a binary-state setting where the true state of the world is denoted by $\theta \in \{0,1\}$. The decision-maker (DM) holds a prior belief $\pi_0 = \mathbb{P}(\theta = 1)$, and receives a signal $s \in \{0,1\}$ that is informative about the state. The informativeness of the signal is governed by a likelihood parameter $w \in [0.5,1]$. In this binary setting, the observed signal $s \in \{0,1\}$ does not by itself determine the signal strength. Rather, informativeness is encoded by the underlying likelihood parameter w, which maps to signal strength via $S(w) = \left|\log\left(\frac{w}{1-w}\right)\right|$. where:

$$p(s = 1 \mid \theta = 1) = w, \quad p(s = 1 \mid \theta = 0) = 1 - w.$$

The closer w is to 0.5, the less informative the signal, the closer it is to 1, the more strongly it favors state $\theta = 1$.

A fully Bayesian agent would use the true value of w to update their beliefs. In contrast, I assume that the individual does not observe the exact diagnosticity w, but instead receives an internal estimate e of its informativeness.

3.2.2 Behavioral Assumption: Noisy Strength Estimation

The decision-maker observes the direction of the signal (whether $s = 1$ or $s = 0$), but does not directly perceive the strength of the evidence. Instead, they form an internal estimate e of the signal's informativeness, which corresponds to the magnitude of the log-likelihood ratio — that is, the strength of the signal in Bayesian terms.

If w denotes the diagnosticity, or the probability of observing $s = 1$ under state $\theta = 1$, then the true signal strength is given by:

$$S(w) = \left|\log\left(\frac{w}{1-w}\right)\right|.$$

The internal estimate e is assumed to be a noisy but unbiased signal of this log-likelihood magnitude. Formally, the behavioral assumption is:

$$\mathbb{E}[e \mid S(w)] = S(w).$$

This means that while the individual does not know the exact value of the log-likelihood ratio, their estimate is correct on average. However, due to cognitive or informational noise, each realized estimate e may deviate from the truth.

Recognizing the imprecision in their estimate, the individual does not fully rely on e, but instead forms a perceived signal strength by averaging it with a default reference level S_0, which represents a default or reference signal strength the individual uses as a cognitive anchor. It may reflect prior experience, institutional norms, or a heuristic for "typical" informativeness. This leads to:

$$\hat{S} = \alpha \cdot e + (1 - \alpha) \cdot S_0,$$

where $\alpha \in (0, 1)$ represents the individual's confidence in their own estimate. A smaller α corresponds to greater caution and a stronger reliance on the default.

This structure captures the central behavioral mechanism: conservative updating in the presence of uncertainty, leading to overinference when signals are weak and underinference when signals are strong.

3.2.3 Constant Weighting and Independence from Strength

A key simplification in my approach is the assumption that α is constant and does not vary with the observed estimate e or the underlying true strength w. This differs from the original model, which allows for more general weighting structures (including cases where α may be signal-dependent). By holding α fixed, the model becomes easier to analyze while still capturing the essential behavioral mechanism: a bias introduced by shrinkage toward a default.

This assumption also implies that the weight placed on the estimate is unaffected by the perceived quality or magnitude of the signal. Psychologically, this can be interpreted as the individual treating all their estimates as equally reliable, a plausible heuristic in situations where cognitive resources are limited or past experience is lacking.

3.2.4 Interpretation of Parameters and Bias

Each of the parameters in this simplified structure carries intuitive meaning. The default signal strength S_0 serves as a fixed reference point in the individual's belief system. This anchor could arise from personal experience, a commonly used heuristic, or institutional defaults. The weighting parameter α reflects the individual's confidence in their internal estimate e, with lower values implying greater caution and heavier reliance on the default. Finally, the estimate e itself, while potentially noisy, is assumed to be unbiased in expectation, that is, although any given estimate may deviate from the truth, the average of many such estimates will converge to the true signal strength S.

These assumptions imply a predictable pattern of inference bias:

$$\mathbb{E}[\hat{S} \mid S] = \alpha \cdot \mathbb{E}[e \mid S] + (1-\alpha) \cdot S_0 = \alpha S + (1-\alpha) S_0.$$

Hence, the expected bias in perception is:

$$\text{Bias}(S) = \mathbb{E}[\hat{S} \mid S] - S = (1-\alpha)(S_0 - S).$$

This equation formalizes the key insight: when $S < S_0$, individuals overestimate the signal's strength, when $S > S_0$, they underestimate it. There is a unique switching point $S^* = S_0$ where perception is accurate on average.

3.2.5 Reduction of Assumptions

An important contribution of my approach is the recognition that the behavioral prediction, the overinference/underinference asymmetry, does not require strong structural assumptions. In particular, I show that the core result, a single crossing point at which inference is unbiased, can be derived solely from the unbiasedness of the internal estimate, the constancy of α, and the presence of a fixed reference value w_0.

In contrast, the original paper's derivation of Proposition 1 relies on more general assumptions like monotonicity and the monotone likelihood ratio property (MLRP). By stripping down the model, I demonstrate that the essence of the result survives under far weaker conditions. This not only improves theoretical clarity but makes the model more portable across different applied contexts, from experimental belief updating to market responses and decision support systems.

3.2.6 Behavioral Updating: Cognitive Constraints and Signal Estimation

While the Bayesian benchmark assumes full rationality, the behavioral model relaxes this by introducing cognitive imprecision. The individual is assumed to perfectly observe the direction s_d, but only has an internal, noisy estimate of the signal strength, denoted by e. Crucially, this estimate satisfies two properties: The estimate e is assumed to be an imperfect signal of the true strength S, but it satisfies two key properties. First, it is unbiased in expectation: across many realizations, the average estimate coincides with the true value, that is, $\mathbb{E}[e \mid S] = S$. This means individuals are not systematically biased upward or downward in their perception of signal strength. Second, the estimate is noisy, meaning individual realizations of e may differ substantially from S, and importantly, individuals are aware of this imprecision.

To illustrate, consider a person interpreting a financial analyst's report. Suppose the true informativeness of the report is $S = 2.0$. The individual might form an internal impression e of this strength based on tone, wording, or heuristics. Sometimes this estimate might be 1.5, other times 2.5, but on average, it centers around 2.0. Knowing that their impressions vary, the person acknowledges the risk of overreacting or underreacting and adjusts accordingly.

Given this uncertainty, individuals do not update their beliefs using the noisy estimate e in isolation. Instead, they adopt a conservative strategy: they average e with a fixed default value S_0, which represents their baseline expectation of signal strength. The resulting perceived signal strength is:

$$\hat{S}(s_d, e) = \alpha \cdot e + (1 - \alpha) \cdot S_0,$$

where $\alpha \in (0, 1)$ captures the individual's confidence in their estimate. A higher α corresponds to greater reliance on e, while a lower α reflects greater caution and anchoring to the default.

This leads to a form of belief revision that resembles Bayesian reasoning but incorporates psychological caution. This style of reasoning is often described as quasi-Bayesian. This updating behavior is not strictly Bayesian in the traditional sense. A Bayesian agent with a well-defined prior and likelihood function would directly apply Bayes' rule to update their beliefs using all available information, in this case, e, and would not average it with a default unless the default emerged naturally as a prior distribution. However, the behavior modeled here can be interpreted as quasi-Bayesian: it resembles Bayesian updating under the assumption that the individual treats e as an uncertain observation and approximates the posterior by shrinking toward a prior expectation. This type of convex combination, sometimes called "anchoring and adjustment", is well documented in behavioral literature and reflects bounded rationality rather than strict Bayesian rationality.

As a result, the model predicts a systematic asymmetry in perceived signal strength. When $e < S_0$, the individual overestimates the signal's informativeness, when $e > S_0$, they underestimate it. This leads directly to overinference in response to weak signals and underinference in response to strong ones.

3.2.7 Formal Result and Single-Crossing Property

The core behavioral assumption that individuals combine a noisy estimate of signal strength with a fixed anchor, leads to a specific and testable prediction about how inference bias varies with signal strength.

Recall that a Bayesian agent would update beliefs according to the log-odds rule:

$$\text{logit}(\pi_1) = \text{logit}(\pi_0) + \log\left(\frac{p(s \mid \theta = 1)}{p(s \mid \theta = 0)}\right) = \text{logit}(\pi_0) \pm S(s),$$

where $S(s)$ is the true signal strength and the sign depends on the signal direction.

In contrast, the behavioral agent does not observe the true value $S(s)$, but instead forms a perceived signal strength $\hat{S}(s_d, e)$, given by:

$$\hat{S} = \alpha \cdot e + (1 - \alpha) \cdot S_0,$$

with e being a noisy, unbiased internal estimate of $S(s)$. The agent then uses this perceived strength in the same log-odds updating formula:

$$\text{logit}(\pi_1) = \text{logit}(\pi_0) \pm \hat{S}.$$

This substitution changes the nature of the inference. Because $\hat{S}$ is a convex combination of the estimate and a default, the agent does not respond proportionally to the actual informativeness of the signal. Instead, their update is conservative, the posterior belief moves less than it would under full Bayesian updating when the signal is strong, and more than it should when the signal is weak.

Formally, the inference bias is defined as the difference between the expected perceived signal strength and the true strength:

$$\text{Bias}(S) := \mathbb{E}[\hat{S} \mid S] - S = (1-\alpha)(S_0 - S).$$

This bias is positive when $S < S_0$ (overinference) and negative when $S > S_0$ (underinference). Therefore, the central theoretical result:

Proposition 1. There exists a unique threshold $S^* = S_0$ such that individuals overinfer when $S < S^*$ and underinfer when $S > S^*$.

This proposition follows from the structure of the model: the estimate e is unbiased, the weight α is constant, and the perceived signal strength is a linear average of e and the anchor S_0. The expectation $\mathbb{E}[\hat{S} \mid S]$ is thus a linear function with slope $\alpha < 1$ that intersects the 45-degree identity line exactly once, a property known as single crossing.

Example. Consider a student interpreting feedback on an assignment. Suppose the true informativeness of the feedback is $S = 0.6$, and the student internally estimates this strength as $e = 0.6$. The student anchors to a default level of informativeness $S_0 = 1.0$, and applies a weighting factor $\alpha = 0.5$. Suppose the feedback score is transformed into a signal strength $S = 0.6$, but due to uncertainty or lack of calibration, the student estimates it as $e = 0.6$. If the student has a default sense of feedback informativeness $S_0 = 1.0$ and uses a weight $\alpha = 0.5$, then the perceived strength becomes:

$$\hat{S} = 0.5 \cdot 0.6 + 0.5 \cdot 1.0 = 0.8.$$

The student reacts more strongly than the actual informativeness warrants, this is overinference. Alternatively, if the feedback were very strong, say $S = 1.4$, the same calculation yields:

$$\hat{S} = 0.5 \cdot 1.4 + 0.5 \cdot 1.0 = 1.2,$$

which is weaker than the true value, the student underreacts. In both cases, the conservatism introduced by anchoring to a fixed reference point leads to predictable deviations from Bayesian updating, consistent with the theory.

3.3 Mathematical Formalization of the Simplified Model

Having laid out the theoretical and behavioral underpinnings of the model, I now present a formal derivation of its central result: that individuals overinfer from weak signals and underinfer from strong ones. This section walks through the logic in a clear, step-by-step manner, highlighting the minimal assumptions required for the result to hold.

3.3.1 Averaging Behavior and Perceived Strength

Let S denote the true signal strength, and let e represent the individual's internal estimate of this strength. As established earlier, the individual does not act solely on this estimate, but instead forms a weighted average between it and a default or reference strength S_0. This perceived strength is given by:

$$\hat{S} = \alpha \cdot e + (1 - \alpha) \cdot S_0,$$

where $\alpha \in (0, 1)$ represents the individual's confidence in their estimate. A lower α indicates a more cautious approach, pulling the perceived signal strength closer to the anchor S_0.

This averaging rule captures a psychologically intuitive behavior: people recognize their internal estimates are noisy, and thus hedge by reverting partially to a typical or average level. This leads to systematic misperception of signal strength, depending on whether the true signal is weaker or stronger than the anchor.

3.3.2 Expectations and Inference Bias

To understand how this behavior leads to inference bias, we consider the expected value of perceived strength conditional on the true signal strength. If we assume that the internal estimate is unbiased, that is,

$$\mathbb{E}[e \mid S] = S,$$

Then it follows that:

$$\mathbb{E}[\hat{S} \mid S] = \alpha \cdot \mathbb{E}[e \mid S] + (1-\alpha) \cdot S_0 = \alpha S + (1-\alpha) S_0.$$

This expression shows how the expected perceived strength depends linearly on the true signal strength. We can now define the inference bias, the difference between perceived and true strength, as:

$$\text{Bias}(S) = \mathbb{E}[\hat{S} \mid S] - S = (1-\alpha)(S_0 - S).$$

This bias is positive when $S < S_0$, indicating overinference, and negative when $S > S_0$, indicating underinference. The bias vanishes precisely at $S = S_0$, yielding the model's central result:

> **Proposition 1 (Simplified).** There exists a unique switching point $S^* = S_0$ such that individuals overinfer for $S < S^*$ and underinfer for $S > S^*$.

3.3.3 Interpretation of the Single-Crossing Property

The function $\mathbb{E}[\hat{S} \mid S]$ is linear in S, with slope α strictly less than one. Consequently, the graph of expected perceived strength lies above the identity line for small values of S, crosses it once at $S = S_0$, and remains below it for large values. This geometric property ensures that inference bias changes sign exactly once. It is this structure that drives the model's prediction of over- and underreaction.

3.3.4 No Additional Assumptions Required

A key strength of this derivation lies in its simplicity and generalization efficiency. In contrast to the original formulation by (Augenblick et al., 2025), which depends on broader conditions such as monotonicity and the monotone likelihood ratio property (MLRP), the simplified version requires only a minimal set of assumptions. Specifically, it assumes that the internal estimate e is unbiased in expectation, so that $\mathbb{E}[e \mid S] = S$. Furthermore, it posits that the weighting parameter α is constant and does not vary with the value of e or the true signal strength S. Finally, the model assumes a fixed and signal-independent default strength S_0 toward which individuals anchor their beliefs.

These three ingredients, unbiased estimation, constant weighting, and a fixed reference point, are sufficient to derive the core result without invoking any distributional assumptions about e or structural features of the signal-generating process. The result is a model that is both analytically transparent and widely applicable across domains.

3.3.5 Link to Perceived Diagnosticity

Although the derivation here uses the abstract quantity S to represent signal strength, it can be readily translated into more concrete contexts. For instance, in an adaptation to a Bayesian updating setting with binary signals, the parameter w, the likelihood of observing signal $s = 1$ conditional on state $\theta = 1$, plays the role of informativeness. The same averaging structure then applies:

$$\hat{S} = \alpha \cdot e + (1 - \alpha) \cdot S_0,$$

with corresponding inference bias:

$$\text{Bias}(S) = \mathbb{E}[\hat{S} \mid S(w)] - S(w) = (1 - \alpha)(S_0 - S(w)).$$

This generality illustrates the robustness of the over-/underinference result: it emerges naturally from a simple structure of conservative belief revision in the face of internal uncertainty.

Chapter 4

Results

4.1 Empirical Validation of the Framework

Accurately evaluating a behavioral model of belief updating requires more than simply observing stated beliefs, it depends critically on designing experiments that isolate the core mechanisms at work while minimizing potential confounds. In their empirical validation of the over-/underinference framework, Augenblick et al. (2025) implement a series of carefully structured experiments to test the theory's central prediction: individuals tend to overinfer from weak signals and underinfer from strong ones, primarily due to cognitive imprecision in estimating signal strength.

Selecting the right experimental designs is essential in this context because subtle features, such as the presentation of information, the perceived informativeness of signals, or the cognitive demands of the task, can all systematically influence outcomes. As highlighted in the broader behavioral literature, deviations from Bayesian updating often emerge from representativeness heuristics, cognitive uncertainty (Enke and Graeber, 2023), and limits on attentional processing (Achtziger et al., 2014). If experiments are not carefully crafted, these biases may either be mistaken for the model's predictions or complicate them altogether.

For example, individuals are known to compress signal strength and regress toward average probabilities under conditions of uncertainty (Enke and Graeber, 2023), and may rely on heuristics that prioritize signal direction over reliability (Mohrschladt et al., 2024). Others may overreact to signal strength when it offers the possibility of certainty effect or when sample representativeness is high, as shown in studies of belief movement and information valuation Ambuehl and Li (2018). These effects highlight the delicacy of inference: even subtle framing or complexity can lead to measurable distortions in updating.

Given these risks, Augenblick et al. (2025) take a robust empirical approach: across multiple experiments, they vary the statistical structure of signals, the mode of presentation, and the degree of noise in signal strength estimates. This diversity helps distinguish genuine behavioral patterns predicted by their model from artifacts of task design or misinterpretation.

Ultimately, the experiments not only validate the central logic of the model, demonstrating the predicted asymmetry in belief updating, but also demonstrate its robustness across settings. This lends credibility to the framework's psychological realism and reinforces the importance of experiment selection in testing cognitively grounded economic theories.

4.1.1 Study 1a: Controlled Abstract Belief Updating Task

The first empirical test of the over-/underinference framework uses a stylized belief-updating experiment based on the classic "bookbag-and-poker-chips" paradigm. Participants infer the identity of a hidden card deck (Green or Purple) based on observed draws of cards. Each deck contains a different ratio of diamonds and spades. The relative proportions control the informativeness of each observed draw.

Signal strength is defined as the log-likelihood ratio of observing the drawn suit under the Green deck versus the Purple deck:

$$S(s) = \left| \log \left(\frac{p(s \mid \text{Green})}{p(s \mid \text{Purple})} \right) \right|.$$

By adjusting deck compositions, say, 90% vs. 10% diamonds in the strong condition vs. 60% vs. 40% in the weak condition, the authors create 32 unique levels of signal precision. In each round, participants observe a single card and report the probability it came from the Green deck. The signal's direction is always clear, but its strength varies parametrically.

This design cleanly tests the model's prediction. Under Bayes' rule, log-odds shift linearly with $S(s)$. Participants, however, exhibit conservative updating: their belief changes are too large when $S(s)$ is small (overinference), and too small when $S(s)$ is large (underinference).

Empirical results confirm this asymmetry. Perceived signal strength inferred from reported posteriors increases with true $S(s)$, but the relationship is attenuated: the slope is significantly below one, indicating compression toward an intermediate strength.

The experiment also reveals individual heterogeneity. Distortion is larger among participants with lower cognitive reflection test (CRT) scores, greater response variance, and lower self-reported certainty. These patterns support the model's foundation: signal direction is perceived correctly, but signal strength is estimated with noise that varies across individuals.

Study 1a provides a clear test in a controlled setting, abstracting away from domain-specific heuristics or content. The results strongly support the model's mechanism and demonstrate that asymmetric updating arises even under idealized inference tasks.

4.1.2 Study 1b: Priors and Cognitive Uncertainty

Study 1b replicates the design of Study 1a while introducing asymmetrical prior probabilities like 1/2, 1/3, 1/4. This tests the robustness of the over-/underinference pattern and examines whether it can be explained by base-rate neglect. Participants again view a single card draw and report posterior probabilities.

The pattern persists: weak signals are overweighted and strong ones underweighted. Regressions correcting for base-rate neglect confirm that the distortion arises from misperception of signal strength, not prior probabilities.

The researchers also collect subjective uncertainty ratings. Individuals who report greater uncertainty exhibit stronger deviations from Bayesian updating, reinforcing the role of noisy strength estimation. This provides a direct proxy for imprecision in perceived informativeness.

Together with Study 1a, these findings show that belief updating asymmetry arises from cognitive imprecision in signal strength, rather than direction or priors, and remains stable across variations in prior information.

4.1.3 Study 2: Naturalistic Prediction of Basketball Outcomes

To test the external validity of the framework, Augenblick et al. (2025) implement a more naturalistic belief-updating task based on NBA basketball games. Participants predict which team will win given simplified game scenarios. Each scenario spans four possessions in a specific game quarter, with prior win probabilities based on historical data provided at the outset. After each event, like a made or missed basket, participants update their forecasts.

The key manipulation is timing: the same type of signal—a scored basket—carries more information in the fourth quarter than in the first. This variation creates a natural proxy for signal strength. Early-game events are relatively uninformative, late-game events are more diagnostic.

True win probabilities are assigned using the Inpredictable.com calculator, which provides a normative benchmark. Participants overwhelmingly update in the correct direction after each event, indicating no confusion about signal sign. However, belief changes deviate systematically from Bayesian predictions: early-quarter signals are overweighted, and late-quarter signals underweighted. The crossover point occurs around the third quarter, mirroring the structure in previous studies.

As in Study 1b, reported uncertainty correlates with greater deviation from Bayesian updating, reinforcing the role of noisy strength estimation. Base-rate neglect is modest and does not account for the asymmetry.

Study 2 demonstrates that the model's predictions extend beyond abstract tasks. Even when signals are narrative and intuitive, like scoring plays in basketball, individuals show the same pattern of asymmetric updating. This supports the view that the over-/underinference mechanism generalizes to real-world decision environments.

4.1.4 Summary of Empirical Evidence

Taken together, the three experiments provide strong and convergent empirical support for the over-/underinference framework. Study 1a establishes the core asymmetry in a highly controlled environment, demonstrating that even with transparent statistical structures and clear signal direction, individuals systematically distort signal strength. Study 1b shows that this effect is robust to variations in prior probabilities and correlates meaningfully with cognitive uncertainty and self-assessed confidence. Study 2 extends the findings to a naturalistic, dynamic setting and confirms that the same updating asymmetries emerge when signals are intuitive rather than numeric. Across all studies, the evidence suggests that belief updating is directionally correct but systematically imprecise in magnitude, strongly validating the framework's central claim that individuals overinfer from weak signals and underinfer from strong ones due to noisy strength perception.

4.1.5 Quasi-Experimental Evidence from Sports Betting and Financial Markets

Beyond the experimental studies, Augenblick et al. (2025) test the predictions of the over-/underinference framework in real-world high-stakes environments. They analyze belief updating dynamics in two domains where probabilistic forecasts are observable and evolve over time: sports betting markets and financial derivatives markets. Although these settings do not permit controlled randomization, they offer natural quasi-experimental variation in signal strength and allow for the use of a powerful diagnostic: the comparison of belief movement to uncertainty reduction.

In both domains, the authors exploit a theoretical result derived from Bayesian updating. Under Bayes' rule, beliefs form a martingale process. This implies that the expected movement of beliefs—measured as the sum of squared changes over time—should equal the expected reduction in uncertainty, defined as the decrease in posterior variance. Deviations from this equality indicate systematic overreaction, like excess movement, or underreaction, such as insufficient movement, relative to the informativeness of signals.

Using over five million transactions from betting markets across five major sports, the authors observe the evolution of market-implied win probabilities during live games. They partition each game into equal-length time chunks and compute, for each period, the average belief movement and uncertainty reduction. The key empirical insight is that signals are generally weaker early in a game, when the outcome is highly uncertain, and stronger as the game nears conclusion. Consistent with the framework, belief movement in early periods systematically exceeds uncertainty reduction, implying overinference, while in later periods, movement lags behind uncertainty reduction, indicating underinference. The crossover point occurs around the third quarter of games like basketball and football, mirroring the structure observed in Study 2.

In parallel, the authors analyze S&P 500 index option prices from 1996 to 2018 to estimate market-implied risk-neutral beliefs about future returns. They again compute belief movement and uncertainty reduction over time to contract maturity. Similar to sports markets, early in a contract's life, when signals about the future are inherently weak, beliefs move more than uncertainty reduction warrants. Near expiration, signals become sharper, but belief movement is muted relative to the information conveyed, implying underinference.

This pattern of excess movement early and conservative updating late replicates across both domains. Importantly, it arises even though individual-level cognitive biases cannot be directly observed in these aggregate settings. The findings thus suggest that the asymmetric updating predicted by the over-/underinference model can manifest at the systemic level, potentially shaping prices and decision-making in high-stakes, information-rich environments.

The market and betting evidence substantially broadens the empirical relevance of the framework. It suggests that the cognitive imprecision documented in controlled experiments is not confined to abstract tasks or unincentivized settings. Instead, it may aggregate into observable distortions in collective belief dynamics. Moreover, the empirical method, based on comparing belief movement and uncertainty reduction, offers a general, model-light diagnostic for detecting asymmetric updating, even in settings where the data-generating process is unobservable.

Together with the experimental results, this evidence supports the view that individuals and perhaps institutions correctly interpret the direction of new information but systematically misperceive its strength, leading to predictable asymmetries in belief updating across domains.

Chapter 5

Discussion

5.1 Discussion and Limitations

This thesis has examined how individuals update their beliefs in response to signals of varying strength, focusing on the empirically documented asymmetry where people tend to overreact to weak signals and underreact to strong ones. Building on the framework proposed by Augenblick et al. (2025), the analysis has explored the idea that this asymmetry stems not from directional bias, but from cognitive imprecision in estimating signal strength.

The main theoretical contribution of this thesis is the development of a simplified version of the original model. By abstracting away from more technical assumptions, such as specific parametric distributions or monotone likelihood ratio properties, this version retains the core mechanism: individuals form noisy internal estimates of signal strength and update conservatively using a weighted average with a default anchor. This stripped-down formulation shows that the characteristic single-crossing pattern in inference emerges naturally from basic structural constraints on cognition. The simplicity of the model enhances its tractability and generalizability, making it potentially useful for integration into broader behavioral economic models.

The empirical setup further supports the framework's relevance. Across multiple experimental environments, ranging from abstract card-draw tasks to intuitive sports prediction settings, belief updating behavior consistently exhibits the predicted asymmetry. Even in naturalistic and dynamic environments such as live basketball game scenarios, participants showed predictable patterns of over- and underinference. Notably, quasi-experimental evidence from financial markets and sports betting complements these findings. In both domains, aggregate belief movement displays the same distortion: excessive updating in response to weak, early signals, and muted reactions to stronger, later ones. These results suggest that the model captures not only individual cognition but also systematic deviations in real-world decision contexts.

The findings carry several implications. First, they provide further evidence that errors in belief updating are not merely noisy or idiosyncratic but exhibit structure rooted in cognitive processing. This has practical relevance for settings like financial forecasting, algorithmic design, or political communication, where users must interpret signals of varying strength and make sound decisions based on the signals. Platforms that present probabilistic information, such as polling aggregators or risk dashboards, may benefit from framing information in ways that reduce perceived strength ambiguity. Second, from a theoretical standpoint, the results highlight the importance of modeling belief formation not only in terms of bias direction but also in terms of perceived signal precision.

Despite these contributions, several limitations remain. The simplified model, while conceptually simpler, has not been formally estimated or tested using structural methods. As such, its quantitative predictions remain unverified. The assumption of a fixed adjustment parameter, α, may also oversimplify heterogeneity across individuals or contexts—different agents may anchor to different priors, or weigh evidence differently based on task familiarity, incentives, or cognitive resources. Moreover, although the experimental and quasi-experimental results align with the model qualitatively, they rely on inferred signal strength, direct measurement of perceived informativeness would better validate the proposed mechanism.

Future research could address these gaps by structurally estimating the model on individual-level data or by developing richer designs that elicit both posterior beliefs and perceived signal strength. Extensions that allow for endogenous anchor formation, dynamic learning about α, or context-dependent adjustment rules could increase both realism and explanatory power. Additionally, applying the model to high-stakes belief environments, such as climate science communication, misinformation dynamics, or algorithmic feedback systems, could enhance its practical and scholarly relevance.

In sum, this thesis reinforces the idea that belief formation is shaped not only by what people know, but by how precisely they perceive the strength of what they know. As such, it contributes to a growing literature on bounded rationality, structured misperception, and the cognitive underpinnings of inference.

Chapter 6

Conclusion

6.1 Conclusion

This thesis has investigated how individuals update beliefs under uncertainty when faced with signals of varying strength. Drawing inspiration from the framework developed by Augenblick et al. (2025), it has contributed a simplified and transparent model that isolates a core cognitive mechanism: individuals form noisy internal estimates of signal informativeness and update conservatively by anchoring to an intermediate default. This minimalist approach shows that the asymmetric pattern of overreacting to weak signals and underreacting to strong ones emerges not from specific structural assumptions, but from a basic and generalizable constraint—cognitive imprecision in strength perception.

The thesis makes several contributions to the literature on behavioral economics and belief formation. First, it clarifies the psychological foundation of asymmetric inference, emphasizing that belief distortions often arise not from misreading the direction of evidence but from an inability to precisely perceive its magnitude. Second, it extends the empirical reach of the framework by synthesizing results from laboratory experiments, naturalistic tasks, and quasi-experimental data from financial and betting markets. This cross-environmental validation underscores the robustness of the core mechanism and strengthens its claim to generality.

For scholars of economic cognition, bounded rationality, and behavioral modeling, this thesis reinforces a key insight: errors in belief updating are not random but systematically structured by the way people process and simplify complex signals. It calls for a shift in focus from correcting directional biases to understanding how individuals internally represent and weigh probabilistic evidence. Models that incorporate noisy strength perception and conservative adjustment offer a promising path forward in explaining both individual decisions and aggregate patterns.

More broadly, the findings have practical implications for domains where beliefs must be formed from uncertain signals, whether in policy communication, investment decisions, algorithmic feedback systems, or public forecasting. The tendency to overweight weak signals may help explain phenomena like hype cycles, misinformation cascades, or excessive volatility in early-stage markets. Conversely, underreaction to strong signals may reflect missed opportunities or policy inertia in the face of compelling evidence.

This thesis has taken a step toward formalizing and understanding these tendencies. Future work can build on this foundation by estimating the model quantitatively, exploring heterogeneity in strength perception, and testing interventions that make informativeness more transparent. Ultimately, by connecting the structure of belief distortions to the structure of cognition, this work contributes to a broader research agenda: modeling human inference not as flawed, but as bounded, patterned, and intelligible.

Bibliography

Anja Achtziger, Carlos Alós-Ferrer, Sabine Hügelschäfer, and Marco Steinhauser. The neural basis of belief updating and rational decision making. *Social Cognitive and Affective Neuroscience*, 9(1):55–62, January 2014. ISSN 1749-5024, 1749-5016. doi: 10.1093/scan/nss099. URL https://academic.oup.com/scan/article/9/1/55/1673542.

Sandro Ambuehl and Shengwu Li. Belief updating and the demand for information. *Games and Economic Behavior*, 109:21–39, May 2018. ISSN 08998256. doi: 10.1016/j.geb.2017.11.009. URL https://linkinghub.elsevier.com/retrieve/pii/S0899825617302191.

Ned Augenblick and Matthew Rabin. Belief Movement, Uncertainty Reduction, and Rational Updating. *The Quarterly Journal of Economics*, 136(2):933–985, March 2021. ISSN 0033-5533, 1531-4650. doi: 10.1093/qje/qjaa043. URL https://academic.oup.com/qje/article/136/2/933/6127317.

Ned Augenblick, Eben Lazarus, and Michael Thaler. Overinference from Weak Signals and Underinference from Strong Signals. *The Quarterly Journal of Economics*, 140(1):335–401, January 2025. ISSN 0033-5533, 1531-4650. doi: 10.1093/qje/qjae032. URL https://academic.oup.com/qje/article/140/1/335/7821262.

Cuimin Ba, J. Aislinn Bohren, and Alex Imas. Over- and Underreaction to Information. *SSRN Electronic Journal*, 2024. ISSN 1556-5068. doi: 10.2139/ssrn.4274617. URL https://www.ssrn.com/abstract=4274617.

Kai Barron. Belief updating: does the 'good-news, bad-news' asymmetry extend to purely financial domains? *Experimental Economics*, 24(1):31–58, March 2021. ISSN 1386-4157, 1573-6938. doi: 10.1007/s10683-020-09653-z. URL https://www.cambridge.org/core/product/identifier/S1386415721012047/type/journal_article.

Daniel J. Benjamin, Matthew Rabin, and Collin Raymond. A Model of Nonbelief in the Law of Large Numbers. *Journal of the European Economic Association*, 14(2):515–544, 2016. ISSN 1542-4766. URL https://www.jstor.org/stable/43965314. Publisher: Oxford University Press.

Daniel J Benjamin, Don A Moore, and Matthew Rabin. Biased Beliefs About Random Samples: Evidence from Two Integrated Experiments. *National Bureau of Economic Research*, October 2017. doi: 10.3386/w23927.

Daniel J. Benjamin, Aaron L. Bodoh-Creed, and Matthew Rabin. Base-rate neglect: Foundations and implications. 2019.

Pedro Bordalo, Nicola Gennaioli, and Andrei Shleifer. Overreaction and Diagnostic Expectations in Macroeconomics. *The Journal of Economic Perspectives*, 2025.

Olivier Coibion and Yuriy Gorodnichenko. What Can Survey Forecasts Tell Us about Information Rigidities? *Journal of Political Economy*, 120(1):116–159, February 2012. ISSN 0022-3808, 1537-534X. doi: 10.1086/665662. URL https://www.journals.uchicago.edu/doi/10.1086/665662.

Benjamin Enke and Thomas Graeber. Cognitive Uncertainty. *The Quarterly Journal of Economics*, 138(4):2021–2067, September 2023. ISSN 0033-5533, 1531-4650. doi: 10.1093/qje/qjad025. URL https://academic.oup.com/qje/article/138/4/2021/7181327.

Daniel Kahneman. *Thinking, Fast and Slow*. Farrar, Straus and Giroux, October 2011.

Mel Win Khaw, Ziang Li, and Michael Woodford. Cognitive Imprecision and Small-Stakes Risk Aversion. *The Review of Economic Studies*, 88(4):1979–2013, July 2021. ISSN 0034-6527, 1467-937X. doi: 10.1093/restud/rdaa044. URL https://academic.oup.com/restud/article/88/4/1979/5880002.

Yea-Seul Kim, Paula Kayongo, Madeleine Grunde-McLaughlin, and Jessica Hullman. Bayesian-Assisted Inference from Visualized Data, August 2020. URL http://arxiv.org/abs/2008.00142. arXiv:2008.00142 [cs].

Hannes Mohrschladt, Maren Baars, and Thomas Langer. Belief Updating Beyond the Two-State Setting. *Management Science*, 70(10):6761–6777, October 2024. ISSN 0025-1909, 1526-5501. doi: 10.1287/mnsc.2022.00513. URL https://pubsonline.informs.org/doi/10.1287/mnsc.2022.00513.

Pietro Ortoleva. Alternatives to Bayesian Updating. *Annual Review of Economics*, 16(1):545–570, August 2024. ISSN 1941-1383, 1941-1391. doi: 10.1146/annurev-economics-100223-050352. URL https://www.annualreviews.org/content/journals/10.1146/annurev-economics-100223-050352.

Lawrence D. Phillips and Ward Edwards. Conservatism in a simple probability inference task. *Journal of Experimental Psychology*, 72(3):346, 1966. doi: 10.1037/h0023653.

Matthew Rabin and Dimitri Vayanos. The Gambler's and Hot-Hand Fallacies: Theory and Applications. *Oxford University Press*, 2010.

Peibei Shi and Annie Qu. Weak signal identification and inference in penalized model selection. *The Annals of Statistics*, 45(3), June 2017. ISSN 0090-5364. doi: 10.1214/16-AOS1482. URL https://projecteuclid.org/journals/annals-of-statistics/volume-45/issue-3/Weak-signal-identification-and-inference-in-penalized-model-selection/10.1214/16-AOS1482.full.

Eric Vansteenberghe. Distinguishing Strong from Weak Signals in Economic Forecasts, November 2024. URL http://arxiv.org/abs/2411.05938. arXiv:2411.05938 [econ].

Michael Woodford. Modeling Imprecision in Perception, Valuation, and Choice. *Annual Review of Economic*, 2020.