Overinference from Weak Signals and Underinference from Strong Signals

Ned Augenblick
Berkeley Haas

Eben Lazarus
Berkeley Haas

Michael Thaler
UCL

September 2024

Background

How does new information change beliefs? Conflicting evidence across context contexts

Often categorized into over- and underinference (generating over- and underreaction)

1. **Financial markets**
   - **Excess volatility** in aggregate valuations: Appears consistent with overreaction
   - **Post-earnings drift** for individual stocks: Appears consistent with underreaction
   - ...

$\Rightarrow$ **lots of theories with context-specific biases**

2. **The lab**
   - Depends on the task. . .but in common tasks, Benjamin (2019):

     "**Stylized Fact 1:** Underinference is by far the dominant direction of bias."

Want to reconcile this evidence & offer a simple theory of updating behavior across settings.

This Paper

What drives over- vs. underinference?

- We argue a key mediating factor is the **strength of information**
- Experiments overwhelmingly study **strong** signals
- Outside the lab, constant stream of **weakly** informative signals about future events
  - New poll about politician's favorability. . .Pr(reelection)
  - **Daily index return**. . .**Pr(annual return > x)**
  - **First minutes of a basketball game**. . .**Pr(win)**
  - Important counterexamples: firm earnings, **last-minute baskets, interest-rate surprises**, . . .

Our story: **Underinference from strong signals, overinference from weak signals**

Basic Intuition

Our broad idea:

- In many settings, people receiving info understand its *direction*, but not its *strength*
  - Good poll numbers: $\nearrow$ Pr(win). . .but how much?
  - Basket: $\nearrow$ Pr(win). . .but how much?
  - Earnings beat expectations: $\nearrow$ fundamental value. . .but how much?
- Can understand if info is good/bad, but must generate noisy estimate of strength
- Shrinkage to moderate strength: after good news, shade toward "average" good news
  - Extreme case: Update the same for all good news

$\Longrightarrow$ relative to full Bayesian, average overreaction to weak, underreaction to strong

What We Do

Systematic study of reaction to information of different strengths:

1. Theory
2. Experiments
   (a) Abstract setting: easy to vary signal strength, but concerns about external validity
   (b) Naturalistic: give people less abstract situations (basketball games) where strength varies
3. Observational evidence: (a) sports betting & (b) option markets
   - Use indirect measures of over-/underreaction, and time to expiration as proxy for signal strength

- Then a few conjectures (+ some evidence) on more general AP/forecasting applications
  - Equity-premium predictions: Overreaction to news [Gandhi, Gormsen, Lazarus 2023]
  - Interest-rate predictions: Underreaction to news [Farmer, Nakamura, Steinsson 2023]
  - Brief discussion in paper, but more to be done on these extensions

Contribution

Underinference:

- Dominant direction in balls-and-urns experiments (Benjamin 2019; Edwards 1968; Griffin & Tversky 1992; Enke & Graeber 2023)
  - But little evidence for weak signals
  - One recent paper does, confirms patterns we find (Ba, Bohren, & Imas WP)
- Also: Neglect signal quantity or setting (Griffin & Tversky 1992; Massey & Wu 2005)

Overinference:

- More common in observational data (e.g., Bordalo et al. 2019)
  - Our argument: Environments with weakly informative signals
- Also common in forecasting experiments, varying horizon (Afrouzi et al. 2022; Fan et al. WP)
- Finance literature on overreaction and excess volatility (e.g. Shiller 1981)

We unite these strands: **underinfer from strong signals** and **overinfer from weak signals**:

- Strong: More common in lab. Weak: More common in forecast revisions.
- Other complementary mechanisms: Fan et al. WP; Kwon and Tang WP

Combine experiment with betting + finance data for external validity (Levitt & List 2006)

- Use tools from Augenblick & Rabin (2021), Augenblick & Lazarus (WP)
- ID settings in which signals strong vs. weak by using time to resolution

Outline

1. Background
2. Theory
3. Experimental Evidence
4. Empirical Evidence & Implications
5. Conclusions

Theory: Setup

- Person sees signal s about binary state $\theta \in \{0,1\}$. Objective prior $\pi_0 \equiv \mathbb{P}(\theta = 1)$.
- How would a full Bayesian update? $LR_{post} = LR_{prior} \times LR_{signal}$, or

$$\frac{\pi_1}{1-\pi_1} = \frac{\pi_0}{1-\pi_0} \times \frac{\mathbb{P}(s|\theta=1)}{\mathbb{P}(s|\theta=0)}$$

$$\log\left(\frac{\pi_1}{1-\pi_1}\right) = \log\left(\frac{\pi_0}{1-\pi_0}\right) + \log\left(\frac{\mathbb{P}(s|\theta=1)}{\mathbb{P}(s|\theta=0)}\right)$$

Theory: Setup

- Person sees signal s about binary state $\theta \in \{0,1\}$. Objective prior $\pi_0 \equiv \mathbb{P}(\theta = 1)$.
- How would a full Bayesian update? $LR_{post} = LR_{prior} \times LR_{signal}$, or

$$\log\left(\frac{\pi_1}{1-\pi_1}\right) = \log\left(\frac{\pi_0}{1-\pi_0}\right) + \log\left(\frac{\mathbb{P}(s|\theta=1)}{\mathbb{P}(s|\theta=0)}\right)$$

- Separate out: (1) *Direction* is $\pm$: $sign\left(\log\left(\frac{\mathbb{P}(s|\theta=1)}{\mathbb{P}(s|\theta=0)}\right)\right)$. (2) *Strength*: $\mathbb{S} \equiv \left|\log\left(\frac{\mathbb{P}(s|\theta=1)}{\mathbb{P}(s|\theta=0)}\right)\right|$.

$$\log\left(\frac{\pi_1}{1-\pi_1}\right) = \log\left(\frac{\pi_0}{1-\pi_0}\right) \pm \mathbb{S}$$

- Behav. assumption: Person knows direction, but unsure about strength $\mathbb{S}$
  - Easy to see "good" vs. "bad," but strength requires fully understanding DGP
  - People have idea of when signals are more informative, but still uncertain (and know this)
- Updates as if strength is $\hat{\mathbb{S}}$. Say *overinference* if $\mathbb{E}[\hat{\mathbb{S}}|s] > \mathbb{S}$, *underinference* if $\mathbb{E}[\hat{\mathbb{S}}|s] < \mathbb{S}$.

Theory: General Updating Assumptions

1. For signal s with direction s_d, the person forms an **estimate** e about strength $\mathbb{S}$ such that:
   (a) e is unbiased: $\mathbb{E}[e|\mathbb{S}] = \mathbb{S}$
   (b) e is well-ordered (satisfies MLRP): $\frac{\mathbb{P}(e\,|\,\mathbb{S}=\mathsf{S}')}{\mathbb{P}(e\,|\,\mathbb{S}=\mathsf{S})}$ strictly increases in e for all $\mathsf{S}' > \mathsf{S}$
   (c) e is imperfect: there is no pair $(e, \mathbb{S})$ s.t. $P(\mathbb{S}|e) = 1$
2. Prior expectation of signal strength given direction is $\hat{\mathbb{S}}(s_d)$. Need not be perfect, but must satisfy $\min \mathbb{S}(s) < \hat{\mathbb{S}}(s_d) < \max \mathbb{S}(s)$.
3. The person updates toward the estimate: posterior $\hat{\mathbb{S}}$ lies strictly between prior $\hat{\mathbb{S}}(s_d)$ and estimate e.

Proposition (*Overinference from Weak Signals, Underinference from Strong Signals*)

There exists a unique switching point $\mathbb{S}^*$ such that the person overinfers from s ($\mathbb{E}[\hat{\mathbb{S}}|s] > \mathbb{S}$) if $\mathbb{S}(s) < \mathbb{S}^*$ and underinfers from s if $\mathbb{S}(s) > \mathbb{S}^*$.

Idea of proof: Express $\mathbb{E}[\hat{\mathbb{S}}|s] - \mathbb{S}(s)$ as expectation of single-crossing function w.r.t. MLRP estimate distribution. Variation-diminishing property (Karlin 1968) implies expectation is also single-crossing.

Note that this setup encompasses quasi-Bayesian updating given estimate, but doesn't require it.

Theory: Parametric Version

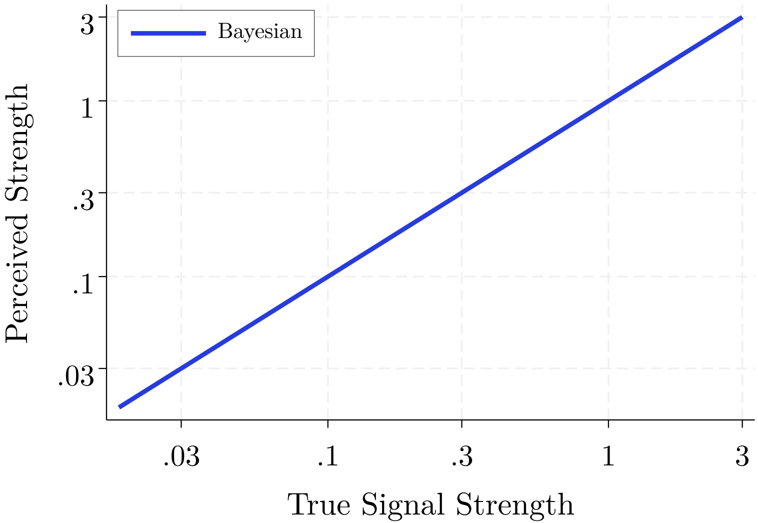
Useful for empirics to parameterize strength & estimate distributions

- **True signal strength**: $\log \mathbb{S} \sim \mathcal{N}(\mu_{\mathbb{S}}, \sigma^2_{\mathbb{S}})$
- **Given** $\mathbb{S}$, person gets estimate $e \sim \mathcal{N}\left(\log \mathbb{S} - \frac{\sigma_e^2}{2}, \sigma_e^2\right)$, then (quasi-)Bayesian updating on $\mathbb{S}$

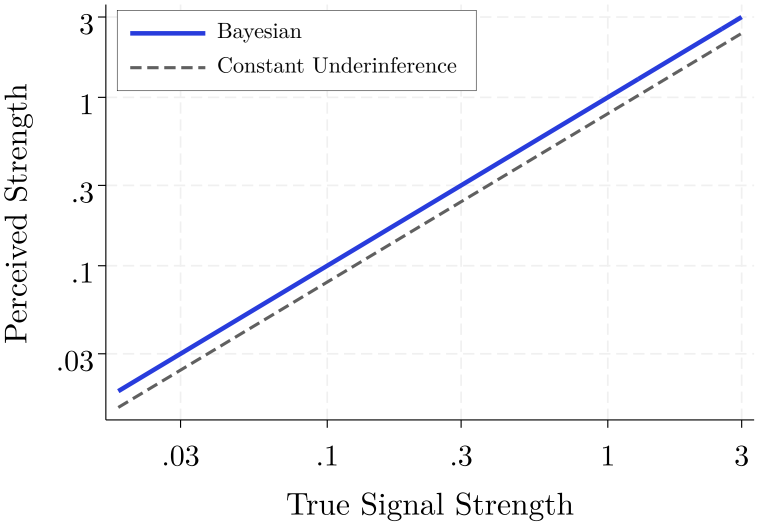
$\Longrightarrow$ **Posterior mean** of perceived $\mathbb{S}$ given e: power law $\mathbb{E}[\hat{\mathbb{S}}|s] = \boldsymbol{k}\mathbb{S}^{\boldsymbol{\beta}}$, with $\beta \equiv \frac{\sigma^2_{\mathbb{S}}}{\sigma^2_{\mathbb{S}}+\sigma^2_e} \in (0,1)$

- Again get shrinkage to "moderate" signal strength
- **Switching point** $\mathbb{S}^* \equiv k^{1/(1-\beta)}$ s.t. person underestimates $\mathbb{S}$ on average iff $\mathbb{S} > \mathbb{S}^*$

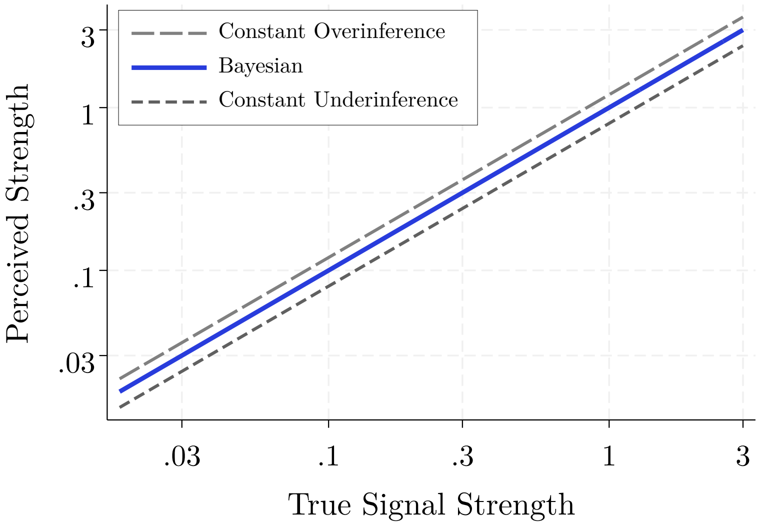
Perceived vs. Actual Signal Strength



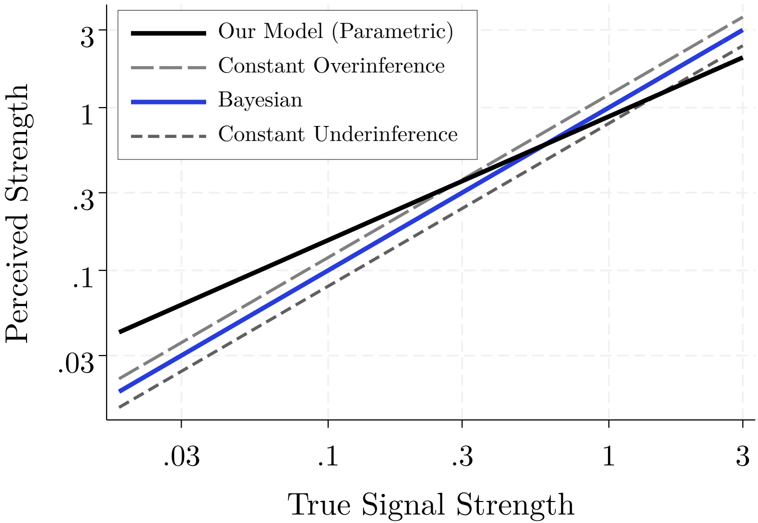
Perceived vs. Actual Signal Strength



Perceived vs. Actual Signal Strength



Perceived vs. Actual Signal Strength



Effective Weight on Signal

- Another way to think about it. Recall Bayesian

$$\log\left(\frac{\pi_1}{1-\pi_1}\right) = \log\left(\frac{\pi_0}{1-\pi_0}\right) \pm \mathbb{S}$$

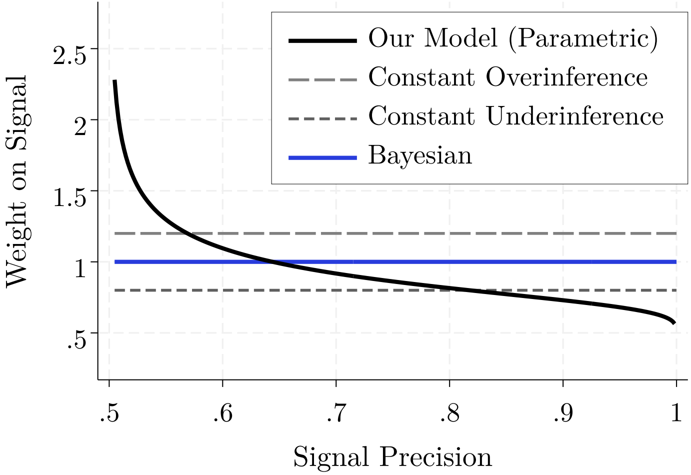
- Common way to parameterize over-/underinference: $\hat{\mathbb{S}} = w \cdot \mathbb{S}$:

$$\log\left(\frac{\pi_1}{1-\pi_1}\right) = \log\left(\frac{\pi_0}{1-\pi_0}\right) \pm w \cdot \mathbb{S}$$

  - $w = 1$: **Bayesian ($w = 1$)**
  - $w > 1$: **Overinference**
  - $w < 1$: **Underinference**
- Past papers: w is a constant.
- Us: w depends on $\mathbb{S}$. $w > 1$ when $\mathbb{S} < \mathbb{S}^*$ and $w < 1$ when $\mathbb{S} > \mathbb{S}^*$

Over- and Underinference by Signal Strength

- y-axis: weight $w(p)$ put on signal of precision p relative to Bayesian (here $\mathbb{S} = \log\left(\frac{p}{1-p}\right)$)



Discussion, Open Questions, and Extensions

Where does estimate come from?

- Simplified model of difficult-to-understand DGP
  - DGP could be difficult to understand for reasons of numerical complexity (as in experiments) or because extremely hard to figure out meaning of signals (as in reality)
- Limited processing: only attend to parts of signal
  - Extension in paper considers this more formally, and considers average behavior across multiple people focusing on different components of the same signal

Other extensions in paper: Incorrect priors, base-rate neglect, uncertainty about direction.

Outline

Experiment 1 in Brief

Standard bookbag-and-poker-chips setup [1,000 U.S. adults total on `prolific.co`, preregistered]

- Quite abstract, but can directly control DGP (and will use quantitative estimates later)

- Two card decks: Green & Purple
- Two card suits: diamond $\diamondsuit$ & spade $\spadesuit$

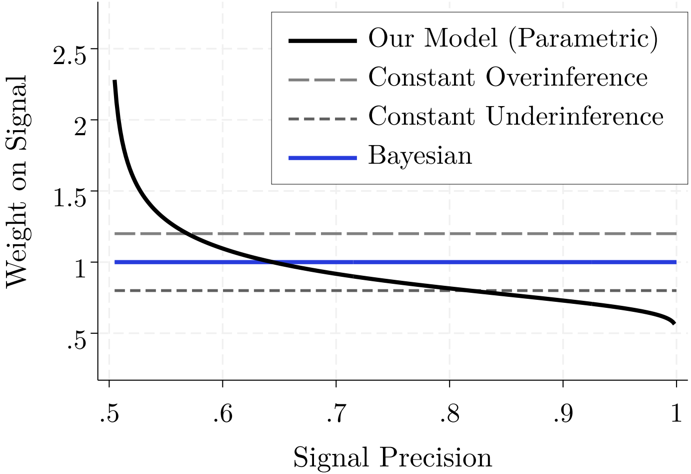
- Green has more diamonds: n_1 $\diamondsuit$ and n_2 $\spadesuit$ $(n_1 > n_2)$
- Purple has more spades: n_2 $\diamondsuit$ and n_1 $\spadesuit$

- One deck chosen (**equally likely for now**) and one card drawn, showing suit but not color
- Elicit $\mathbb{P}(\text{deck} \mid \text{suit drawn})$, with answers incentivized (binarized scoring rule) Example screen
- Example: it's a $\diamondsuit$: Green is more likely, but how much?

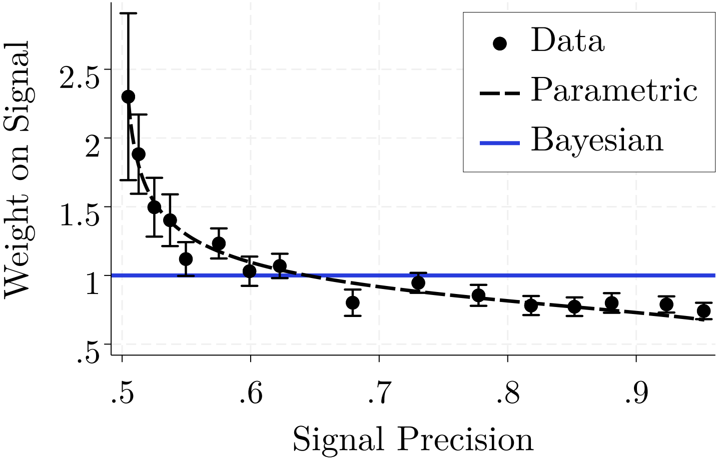
- Use non-round deck size $N = n_1 + n_2$ to make magnitude unclear
- How does updating vary with signal strength $(\frac{n_1}{N})$?

Over- and Underinference by Signal Strength: Theory

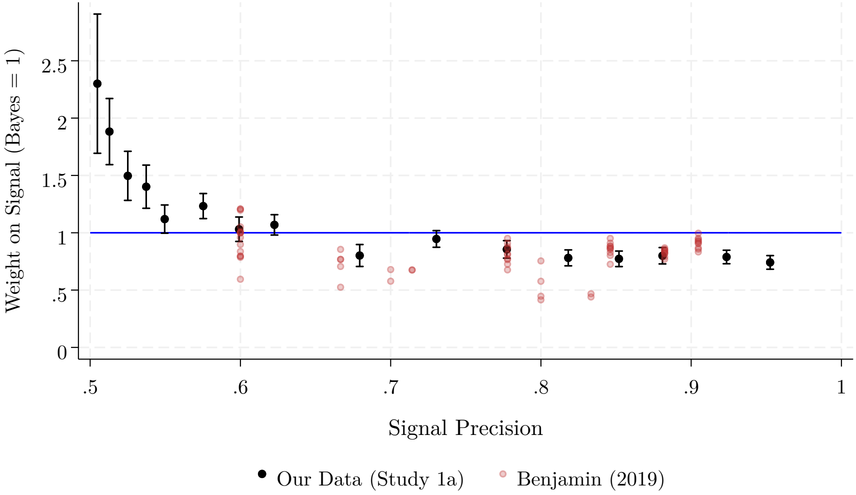
- y-axis: weight $w(p)$ put on signal of precision p relative to Bayesian (here $\mathbb{S} = \log\left(\frac{p}{1-p}\right)$)



Over- and Underinference by Signal Strength: Data

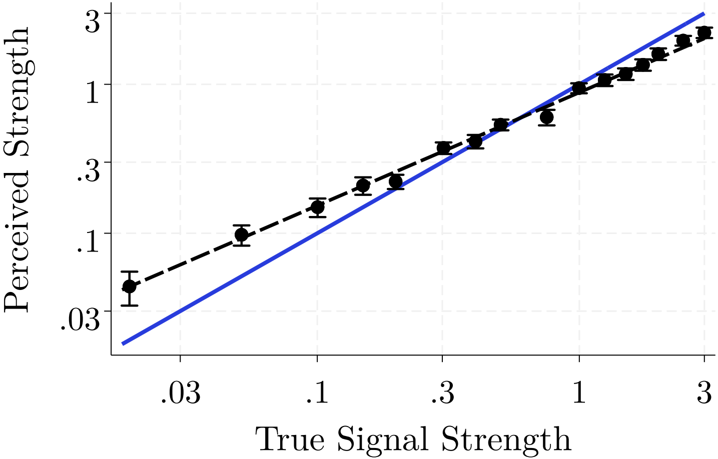


Over- and Underinference by Signal Strength: Comparison to Lit.



- **Estimates on medium/strong signals similar to 70 comparable arms from Benjamin (2019)**

People Are Responsive to Signal Strength...Just Not Enough



Model Estimates & Other Tests

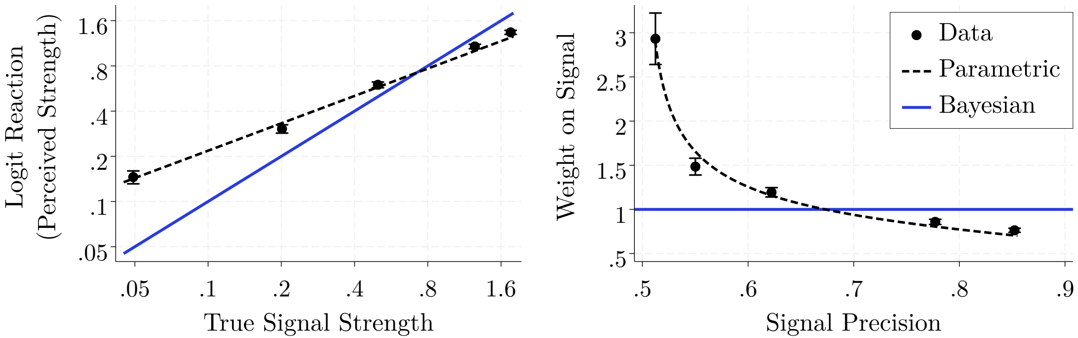
Recall: model predicts people perceive signal strength $\mathbb{S} \equiv \left|\log\left(\frac{P(s|\theta=1)}{P(s|\theta=0)}\right)\right|$ as strength $\hat{\mathbb{S}} = k\mathbb{S}^{\beta}$

- Nonlinear least squares estimate: $k = 0.88$ (s.e. 0.02); $\beta = 0.76$ (s.e. 0.03)
- Switching point $\mathbb{S}^*$ is such that $\hat{\mathbb{S}}^* = \mathbb{S}^*$, occurs when signal precision $P(s_H|\theta = 1) = 0.64$:
  - Signals with precision $> 0.64 \rightarrow \hat{\mathbb{S}} < \mathbb{S}$, i.e. underinference
  - Signals with precision $< 0.64 \rightarrow \hat{\mathbb{S}} > \mathbb{S}$, i.e. overinference
- Explains why hard to detect overinference when looking at signals with precision $\geqslant 0.6$

Experiment 1b: Replication Study

- Previous study started with priors on decks of 50/50
- What if prior is different?
- Basically same study but more decks:
  - 1 Green and 1-3 other (same composition) Blue-shaded decks
  - Prior on Green then 25%, 33%, 50%
  - Draw card, "What is likelihood Green deck was chosen?"
- Find same main effects (next slide), including when controlling for weight on prior Regressions

Experiment 1b: Replication with Different Priors Over Decks



Additional Analysis and Robustness

In paper:

- Distributions: FOSD in signal strength, but lots of heterogeneity Heterogeneity
- Formal tests Regressions
  - Heterogeneity tests: Estimation noise, inexperience, cog. uncertainty Heterogeneity regressions
- Willingness to pay to acquire new information WTP
  - Too high for weak signals
  - Too low for strong signals
- Inference from multiple signals Multiple signals
- Inference from asymmetric signals Asymmetric details
- Additional tests of potential alternative explanations Robustness details

Naturalistic Experiment

- Experimental evidence matches our predictions. . .but clear questions:
  - How does this relate to empirical evidence outside the lab?
  - Are findings just about numerical (rather than substantive) complexity of updating task?
- Challenge: need DGP that people understand and some sense of correct answer
- Solution: recruit NBA fans and ask about basketball game situations
- Ex.: Suppose home team up 7 in 1st quarter. P(win)? Now, after other team scores 2: P(win)?
- Direction of updating is clear, but correct amount less so
- Will help bridge to observational data

Naturalistic Experiment

Details:

- 500 self-identified NBA fans on `prolific.co`
- Multiple scenarios: Made/missed basket, home/away team, point differential
- **Main variation:** Which quarter the event takes place in (1st/2nd/3rd/4th)
- Going from $2 \to 4$ point lead more meaningful w/ 2:00 left in the 4th than the 1st
- To estimate objective signal log likelihood, use 3rd party win prob. estimates
  (via `inpredictable.com`, estimated using $>$ decade of NBA games)

Experiment 2: Example Screen

Prediction

We will give you some more information about the game. With 2:25 left in the first quarter, Team B made a shot to cut the lead to 5 points.

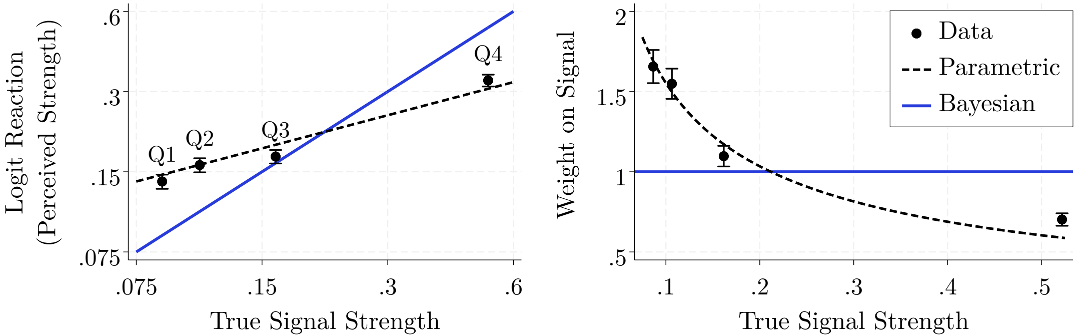
Time left	Last action	Score	Your guess
2:40	Team A made a shot	Team A is up by 7 points	Team A: **67%** chance of winning
2:25	**Team B made a shot**	**Team A is up by 5 points**	?

Now, what do you think is the percent chance that Team A wins the game?

0 10 20 30 40 50 60 70 80 90 100

62

Experiment 2: Main Results



- Takeaways: consistent results in non-math-like problems
- Concerns: Weakly incentivized, unclear external validity or applicability to finance settings

Outline

1. Background
2. Theory
3. Experimental Evidence
4. Empirical Evidence & Implications
   - Betting Markets
   - Index Options
   - Broader Implications
5. Conclusions

Evidence from Options & Betting Markets

Now consider market-implied probabilities in binary betting markets and index options.

New challenges:

1. Interpretation as "beliefs"?
2. No longer control DGP $\Longrightarrow$ true signal informativeness is unknown, as are correct beliefs

- Need to develop measure of over-/underreaction that doesn't require objective prob.
- And need an ex ante proxy for signal strength
- Key feature of setting we focus on:
  Fix expiration date and observe implied beliefs over time within contract
  - Will allow us to develop proxy for signal informativeness based on time to resolution

Belief Movement and Uncertainty Reduction

For belief π_t over binary state θ (e.g. "Warriors win," or "$R_{mkt,0\to T} \in [10\%, 15\%]$") realized at T:

- **Belief movement** (quadratic variation):
$$m_{t,t+1} = (\pi_{t+1} - \pi_t)^2$$
- **Uncertainty reduction** (decrease in subjective variance):
$$r_{t,t+1} = \pi_t(1 - \pi_t) - \pi_{t+1}(1 - \pi_{t+1})$$
- **Result** (Augenblick & Rabin 2021): For a Bayesian, for any DGP,
$$\mathbb{E}_t[m_{t,t+1}] = \mathbb{E}_t[r_{t,t+1}]$$
- Allows for agnostic test of Bayes' rule, with rejections having clean interpretation
  - For non-Bayesian, $\mathbb{E}[m] \neq \mathbb{E}[r]$
  - $\mathbb{E}[m] > \mathbb{E}[r] \iff$ too much movement, overreaction (& vice versa)
  - For $T = 2$, can prove formally that overinference $\implies \mathbb{E}_t[m] > \mathbb{E}_t[r]$. For $T > 2$, simulations.
- Crucial features: Valid regardless of DGP, for any subset of data (as long as cut is ex ante)

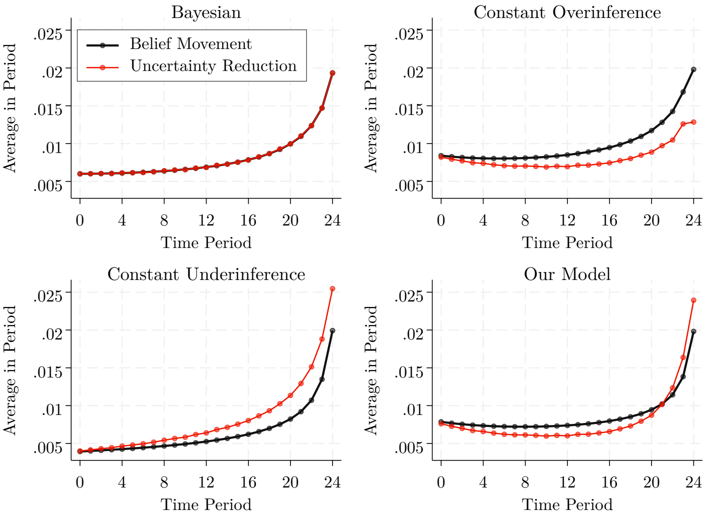
$\implies$ **measure of over/under-inference not requiring true beliefs**...

Belief Movement and Uncertainty Reduction

. . .**also need proxy for signal strength** (despite not knowing DGP)

- Theory says key mediating variable for over-/underinference is signal strength
- Need something that picks it up ex ante (not sorting on ex post movement), since $\mathbb{E}_t[m] = \mathbb{E}_t[r]$ holds only for conditional expectations
- We exploit the *fixed-event* feature of our data, and use **time to resolution** as our ex ante strength proxy:
  - First-quarter baskets less informative than fourth-quarter baskets (on avg)
  - Today's info is less informative about equity index value in a year than about tomorrow (on avg)
  - True in any standard option-pricing model (option delta decreases with time to resolution), in simple simulations (next slide), and quite apparent in our data

Simulated Belief Streams



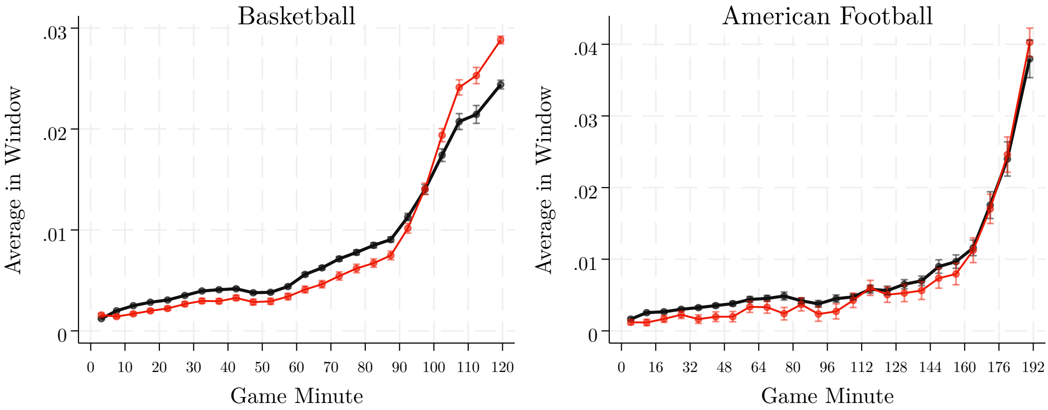
- Simulate one million binary option contracts where underlying follows random walk (± 1 each period, 50/50 chance) and state 1 realized if underlying > 0 at $T = 24$
- **Movement** (Black)
  Uncertainty reduction (Red)
- **Bayesian**: $m = r$
- **Overreaction**: $m > r$
- **Underreaction**: $m < r$
- **Our model** (exp. params): $m > r$ early and $m < r$ late

Data

1. Sports betting (Moskowitz 2021): Betfair
   - Large UK-based prediction exchange matching individual bettors on contracts over game outcome
   - Within-game binary bet prices for soccer, American football, baseball, basketball, hockey games

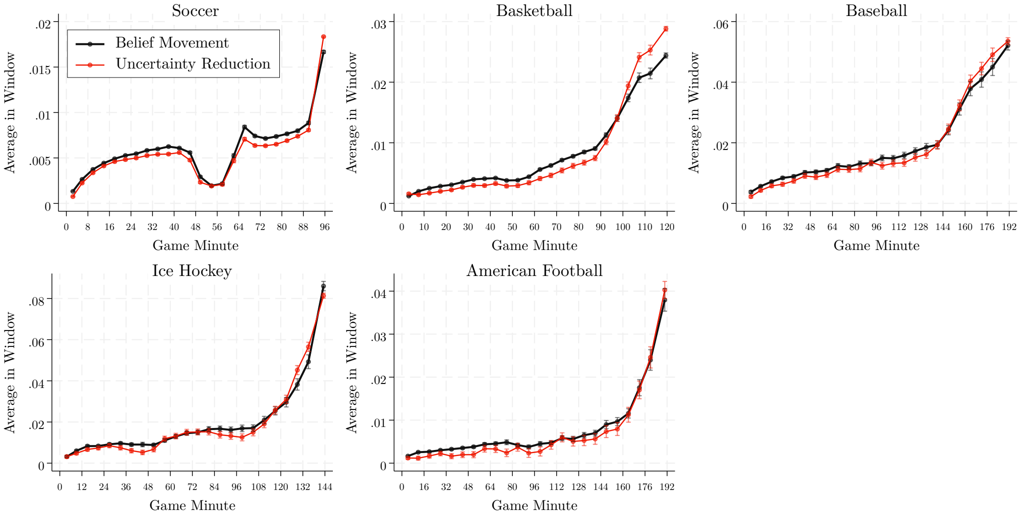
Betfair Data: NBA and NFL

Using implied probabilities from Betfair over the course of sports games:



Over course of game, $\boldsymbol{m}$ and $\boldsymbol{r}$ increase, and $\boldsymbol{m} - \boldsymbol{r}$ moves from $+$ to $-$.

Betfair Data: All Sports

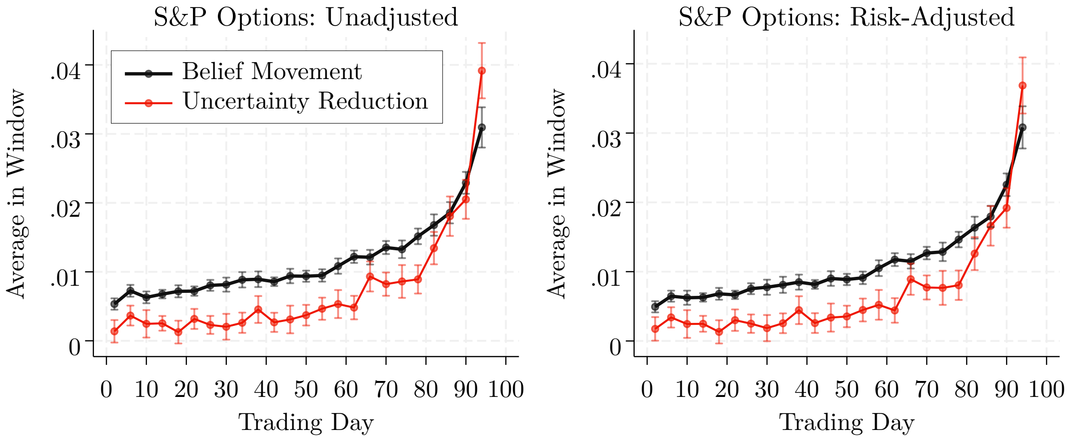


Over course of game, m and r increase, and $m - r$ moves from $+$ to $-$.

Data

1. Sports betting (Moskowitz 2021): Betfair
   - Large UK-based prediction exchange matching individual bettors on contracts over game outcome
   - Within-game binary bet prices for soccer, American football, baseball, basketball, hockey games
2. Risk-neutral beliefs from index options:
   - OptionMetrics: Daily S&P index option prices traded on CBOE, 1996–2018 (Augenblick & Lazarus 2023)
   - Idea: Options allow for bets on value of market at $T \Rightarrow$ their prices give an implied distribution (*risk-neutral* distribution) for value of market at T
   - Use histogram of risk-neutral beliefs for 5-percent return bins from 0 to $T = 100$ trading days, and then calculate m and r using daily belief changes for a given return bin (summed over all bins)
   - Augenblick & Lazarus (2023) show how risk-neutral beliefs are informative about excess vol... less important for our analysis, because looking for relative over-/underreaction within a contract
   - Also translate RN beliefs to **physical** beliefs under hundreds of parameterizations for risk aversion $\phi_{t,j}$ (nearly identical results). Main parameterization uses CRRA utility, estimating γ via ML s.t. resulting physical beliefs are unbiased on average

Options Data



As move closer to expiration, $\boldsymbol{m}$ and $\boldsymbol{r}$ increase, and $\boldsymbol{m} - \boldsymbol{r}$ moves from $+$ to $-$.

Empirical Evidence: Formal Regression Framework

- Our hypothesis:
  - Signals are weak $\implies \mathbb{E}[m] > \mathbb{E}[r]$
  - Signals are strong $\implies \mathbb{E}[m] < \mathbb{E}[r]$
- Sorting variable for signal strength: time to resolution
- Cut market-implied belief streams into **24** time windows
  - Paper has results for alternative choices
- Regress **average** movement on **average** uncertainty reduction within each window
  - If used individual observations, would have severe attenuation bias
  - Bayesian: $\mathbb{E}_t[m_{t,t+1}] = \mathbb{E}_t[r_{t,t+1}] \implies$ should observe intercept of 0, slope of 1
  - Our theory: positive intercept, slope below 1
  - Results align with these theoretical predictions

Empirical Evidence: Formal Regression Results

Regressions of Movement on Uncertainty Reduction

Dep Var:	Betting					Options	
Movement	Soccer	Basketball	Baseball	Hockey	Football	Raw	Risk-Adj.
Uncert. Red.	0.918	0.806	0.889	0.945	0.912	0.680	0.733
	(0.005)	(0.008)	(0.013)	(0.013)	(0.027)	(0.040)	(0.041)
Constant	0.0009	0.0018	0.0026	0.0018	0.0015	0.0065	0.0060
	(0.0001)	(0.0001)	(0.0002)	(0.0002)	(0.0002)	(0.0003)	(0.0003)
R^2	0.977	0.985	0.995	0.976	0.995	0.944	0.941
Time Chunks	24	24	24	24	24	24	24
Events	6,584	5,176	3,927	4,123	1,390	955	955
Observations	4,589,289	867,567	166,346	109,751	86,193	58,864	58,864
p-val.: $\beta_1 = 1$	<0.001	<0.001	<0.001	<0.001	0.007	<0.001	<0.001
p-val.: $\beta_0 = 0$	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001

Parentheses show bootstrapped standard errors with resampling clustered by stream.

- Overreaction when signal informativeness is low...but decreases systematically as signal strength increases, with underreaction close to expiration. 12 time chunks 36 time chunks

Broader Implications

How does our framework help us understand more general settings?

- Have studied over-/underreaction in (subjective or market-implied) **probabilities**...
- ...using **time-series variation** in signal strength **within** a given series
- For finance, often concerned with variation in (subjective or risk-neutral) **expectations**
- Claim: **Cross-sectional variation** in signal strength **across** processes helps us understand expectation puzzles
- See news, need to figure out how much to update forecast $\mathbb{E}_t[X_T]$
- Often clear whether news is good or bad, but less clear precisely how much to update
  - **Less persistent** series $\Longrightarrow$ innovations today are weak news about future $\Longrightarrow$ overreaction
  - **More persistent** series $\Longrightarrow$ innovations today are strong news about future $\Longrightarrow$ underreaction

    [relative to fully signal-understanding Bayesian, even if forecasts are (boundedly) rational]
- Next: Suggestive evidence from recent work

Forecasts of the Future Equity Premium

Gandhi, Gormsen, Lazarus (2023): "Forward Return Expectations"

- Measure and study **expected future equity premium** $\mathbb{E}_t[\mu_{t+n}]$
  - Using **options**
  - Using **surveys**: Equity return expectations at multiple horizons $\rightarrow \mathbb{E}_t[\mu_{t+n}]$
- Main finding: In all settings, $\mathbb{E}_t[\mu_{t+n}]$ is **excessively countercyclical** relative to realized future equity premium μ_{t+n}
- In bad times, $\mathbb{E}_t[\mu_{t+n}] \nearrow$ by more than **investors' own** stated equity premium at $t+n$ Figure
- **Interpretation:** Equity premium has low/moderate persistence $\Longrightarrow$ relatively weak news & overreaction
- Tests: Forecasted persistence ($\hat{\rho}$) exceeds persistence in spot expectations (ρ)

Forecasts of the Future Risk-Free Rate

Farmer, Nakamura, Steinsson (2023): "Learning About the Long Run"

- Measure and study forecast errors in **expected future 3-month T-bill yields** $\mathbb{E}_t[y_{t+n}]$
  - Using SPF survey responses
  - Separately consider GDP forecasts by CBO
- Forecast errors are positively autocorrelated Figure
- **Interpretation:** Risk-free rate has very high persistence $\implies$ relatively strong news & underreaction *(their interpretation: slow learning)*
- Tests: Forecasted persistence ($\hat{\rho}$) underestimates objective persistence (ρ)

Implied Persistence in Equity-Premium and T-Bill Forecasts

Regressions of Forecasted and Realized Values on Current Values

$$\mathbb{E}_t[\mu_{t+n}] = \alpha_{\text{subj}} + \hat{\rho}\,\mu_t + e_{t,\text{subj}}, \qquad \mu_{t+n} = \alpha_{\text{obj}} + \rho\,\mu_t + e_{t,\text{obj}}$$

	Eq. Prem. Forecasts		T-Bill Forecasts
	Options	Livingston Survey	SPF Survey
$\hat{\rho}$	0.65	0.68	0.84
	[0.02]	[0.08]	[0.04]
ρ	0.48	0.64	0.91
	[0.05]	[0.10]	[0.03]

- Overreaction for equity-premium forecasts, underreaction for yield forecasts
- Appears to be **partial** shrinkage to moderate persistence. Consistent with forecasters understanding directional impact of news, but not knowing precise strength.
- Speculative conjecture: Helps tie together evidence across decent range of settings. More to be done, and think this is one (of a few — see paper conclusion!) fruitful avenue for future work.

Outline

1. Background
2. Theory
3. Experimental Evidence
4. Empirical Evidence & Implications
5. Conclusions

Final Notes

Main finding: **Overreaction from weak signals & underreaction from strong signals**

- Robust & consistent evidence across domains:
  Abstract experiment, naturalistic experiment, sports betting, financial markets
  (each with own strengths and weaknesses)
- Basic theory: people know direction, unsure of magnitude $\implies$ shade toward average
  (doesn't apply everywhere, but we think common situation)
- Helps reconciles disparate findings in experimental lit. & finance lit.
- Future work: More to be done teasing out implications in dynamic settings,
  and lots of open questions

Appendix

Example Screen from Experiment 1a

Back

You draw a card from one of two modified decks of cards; a **Green** deck or a **Purple** deck.

The **Green** deck has **853 Diamonds** (♦) and **812 Spades** (♠).
The **Purple** deck has **812 Diamonds** (♦) and **853 Spades** (♠).

*The card you draw is a **Spade** (♠).*

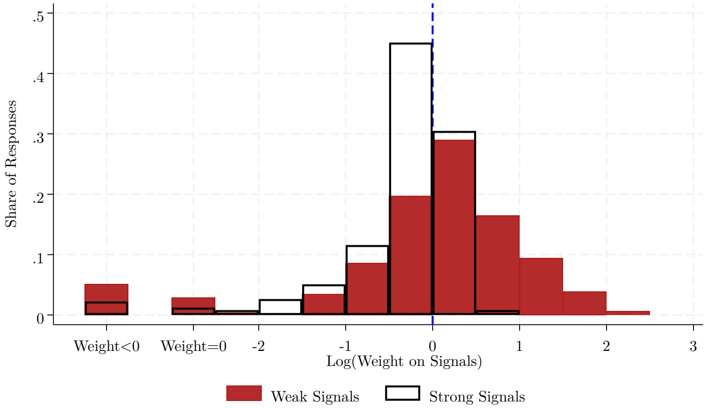
What do you think is the percent chance that your **Spade** (♠) came from the **Green** deck vs. the **Purple** deck?

Please answer between 0 and 100 for each question and have your answers sum up to 100, where higher numbers mean you think that deck is more likely. Your answer may include decimals.

Percent chance that your **Spade** (♠) came from the **Green** deck	0	percent
Percent chance that your **Spade** (♠) came from the **Purple** deck	0	percent
Total	0	percent

Vary deck size (small/**large**), suit (diamond/**spade**), signal strength (**weak**/strong)

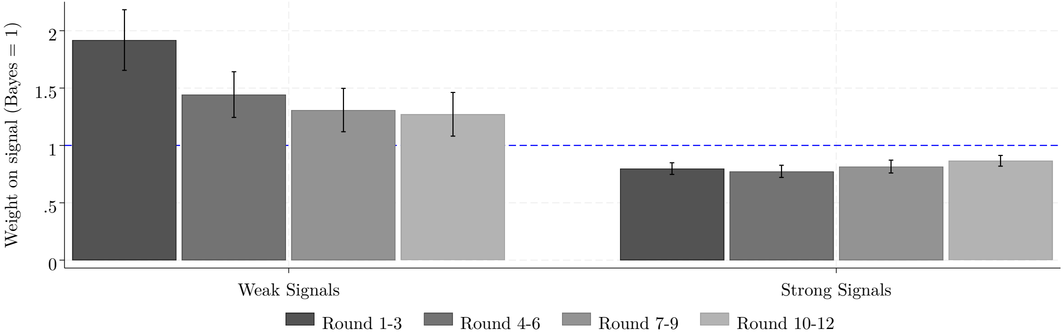
Distributions ◂ Back



- Distribution of weights for **weak signals** $\sim$FOSD distribution for **strong signals**
- But **lots of heterogeneity**. Theory makes predictions:
  - More strength estimation uncertainty $\rightarrow$ further from Bayes
- **Proxies:** task experience, cognitive reflection test, noisiness of answers

Heterogeneity and Learning: Experience

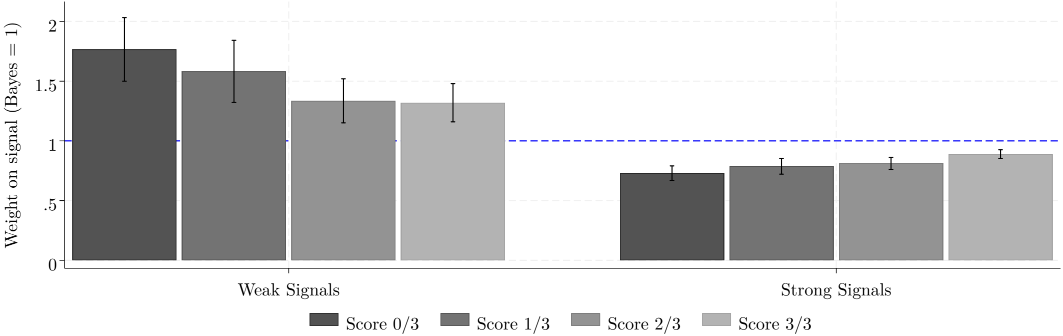
Back



- **Less task experience:** Infer more from weak signals and less from strong signals
  - No feedback between rounds
  - Suggests experience increases sophistication

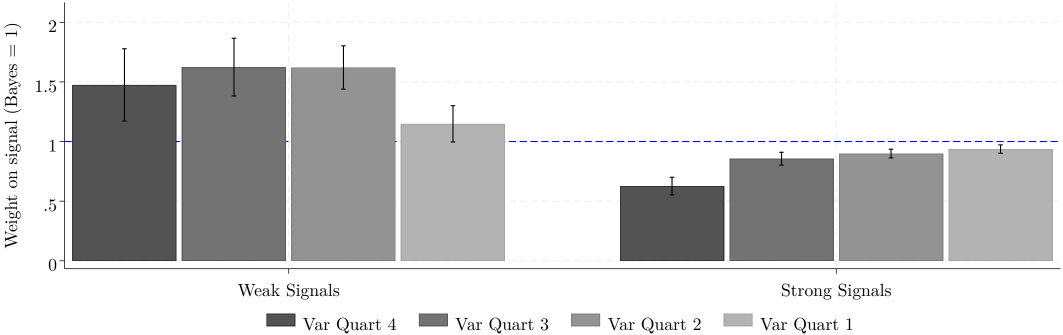
Heterogeneity and Learning: Cognition Task

Back



- **Lower score on cognitive reflection test** (Frederick 2005): Infer more from weak signals and less from strong signals

Heterogeneity and Learning: Noisy Inference ◂ Back



- **Group subjects by *noise*:** SD of answers given when signal strength happens to be same
- **Noisier answers:** Infer more from *other* weak signals, less from *other* strong signals

Overinference Effects: Regression Tests

Back

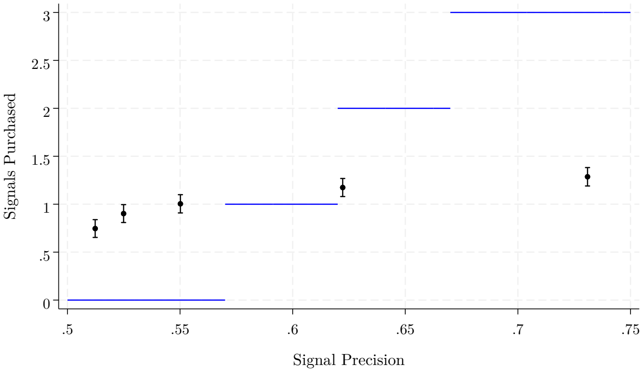
	Bayes	Theory	Study 1a	Study 1b		Study 2	
Dep. Var.: Weight on Signal			(1)	(2)	(3)	(4)	(5)
Constant	1	> 1	1.420	2.180	2.182	1.706	1.700
			(0.030)	(0.049)	(0.048)	(0.024)	(0.025)
Signal Strength	0	< 0	-0.308	-0.957	-0.958	-2.078	-2.060
			(0.031)	(0.065)	(0.065)	(0.111)	(0.112)
Weight on Prior	1				0.980		0.976
					(0.013)		(0.009)
Participant FE			Yes	Yes	Yes	Yes	Yes
Round FE			Yes	Yes	Yes	Yes	Yes
Observations			3964	7500	7500	8000	8000
R^2			0.23	0.16	0.17	0.24	0.24
p-val.: Const. $= 1$			<0.001	<0.001	<0.001	<0.001	<0.001
p-val.: Slope $= 0$			<0.001	<0.001	<0.001	<0.001	<0.001

Overinference Effects: Heterogeneity Tests

Back

	Study 1a			Study 1b		
Dep. Var.: Weight on Signal	(1)	(2)	(3)	(4)	(5)	(6)
Constant	1.395	1.416	1.421	2.147	2.183	2.181
	(0.021)	(0.030)	(0.030)	(0.037)	(0.048)	(0.048)
Strength	0.118	-0.072	-0.175	-0.002	-0.698	-0.756
	(0.031)	(0.051)	(0.040)	(0.070)	(0.089)	(0.120)
Strength × Noise	-0.383			-0.433		
	(0.036)			(0.046)		
Strength × Inexperience		-0.042			-0.037	
		(0.009)			(0.012)	
Strength × CRT Incorrect			-0.102			
			(0.028)			
Strength × Cog. Uncertainty						-0.542
						(0.288)
Weight on Prior				0.980	0.980	0.980
				(0.013)	(0.013)	(0.013)
Participant FE	Yes	Yes	Yes	Yes	Yes	Yes
Round FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3962	3964	3964	7500	7500	7500
R^2	0.28	0.24	0.23	0.20	0.17	0.17

Demand for Information Back



WTP bias similar to beliefs bias:

- Excess demand for weak signals
- Too little demand for strong signals
- Similar to Ambuehl and Li (2018)

Departure even more severe for WTP:

- One explanation: Subjects also undervalue multiple signals.
- Run main treatment with 3 signals and see this effect. Details

Alternative Explanations

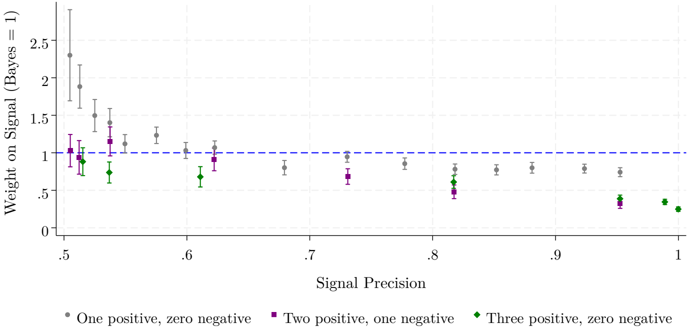
Back

- Misperceptions of probabilities rather than signal strengths (*relative* probabilities)?
  - E.g. Maybe people overestimate probabilities that are $1/2+\epsilon$ vs. $1/2-\epsilon$.
  - Test using **asymmetric** signals, where one signal is close to uninformative
  - $P(s|\theta=1) \gg 1/2$, but $P(s|\theta=0) = 1/2 + / - \epsilon$.
  - Whether + or - ϵ would affect agent who misperceives probabilities, but not agents in our model
  - Find little effect of changing + to - ϵ. Details

- Other explanations: Aversion to 50%, # of cards, preference for one deck or color, preference for round answers
  - No evidence these are driving results
  - E.g. 72 subjects see decks with P(Green) = 1/2 and 69 (96%) have posteriors of exactly 1/2

Underinference from Multiple Signals

◂ Back to main ◂ Back to WTP



Effect of signal strength similar across groups, but switching point changes:

- Gray: 0.644
- Purple: 0.523
- Green: 0.507

Consistent with past literature (Griffin and Tversky (1992) and others)

Asymmetric Signals

Back to main

- Signal in one state has likelihood far from 1/2:
  - Vary $P(s|\theta = 1) = 0.65$ or 0.80
- Signal in other state has likelihood close to 1/2:
  - Vary $P(s|\theta = 0) = 0.495$ or 0.505
- Estimate β from power law model for each $P(s|\theta = 0)$
- If effects due to probabilities, would expect significantly higher β when $P(s|\theta = 0) = 0.505$
- Instead: $\beta = 0.54$ when $P(s|\theta = 0) = 0.505$ and $\beta = 0.60$ when $P(s|\theta = 0) = 0.495$
  - Difference small, in the opposite direction, and not statistically significant (p-value $= 0.404$)

Robustness: 12 Time Periods Back

Regressions of Movement on Uncertainty Reduction

Dep Var:	Sports					Finance	
Movement	Soccer	Basketball	Baseball	Hockey	Football	Raw	Risk-Adj.
Uncert. Red.	0.839	0.797	0.903	0.987	0.912	0.796	0.861
	(0.006)	(0.007)	(0.012)	(0.012)	(0.027)	(0.054)	(0.063)
Constant	0.0014	0.0018	0.0024	0.0013	0.0015	0.0060	0.0054
	(0.0003)	(0.0003)	(0.0004)	(0.0009)	(0.0002)	(0.0005)	(0.0005)
R^2	0.984	0.991	0.996	0.990	0.997	0.945	0.941
Time Chunks	12	12	12	12	12	12	12
Events	6,584	5,176	3,927	4,123	1,390	955	955
Observations	4,589,289	867,567	166,346	109,751	86,193	58,864	58,864
p-val.: $\beta_1 = 1$	<0.001	<0.001	<0.001	0.274	0.002	0.004	0.025
p-val.: $\beta_0 = 0$	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001

Parentheses show bootstrapped standard errors with resampling clustered by stream.

- Overreaction when signal informativeness is low. . .but decreases systematically as signal strength increases, with underreaction close to expiration

Robustness: 36 Time Periods Back

Regressions of Movement on Uncertainty Reduction

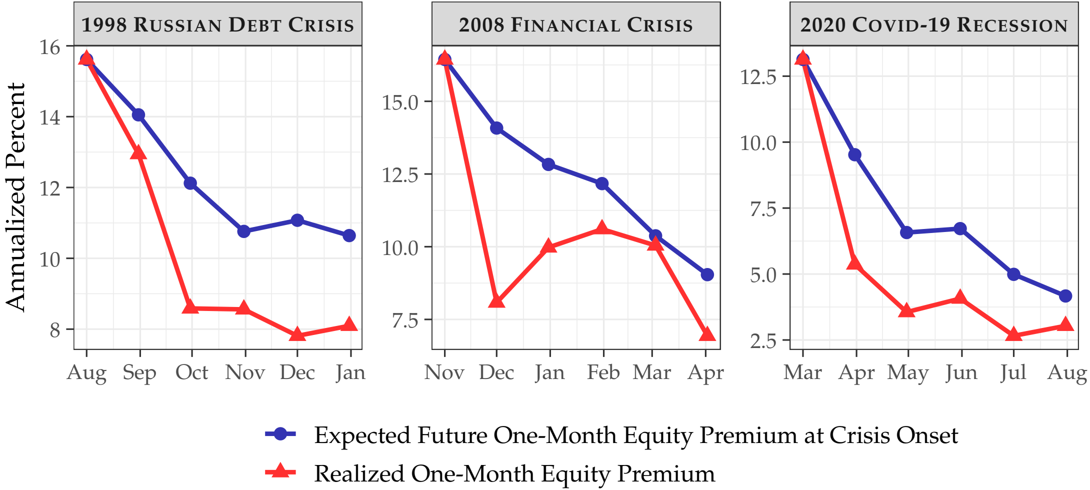
Dep Var:	Sports					Finance	
Movement	Soccer	Basketball	Baseball	Hockey	Football	Raw	Risk-Adj.
Uncert. Red.	0.847	0.849	0.883	0.925	0.920	0.705	0.751
	(0.003)	(0.008)	(0.015)	(0.013)	(0.026)	(0.035)	(0.040)
Constant	0.0014	0.0016	0.0027	0.0020	0.0015	0.0063	0.0058
	(0.0001)	(0.0001)	(0.0002)	(0.0002)	(0.0001)	(0.0003)	(0.0003)
R^2	0.955	0.974	0.993	0.975	0.982	0.932	0.928
Time Chunks	36	36	36	36	36	36	36
Events	6,584	5,176	3,927	4,123	1,390	955	955
Observations	4,589,289	867,567	166,346	109,751	86,193	58,864	58,864
p-val.: $\beta_1 = 1$	<0.001	<0.001	<0.001	<0.001	0.054	<0.001	<0.001
p-val.: $\beta_0 = 0$	<0.001	<0.001	<0.001	<0.001	0.051	<0.001	<0.001

Parentheses show bootstrapped standard errors with resampling clustered by stream.

- Overreaction when signal informativeness is low. . .but decreases systematically as signal strength increases, with underreaction close to expiration

Illustration: Expected and Realized Equity Premia in Crises

Back



. . . and pattern holds systematically in both options & surveys (Livingston & CFO).

Expected and Realized Risk-Free Yields ◂ Back

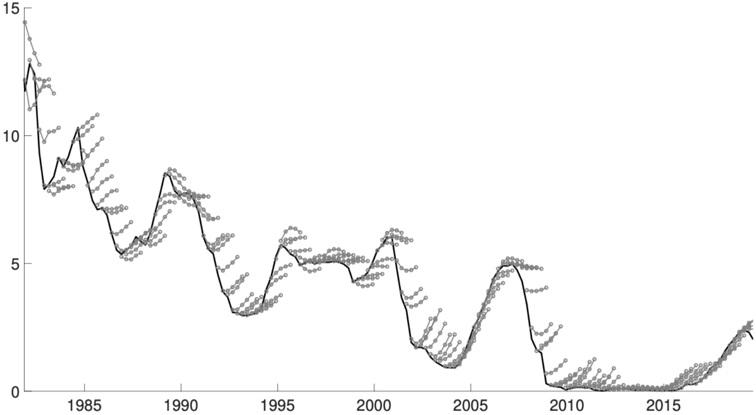


Figure 1: SPF Forecasts of the 3-Month T-Bill Rate

Note: The black solid line is the 3-month T-bill rate. Each short gray line with five circles represents the SPF forecasts made in a particular quarter about the then present quarter (first circle) and following four quarters (subsequent four circles).

- Figure from Farmer, Nakamura, and Steinsson (2023)