Croyances Individuelles : Rationalité, Subjectivité et Motivations Psychologiques

Roland Bénabou

Princeton University

Cours au Collège de France - Nov.-Déc. 2017

(Basé en partie sur des travaux joints avec Jean Tirole)

I - How Do People Form Their Beliefs?

1. Backward-looking expectations, naive extrapolation, adaptive learning

2. Rational expectations in macro, Bayesian equilibrium & refinements in micro

3. Behavioral: wired-in "biases and heuristics": confirmation bias, availability bias, salience, base-rate neglect, hot hand fallacy, probability weighting...

4. Motivated beliefs, cognition, reasoning: forms of self-deception

   ▸ Held (or more likely to be) due to emotional or functional value

   ▸ Resistant to evidence, but respond to costs, benefits and stakes

   ▸ Other telltale signs:
     – Information aversion: willful blindness. Heat vs. light
     – Selective attention, retrieval, or memory
     – Neural signatures?

Motivated cognition / thinking / beliefs

- About the self:
  - Talent, intelligence, willpower, beauty, morality
  - Future prospects: rich vs. poor, healthy vs. sick, happy vs. unhappy
  - Identity (where do I belong? what are my values, goals?)

- About how the world works:
  - Causes of inequality (effort vs. luck), social mobility, "Belief in Just World"
  - Ideology, e.g. merits of state vs. market, proper scope of government
  - What is moral or immoral, "taboo"
  - Other people: trust, in-group / out-group stereotypes
  - Religion, culture

- Much evidence that often not formed and revised in a neutral, objective manner, but in part to serve important "needs"
  - Purely psychological, consumption value
  - Functional, instrumental value

  ⇒ Beliefs as assets that people invest in, value, consume, defend, repair, etc.

Beliefs and misbeliefs: some examples

- Much apparent overoptimism, overconfidence, "better-than-average effect": driving, intelligence, sense of humor, likelihood of good / bad life events, etc.

  - But such snapshots of reported beliefs (or even elicited, from choices with real stakes), e.g. 75% think above average, 2/3 think are in top 1/3, etc., may in fact be consistent with rational, Bayesian model

  - Depends what signals people have received: we don't know what they know. Need much more stringent tests (Benoit-Dubra 2011, Merkle-Weber 2011)

  - More convincing: study process by which beliefs are formed and come to be distorted. Relate to "demand" and "supply" -side predictors

- Implausible beliefs about rising asset prices during bubbles (Shiller 2005)

- Ideology: wide divergences in economic and political beliefs across otherwise similar countries (and also within)

- "Alternative facts," conspiracy theories, global warming, etc.

# Case-Shiller (2003): expectations of housing price increases

Question	Los Angeles 1988	Los Angeles 2003	San Francisco 1988	San Francisco 2003	Boston 1988	Boston 2003	Milwaukee 1988	Milwaukee 2003
Do you think that housing prices in the [city] area will increase or decrease over the next several years?								
Increase	98.3	89.7	99.0	90.5	90.2	83.1	87.1	95.2
Decrease	1.7	10.3	1.0	9.5	9.8	16.9	12.9	4.8
No. of responses	240	145	199	158	194	201	233	187
How much of a change do you expect there to be in the value of your home over the next 12 months?								
Mean response (percent)	15.3	10.5	13.5	9.8	7.4	7.2	6.1	8.9
Standard error	0.8	0.6	0.6	0.6	0.6	0.4	0.5	1.0
No. of responses	217	139	185	147	176	179	217	160
On average over the next 10 years, how much do you expect the value of your property to change each year?								
Mean response (percent)	14.3	13.1	14.8	15.7	8.7	14.6	7.3	11.7
Standard error	1.2	1.2	1.4	1.8	0.6	1.8	0.5	1.3
No. of responses	208	137	181	152	177	186	211	169

- Key question with most such measures: do they really believe (act on) it?
- Need better / more discriminating tests
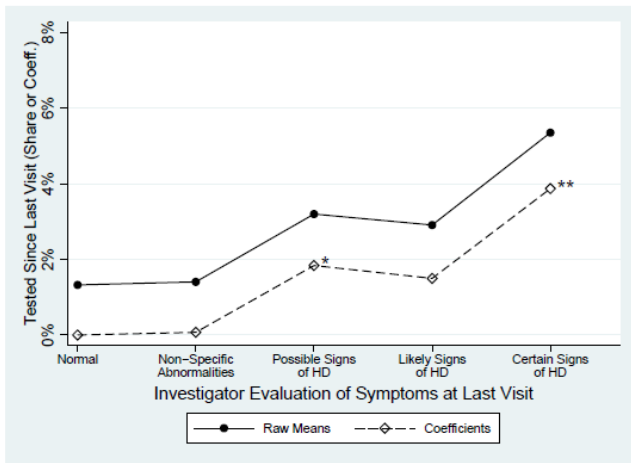
Wishful perceptions of health risks

- Oster et al. (2013): follow untested people at risk for risk for Huntington's disease (1 parent has gene variation $\Rightarrow$ 50% ex-ante chance; updated based on symptoms)

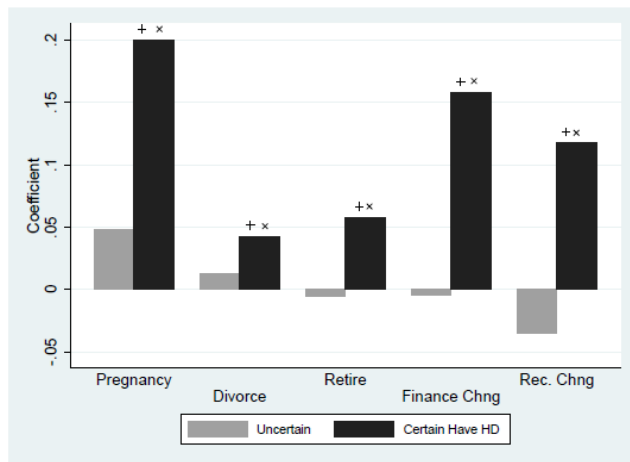Figure 4. : Perceived and Actual Risk of HD, by Motor Score
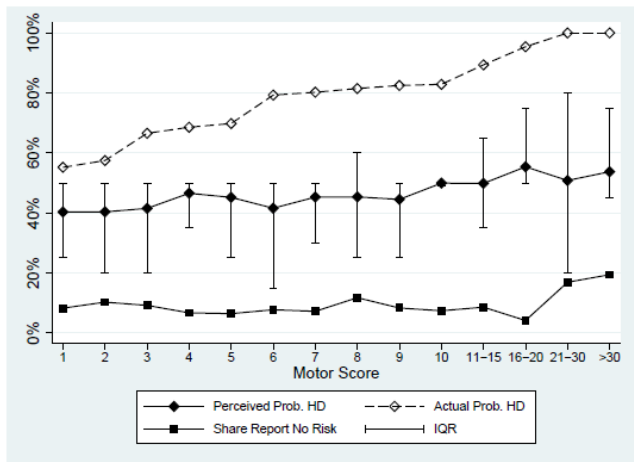
(Non) Demand for testing

Figure 1. : Testing Behavior and Investigator Evaluation of Risk

Behavior consistent with stated beliefs

Figure 5. : Behavior Choice Relative to Individuals without HD Expansion

What would you like to (not) know?

- Ganguly and Tasoff (2016)

Table 4: Information Preferences in the STD Experiment

	Proportion who forgo $10 to get information	Proportion who forgo $10 to avoid information	Proportion who forgo $10 to get preferred test		
			Seekers	Non-Seekers-Non-Avoiders	Avoiders
HSV-1	0.292	0.052	0.061	0.04	0.769
HSV-2	0.292	0.156	0.364	0.04	0.154
Difference	0.000	-0.104***	-0.303***	0.000	0.615***
	(0.066)	(0.043)	(0.093)	(0.039)	(0.154)
N	96	96	33	50	13

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Sample is those who took the STD Experiment first. Standard errors are in parentheses. Significance is determined using a two-sample one-sided proportions test. Seekers are defined as those who forwent $10 to get either HSV test, avoiders are defined as those who forwent $10 to avoid either HSV test, and "non-seeker-non-avoider" are those who are neither seekers nor avoiders. Five subjects exhibit both avoidance and seeking behavior and we place these in the non-seeker-non-avoider category.

II - Understanding Motivated Beliefs

1. Why? (Demand side)

- Standard decision theory, economics: better info $\Rightarrow$ always better off
- Hedonic value of beliefs: Schelling's (1984) "mind as a consuming organ"
  - ▸ Self-esteem, ego (B & T 2002, Koszegi 2006)
  - ▸ Anticipatory emotions/utility, reassurance about future (Ackerlof & Dickens 1982, Loewenstein 1989, Caplin & Leahy 2010, Brunnermeier & Parker 2005, B & T 2011)
- Functional value of beliefs
  - ▸ Self-motivation, self-control: worry about future selves' actions (B & T 2006)
  - ▸ Signaling: convincing oneself makes it easier to convince others

2. How? (Supply side)

- Ex-ante information acquisition or avoidance
- Ex-post signal distortion: "management" of attention, interpretation, recall
  - ▸ Either direct or via self inference (use own actions as diagnostics)

3. Welfare? Ultimately good/bad, functional or dysfunctional

- Alloy & Abrahamson (1979): depressed people are more realistic. Puri & Robinson (2007): optimists work more, save more, expect to retire later, more likely to remarry after divorce

Motivated cognition vs. fixed heuristics & biases

- Very different from mechanical biases and heuristics ("System I")
    - Kahneman ("Thinking, Fast and Slow" 2011); Rabin & Schrag (1999), Eyster & Rabin (2005), Madarasz (2012)
    - Here: critical role of emotions/desires, both current and anticipated, interacting with cognition
    - Responds to incentives and stakes, whether economic or psychological / hedonic. Example: self-esteem vs. confirmation bias
    - More cognitively sophisticated or educated people may be better at maintaining, defending desired beliefs (Kahan 2012: gun control, climate)

- Consistent with important new line in psych. that re-emphasizes role of emotions and their interlinkages with cognition
    - Damasio (1994): emotions, esp. in anticipating future situations, are critical to making even good decisions; sometimes, bad ones
    - Neuroscience; growing literature on processes underlying motivated beliefs, selective memory / asymmetric updating (Benoit & Anderson 2012, Sharot et al 2012)
    - "Cognitive Revolution" ⤳ "Affective Revolution, or "Second Cognitive Revolution"

III - A Simple Unifying Framework
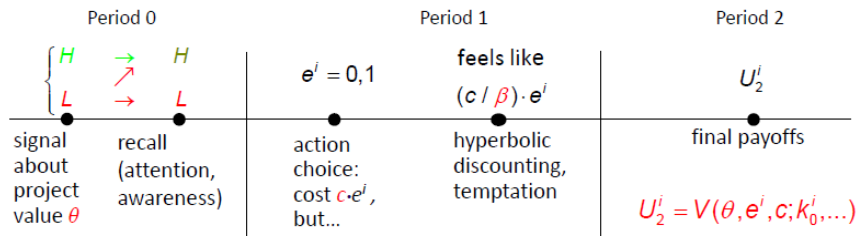
1. Self efficacy => motivated cognition (B&T 2006)

> *"Believe what is in the line of your needs, for only by such belief is the need fulfilled. . . Have faith that you can successfully make it, and your feet are nerved to its accomplishment."*
>
> *(William James, Principles of Psychology, 1890)*

- Individual fears that, later on, will face temptation to slack off, give up, cheat, overconsume
- Return (long-term value, effectiveness) of endeavor is imperfectly known: ability, talent, external events, market conditions, etc.
- Maintaining a "positive view" of that return has functional value: helps enhance, preserve motivation, resilience, "not giving up"
- Hence benefit to selectively process (encoding, recall, awareness) good vs. bad news. But also risk of costly mistakes.

III - A Simple Unifying Framework

1. Self efficacy => motivated cognition (B&T 2006)

| | Period 0 | Period 1 | Period 2 |

$$\begin{cases} H & \to & H \\ L & \to & L \end{cases}$$

$e^i = 0, 1$

feels like

$(c / \beta) \cdot e^i$

U_2^i

signal about project value θ

recall (attention, awareness)

action choice: cost $c \cdot e^i$, but...

hyperbolic discounting, temptation

final payoffs

$U_2^i = V(\theta, e^i, c; k_0^i, ...)$

- In period 1, will face temptation to slack off, give up, cheat, overconsume: β
- Return (long-term value, effectiveness) of endeavor is imperfectly known: θ depends on ability, talent, external events, market conditions, etc.
- Maintaining a "positive view" of that return has functional value: helps enhance, preserve motivation, resilience, "not giving up": $e^i \nearrow$
- Hence benefit to selectively process (encoding, recall, awareness) good vs. bad news (H vs. L). But also risk of costly mistakes.

A Simple Unifying Framework

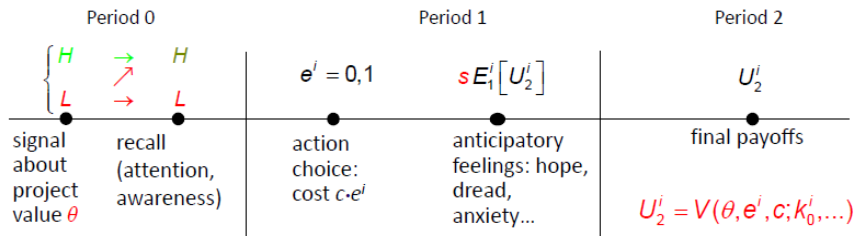2. Anticipatory feelings / self-esteem => motivated cognition (B&T 2011, B 2013)

> "I have done this, says my memory. I cannot have done that, says my pride, remaining inexorable. Finally –memory yields." (F. Nietzsche).

> "Pour briser les assauts renouvelés de ma mémoire, travaillait utilement en sens inverse mon imagination." (M. Proust)

- Individual experiences hope, dread, anxiety about long-term outcomes: wealth, health, social capital. "Consuming beliefs"
- Future prospects depend on decisions, returns (ability, talent, market), and initial endowment of (human, social, professional) capital at stake
- Maintaining a "positive view" of future outcomes has hedonic value
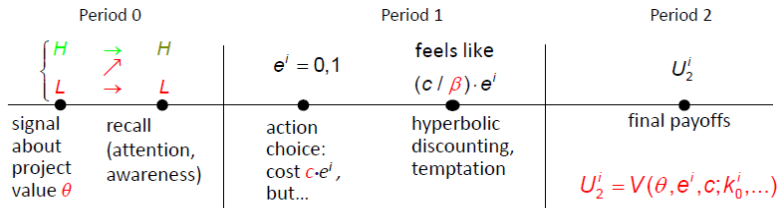- Hence benefit to selectively process (encoding, recall, awareness) good vs. bad news. But also distorts decisions

A Simple Unifying Framework

2. Anticipatory feelings / self-esteem => motivated cognition (B&T 2011, B 2013)

Period 0	Period 1	Period 2

$$\begin{cases} H & \rightarrow & H \\ L & \rightarrow & L \end{cases}$$

$e^i = 0,1$ — $sE_1^i\left[U_2^i\right]$ — U_2^i

signal about project value θ | recall (attention, awareness) | action choice: cost $c \cdot e^i$ | anticipatory feelings: hope, dread, anxiety... | final payoffs

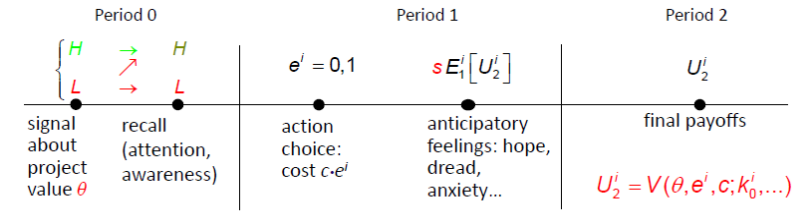$$U_2^i = V(\theta, e^i, c; k_0^i, \ldots)$$

- In period 1, will experience hope, dread, anxiety about long-term outcome, welfare "consuming" beliefs: s

- Future prospects depend on decisions, returns θ (ability, talent, market), and initial endowment of (human, social, professional) capital at stake k_0^i

- Maintaining a "positive view" of future outcomes has hedonic value

- Hence benefit to selectively process (encoding, recall, awareness) good vs. bad news (H vs. L). But also distorts decisions

1. Self-Motivation and Belief Distortion

	Period 0	Period 1	Period 2

$$\begin{cases} H \quad \rightarrow \quad H \\ L \quad \rightarrow \quad L \end{cases}$$

$e^i = 0, 1$

feels like
$(c / \beta) \cdot e^i$

U_2^i

signal about project value θ — recall (attention, awareness) — action choice: cost $c \cdot e^i$, but... — hyperbolic discounting, temptation — final payoffs

$$U_2^i = V(\theta, e^i, c; k_0^i, ...)$$

2. Anticipatory Utility and Belief Distortion

$$\begin{cases} H \quad \rightarrow \quad H \\ L \quad \rightarrow \quad L \end{cases}$$

$e^i = 0, 1$

$s E_1^i \left[U_2^i \right]$

U_2^i

signal about project value θ — recall (attention, awareness) — action choice: cost $c \cdot e^i$ — anticipatory feelings: hope, dread, anxiety... — final payoffs

$$U_2^i = V(\theta, e^i, c; k_0^i, ...)$$

- Integrates economic variables (e, c, θ), psychological variables both cognitive and affective (β, s, $E_1[\theta]$, $H \rightarrow L$,), social variables (k_i^0, others' actions e^j)
- All potentially measurable / experimentally manipulable.

A Simple Unifying Framework: Synthesis

- Period 1: makes decisions e^i to maximize

$$U_1^i = -c/\beta e^i + sE_1[U_2^i] + \delta E_1[U_2^i]$$

- Period 0: cognitive "choices" or tendencies, $L \rightsquigarrow H/L$, aiming to maximize

$$U_0^i = - \text{ info costs}/\beta + \delta E_0 \left[-ce^i + sE_1[U_2^i] \right] + \delta^2 E_0 \left[U_2^i \right]$$

  - Nests anticipatory utility $(\beta = 1, s > 0)$ & self-motivation $(\beta < 1, \ s = 0)$
  - Positive results similar, normative implications potentially different

- Useful to decompose final payoffs:

$$U_2^i = \alpha \cdot \theta_\sigma \cdot e^i + (1 - \alpha) \cdot \kappa_\sigma^i, \quad \text{for} \quad \sigma = H, L,$$

- κ_σ^i : fixed stakes, resulting from
  - Agent i's previous investments, sunk decisions: exogenous stakes
  - Other agents' $j \neq i$ equilibrium actions in state $\sigma = H, L$, affecting organization, market: endogenous stakes (Bénabou 2013)

Motivated cognition as biased information processing

- Signal $= H$ or $L \Rightarrow$ how much attention to pay, how to interpret, whether to "keep it in mind" or "not think about it"

- Wishful thinking: game with (future) self, via attention, memory, awareness, interpretation, rationalization (B & T 2002)

  - Realism: acknowledge - encode - recall $H \to H$ and $L \to L$

  - Denial: ignore - miscode - misremember $L \rightsquigarrow H$ (or $H \rightsquigarrow L$)

    Involves direct cost $+$ risk of mistakes / regret

  - Partial awareness: $0 <$ recall rate rate < 1

- Willful blindness: ex-ante information avoidance

  - At $t = 0$, agent chooses whether or not to learn the signal, take the test

  - Anticipatory utility concave in beliefs $\Rightarrow$ preferences for later / lesser resolution of uncertainty (Kreps-Porteus 1978, B. 2013)

  - Again, tradeoff with decision value of information.

Cognition and metacognition / sophistication

"But if I say I'm in denial then, by definition, I'm _not_ in denial."

Cognition and metacognition / sophistication

- Classical conundrum of self deception: simultaneously knowing *P* and not *P*?

- Resolved and modeled by taking into account actual, dynamic processes of attention, memory encoding, retrieval

- Makes clear difference between knowing *that* you forget selectively and knowing *what* you forgot

- Depending on long-run welfare consequences, individual may want to facilitate, or try to prevent, such selective forgetting or attention.

- "*I had during many years followed the Golden Rule, namely, that whenever a published fact, a new observation or thought came across me, which was opposed to my general results, to make a memorandum of it without fail and at once; for I had found by experience that such (contrary and thus unwelcome) facts and thoughts were far more apt to escape from memory than favorable ones*" (Charles Darwin).

Cognition and metacognition / sophistication

- Agents not free to "choose beliefs". Process information, optimally ($\neq$ objectively) at every stage

- At $t = 0$, aims to maximize

$$U_0^i = -m(1 - \lambda) + \delta E_0 \left[-ce^i + sE_1[U_2^i] \right] + \delta^2 E_0 \left[U_2^i \right]$$

- At $t = 1$,

  ▷ Being aware of / recalling signal L means state is L for sure

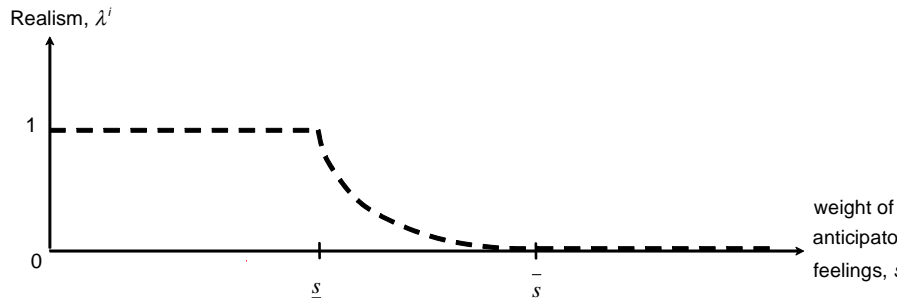 ▷ Being unaware of L / aware of H only leads to posterior

  $$\Pr \left[\text{ state was } H \mid \text{recall } H \right] = \frac{q}{q + \chi(1 - q)(1 - \lambda^i)} \equiv r(\lambda^i)$$

  λ^i : agent's equilibrium (habitual) rate of realism

  χ : degree of sophistication. Will use $\chi = 1$ benchmark here

Processing of bad news

- λ^i : probability with which recall / updating is accurate

Realism, λ^i

1

0

$\underline{s}$

$\bar{s}$

weight of
anticipato
feelings, s

- Individual trades off costs vs. benefits (hedonic, functional) of disregarding, rationalizing, not paying attention to bad news

- Self-efficacy version: replace salience s with weakness of will $1/\beta$

- Conversely: cost of mistakes, c acts like a decrease in s or an increase in β

IV – Main Results and Evidence

- Ex-post, asymmetric updating for good vs. bad news: denial, rationalization, wishful thinking. Evidence on recall, awareness, updating?

- Ex-ante, information avoidance: willful blindness. Evidence?

- Predictions: motivated thinking more likely for beliefs relevant to:

    ‣ Tasks for which perseverance in spite of temptation is more of an issue (β)

    ‣ Fixed or long-lasting forms of "capital": intelligence, health, attractiveness, honesty, social or cultural capital, ethnic identity, specialized human capital, illiquid assets (s). Equivalently, greater salience

    ‣ Issues on which final resolution ("day of reckoning") further into the future (δ)

    ‣ Higher initial stock of illiquid asset with uncertain return: $\kappa_\sigma^i \equiv \theta_\sigma \cdot k_0^i \Rightarrow$

    $\Rightarrow$ incentive for denial $\sim s(\theta_H - \theta_L) \times$ fixed stakes $= s(\theta_H - \theta_L)(1 - \alpha) \cdot k_0^i$

    Similar to "Stockholm syndrome"

- $\Rightarrow$ Stakes-dependent beliefs

Main results

- Decisions for which cost of mistakes (c) is smaller, e.g. because individual less likely to be pivotal: e.g. voting

With anticipatory utility / motives, further results:

- Endowment effect: have k_0^i (wealth, social or cultural capital, etc.) $\Rightarrow$ persuade myself will yield high return or future utility

- Escalating commitment: once think k^i asset is good for me, accumulate more of it, hence higher stakes in being optimistic about its long-term value to me, etc.

- Hedonic treadmill: such escalation may actually reduce utility, yet be unavoidable. Futile pursuit of wealth, fame, "purity", happiness…

- Self-traps: multiple intrapersonal equilibria. Same person can be in mode of
  - "Positive thinking": maintained optimism, "can-do" attitude, rosy glasses…
  - "Being honest with yourself": acknowledge failures and limitations; cautious, even "depressed"

Asymmetric updating about oneself

- "The Good News-Bad News Effect" (Eil & Rao AEJ 2011)

- Link to tradition in psychology: evidence of self-serving / selective / biased use or recall of information
  - Kunda (1987). Babcock-Loewenstein-Issacharoff (1995) on bargaining

- Stage 1: collect info to rank the subjects on intelligence (IQ tests) or beauty (speed dating). Control condition: card with random number from 0 to 9

- Stage 2:
  - Subjects state their prior belief, in %, for being in each of 10 ranks on task
  - Two rounds of: (a) learn if rank above of below other randomly selected, anonymous participant; (b) state updated belief, incentivized for accuracy
  - At the end: elicit willingness to pay to learn / not learn true rank

Summary of main findings

1. Update close to Bayes' rule for positive signals, underupdate for negative signals. But only when signals are about something have a lasting stake in.

2. Buy information when have relatively optimistic beliefs about, will pay to avoid it when have pessimistic beliefs

3. No evidence of "mechanical" confirmatory bias: "desirability" (valence) of signal is key

4. Möbius et al. (2010) "Self-Confidence Management: Theory & Experimental Evidence":

   ▶ Similar experiment, on IQ only. Beliefs elicitation mechanism more robust + subjects state beliefs only about being in top 50% rather than full posterior distribution: easier to compute what Bayesian updating should be.

   ▶ Find underadjustment even to good signals, but significantly more so in response to negative signals / bad news

Sharot-Korn-Dolan : "How Unrealistic Optimism is Maintained in the Face of Reality" (Nature Neuroscience 2012)

- For 80 "bad life events" (e.g., cancer, accident, etc.): self-ratings of own risk, both before and after receiving accurate information about true probability for a person of same age, gender, ethnic and socioeconomic characteristics

- Examine whether updating displays good/bad news asymmetry
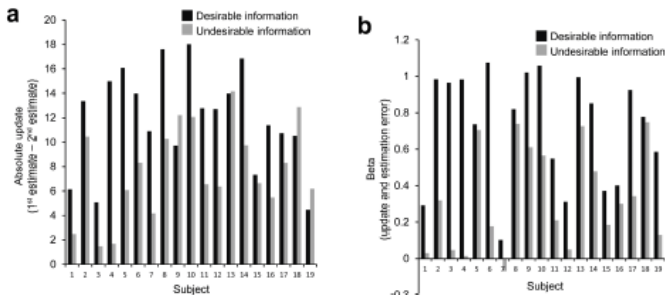
Mechanism

- Examine whether prediction error has explanatory power for extent of belief revision: it does
- See what regions of brain activated by + or - prediction error: different ones
- Across subjects: high optimists (based on prior questionnaires) show systematically less activation of area detecting negative prediction errors; no difference for positive ones
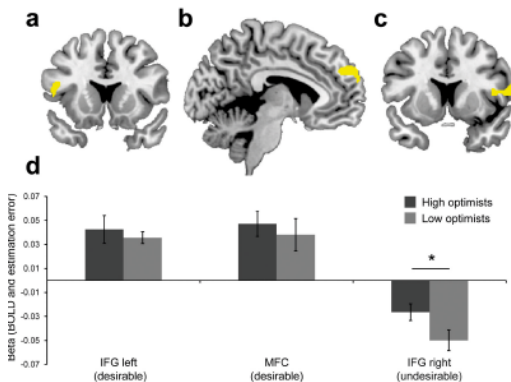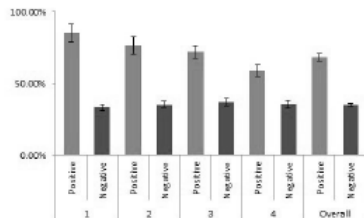
Asymmetric recall of past performance

- "Selective Memory & Motivated Delusion" (Chew, Huang & Zhao 2012)

- Stage 1: 621 subjects, each answers 4 questions from Ravens IQ test; incentive = lottery for \$100, worth $\approx$ \$1 in expectation

- Stage 2: Two months later, called back, showed same 4 questions + 2 had not seen, with the answers
  - Asked to recall whether answered correctly, incorrectly, had not seen, or can't remember. +\$1 for correct response, -\$1 for incorrect, 0 for "can't remember"

- 8 possible types of recall errors: +/- "Amnesia" ($\sigma \to \varnothing$), "Confabulation," ($\sigma \to \sigma'$), "Delusion" ($\varnothing \to \sigma$)

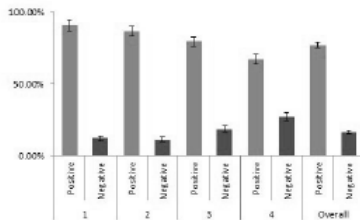|	a	b	c	d
$s = G$	a_G : CR	b_G : Negative C	c_G : Negative A	d_G : Weak Negative A
$s = B$	a_B : Positive C	b_B : CR	c_B : Positive A	d_B : Weak Positive A
$s = \varnothing$	$a_\varnothing$: Positive D	$b_\varnothing$: Negative D	$c_\varnothing$: CR	$d_\varnothing$: Weak CR

Memory biases conditional on performance

Positive Amnesia vs. Negative Amnesia

Positive Delusion vs. Negative Delusion
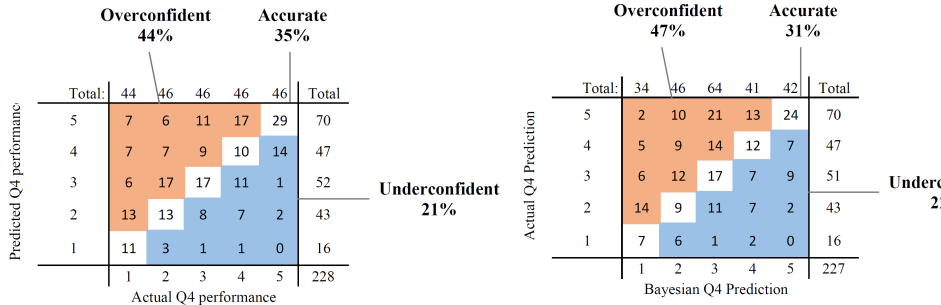
Positive Confabulation vs. Negative Confabulation
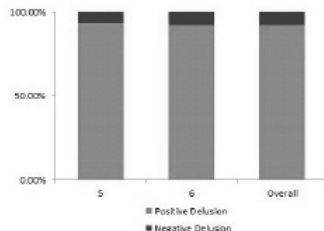
From the lab to the real world

- Huffman, Raymond and Shvets (2017) Persistent Overconfidence and Biased Memory: Evidence from Managers

- 239 managers, each running separate food/beverage store. Each quarter, firm runs tournament incentive scheme, regional/national, determining bonuses

  - High stakes: range up to 150% of base salary, 50% of monthly income
  - Repeated feedback, well-defined performance criteria, familiar activity known group of competitors
  - Researchers observe 31 quarters of performance & feedback

- Measure managers'

  - Predictions (confidential vis a vis firm) of own next-quarter performance (most likely quintile) in this real workplace tournament - Incentivized
  - Recall: previous quarter's performance (rank +/ 10) - Incentivized
  - Types: traits, preferences, and ability of managers measured using surveys, experiments, and historical data

Overconfidence about future (2015-Q4) performance

Left table (Predicted Q4 performance × Actual Q4 performance):

Overconfident 44% Accurate 35%

Total:	44	46	46	46	46	Total
5	7	6	11	17	29	70
4	7	7	9	10	14	47
3	6	17	17	11	1	52
2	13	13	8	7	2	43
1	11	3	1	1	0	16
	1	2	3	4	5	228

Predicted Q4 performance / Actual Q4 performance

Underconfident 21%

Right table (Actual Q4 Prediction × Bayesian Q4 Prediction):

Overconfident 47% Accurate 31%

Total:	34	46	64	41	42	Total
5	2	10	21	13	24	70
4	5	9	14	12	7	47
3	6	12	17	7	9	51
2	14	9	11	7	2	43
1	7	6	1	2	0	16
	1	2	3	4	5	227

Actual Q4 Prediction / Bayesian Q4 Prediction

Underc... 2...

- Less than 0.1% chance that Bayesian model would generate observed predictions

Biased recall of past (2015-Q2) performance
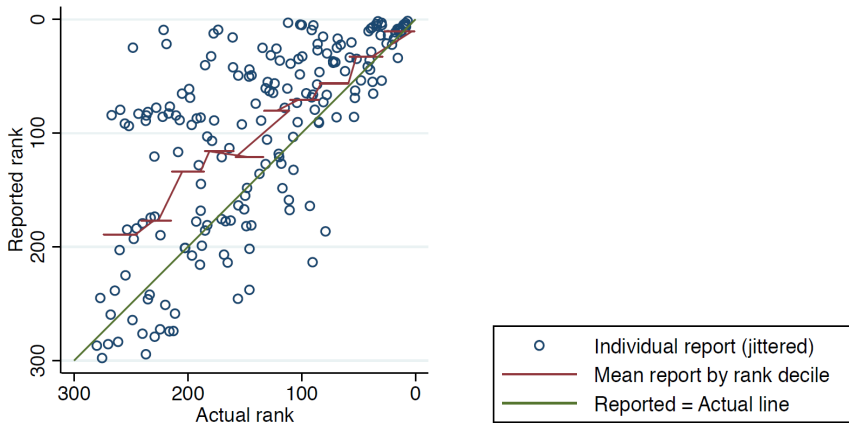
- Recall correlated with actual, 0.67, but significantly different ($p < 0.1\%$). Forgetting more prevalent for worse ranks
- 68% of managers recall better-than-actual rank. Median deviation is 9%

Main Findings

1. Overconfidence in own performance predictions, even for managers with substantial experience, and too extreme to be rational

2. Inaccuracy in recalling recent past performance

   (incentives + they know that experimenter knows it!)

   - Forgetting is more likely if recent performance was bad (controlling for historical ability)

   - Errors in recall are skewed heavily towards better than actual performance

3. Predictions about the future are related to recalled past performance (conditional on actual performance) $\implies$ see "reality denial" in action

4. In progress: How does overconfidence about 2015-Q4 relate to performance in early quarters of 2016?

   - No significant relationship to overall rank, but signs of $\neq$ "managerial styles":

   - Overconfident managers do better in terms of quarterly profit, but have lower ratings for customer service

Stakes-dependent beliefs: hedonic

- Mijovic-Prelec and Prelec (2010): "Self-Deception As Self-Signalling: A Model And Experimental Evidence"

- Mayraz (2011): "Wishful Thinking"

  145 subjects, observe chart of historical wheat prices, then must predict what price would be at date 100. Also state a level of confidence (1-10) in their prediction

  - Paid accuracy bonus. Do this 12 times
  - All prices normalized to lie in $[\$4000, \$12000]$

- Randomly assigned to being Farmers, whose payoff is $P - 4000$, or Bakers, whose payoff is $16000 - P$

- Stakes $= \$0.5$ or $\$1$ for each $\$1,000$ of notional profit

Stakes-dependent beliefs
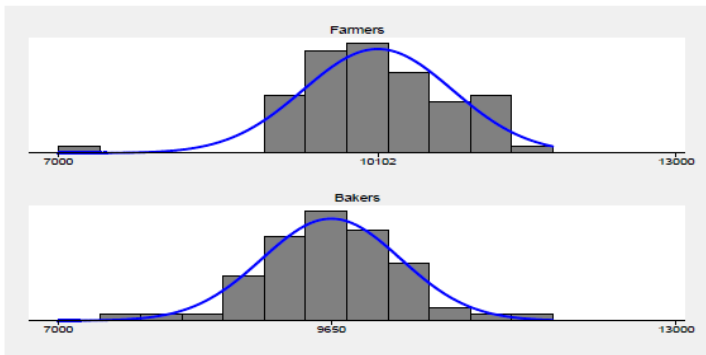
Figure 4: Histogram of the mean predictions made by *Farmers* and *Bakers*. A normal distribution curve was fitted to both histograms. The mean prediction was 10102 and 9650 respectively. 16 of the 20 subjects making the highest (lowest) mean predictions were *Farmers* (*Bakers*).

- Not consistent with rational expectations, fixed cognitive bias, or ego utility
- Consistent with anticipatory utility, broadly defined

Stakes-dependent beliefs: instrumental

- Schwardmann & van der Weele (JEEA 2017): "Deception & Self-Deception"

    ▸ Test channel of "deceiving oneself to deceive others" (Trivers & von Hippel)

- Design: 288 subjects = 18 sessions of 16. Each session = 4 × 4 groups. First, take IQ (Raven's test). Then:

- "Self -deception" stage:

    ▸ Elicit incentivized beliefs that are among top 2 in group, under "control' and "contestant" conditions. Contestants told that will be matched with "employer," who will decide whom to "hire," with incentives for picking top performers. Being hired is valuable

    ▸ Give noisy feedback about performance, elicit posterior beliefs. Also a clever source of exogenous variation in self-confidence.

- "Deception stage": face to face "interview" with employer

    ▸ Additional conditions (i) give employers lie-detection tutorial, warn/not warn contestants about it; (ii) measure lying aversion
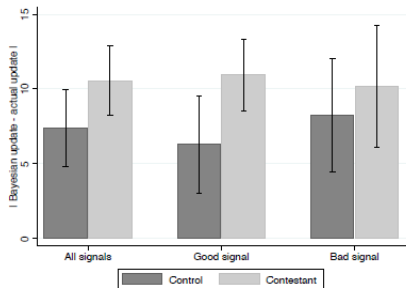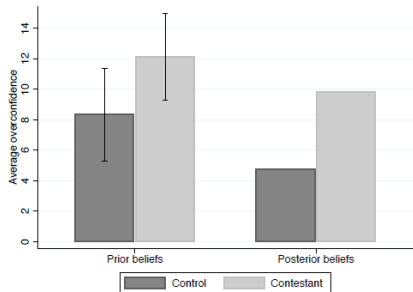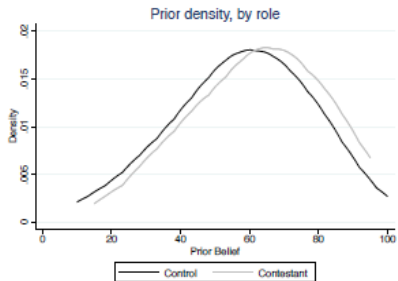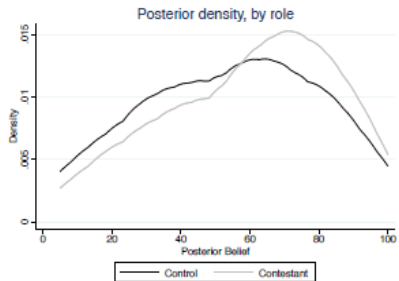
(a)   (b)

Results

- Prior confidence about 50% higher among contestants

- Unbiased feedback lowers it in both groups, but much less so among contestants: posterior confidence more than twice as high as for controls

- Higher self-confidence (exogenous = noise in feedback signal) leads to large increase in employer evaluations / hiring when employers not given lie detection tutorial (coefficient = .6)

- When employers are trained, effect disappears, ability (true performance) is what matters most matters for evaluations. When contestants are warned about training, neither ability nor confidence affects evaluations.

- Related study: Charness-Rusctichini-van de Ven: "Self Confidence and Strategic Behavior"

Self-Serving Beliefs in the Moral Domain

- Babcock et al. (AER 1995) "Biased Judgements of Fairness in Bargaining," Thompson & Loewenstein (1992) "Egocentric Interpretations of Fairness in Negotiations".   ▷ Summary

- Di Tella, Perez-Truglia, Babino, and Sigman (AER 2015): Conveniently Upset: Avoiding Altruism by Distorting Beliefs About Others' Altruism"

    ▶ Variation on standard "Trust Game" between Investor and Trustee. "Allocator" and "Seller" each start with 10 tokens. Allocator decides how to redistribute the combined 20 tokens between herself and Seller: below

    ▶ Simultaneously, Seller chooses "price" at which all the tokens are sold to experimenter: $1.5 or $0.5. In latter case, gets "bribe" of $5

    ▶ Neither side knows anything about the other's choice. Elicit incentivized Allocators' beliefs ($5) about what her Seller chose, and % of corrupt Sellers

    ▶ Allocator can appropriate either ≤ 2 or ≤ 8 of Seller's 10 tokens.

    ▶ Does this exogenous opportunity affect her beliefs about Sellers' degree of corruption, and/or what allocation she chooses?

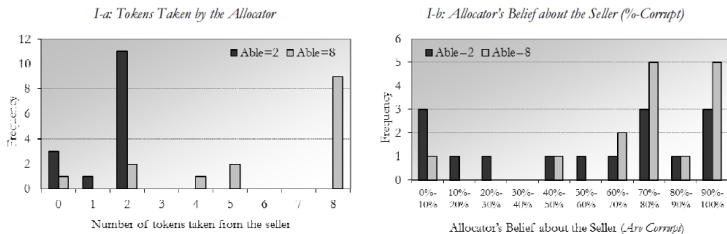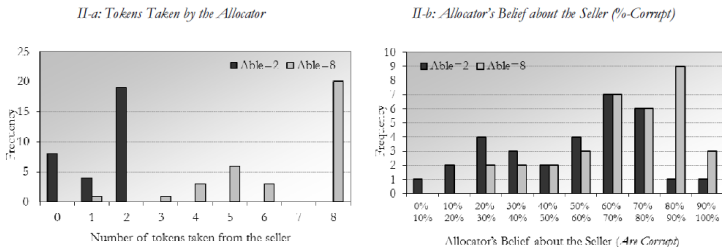Figure I: Distribution of Actions and Beliefs by Treatment Group, Basic Game

I-a: Tokens Taken by the Allocator

I-b: Allocator's Belief about the Seller (%-Corrupt)

Figure II: Distribution of Actions and Beliefs by Treatment Group, Modified Game

II-a: Tokens Taken by the Allocator

II-b: Allocator's Belief about the Seller (%-Corrupt)

- "Forced Seller" or "Forced Allocator" treatments: the effect disappears.

Back to the field

- Di Tella et al. (AER 2004) "Property Rights and Beliefs: Evidence from the Allocation of Land Titles to Squatters"

  - ▸ Exploit rare "natural experiment": after years of disputes stuck in a very inefficient and arbitrary legal process, Argentinian government granted land titles to some very poor households and others not, in a quasi-random manner

  - ▸ Compared to their less lucky neighbors, beneficiaries of this pure windfall responded by adopting increased "pro-market" beliefs that one can succeed on one's own, money is important for happiness, and others can be trusted

- Karlsson et al. (JRU 2009): " "The Ostrich Effect: Selective Avoidance of Information"

  - ▸ Have access to (anonimized) online activity of millions of retirement investors via Swedish Premium Pension Authority and, in US, Vanguard Group. Can see when people log in, rebalance portfolios, etc.

  - ▸ Main finding: investors go online to look up the value of their portfolios much more on days when the market as a whole is up (on average, "good news").

  - ▸ Relates to well-know "disposition effect": investors (stocks, housing), have clearly asymmetric propensity to "sell winners and hold on to losers"

V -Concluding Thoughts

- Beliefs, and how they respond to information, reflect a variable mix of:

  - ▶ Standard Bayesian rationality

    With often cognitive limits: biases and heuristics

  - ▶ Standard psychological mechanisms:

    - Demand side: arise in response to functional or affective needs

    - Supply side: reflect workings of attention, interpretation, memory processes

- The factors can, ultimately, be modeled and measured, experimentally varied

- Evidence shows that beliefs are neither fully responsive to data, nor rigidly fixed by either cognitive heuristics or social determinants

- In line with theory, they are sensitive to tradeoffs between the costs and benefits of accuracy vs. "desirability"

- Broad combination of methods: surveys, experiments in lab and field, neuro, formal modeling, empirical studies on large datasets...

V -Concluding Thoughts

- Motivated cognition has important consequences for actual behaviors: health, business decisions, investments, politics

- What is still missing? Surely, a lot

  - More evidence on motivated cognition, in the lab + especially in the field

  - Other motives / mechanisms not yet captured?.

  - Excessive pessimism, panics, irrational fears

- Collective beliefs, "social cognition": next.