Beliefs and Misbeliefs

Roland Bénabou

Princeton University

1st Tony Atkinson Lecture

LSE - STICERD – October 2014

Based in part on joint work with Jean Tirole (TSE),
and with Davide Ticchi (Lucca) & Andrea Vindigni (Lucca)
How do people form their beliefs?

1. Backward-looking expectations, adaptive learning

2. Rational expectations, Bayesian equilibrium (with refinements)

3. Fixed (wired-in) “biases and heuristics”: base rate neglect, confirmation bias, law of small numbers, hot hand fallacy, probability weighting...

4. Motivated beliefs, cognition, reasoning:
   - Held (or more likely to be) due to emotional or functional value
   - Resistant to evidence, but respond to costs, benefits and stakes
   - Several other distinctive “signatures”
Motivated beliefs / cognition

- **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

- Such beliefs are central to individual and collective performance, social relationships, well-being

- 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

  ⇒ Beliefs as assets that people invest in, value, defend, expend, repair, etc.
Beliefs and misbeliefs: some examples

- 90% of US drivers think they are better than average (Svensson 1981). 94% of professors at large US university thought were better than average professor, etc.

- People think less likely than similar others to suffer adverse life events, more likely to experience good ones (Weinstein 1981)

- 47% of Americans think humans were created instantaneously, 52% believe that humans and dinosaurs coexisted (2001 NSF survey).

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

- Wide divergences in economic and political beliefs across otherwise similar countries (and also within): ideologies
Case-Shiller (2003): expectations of housing price increases

<table>
<thead>
<tr>
<th>Question</th>
<th>Los Angeles</th>
<th>San Francisco</th>
<th>Boston</th>
<th>Milwaukee</th>
</tr>
</thead>
<tbody>
<tr>
<td>Do you think that housing prices in the [city] area will increase or decrease over the next several years?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Increase</td>
<td>98.3</td>
<td>89.7</td>
<td>99.0</td>
<td>90.5</td>
</tr>
<tr>
<td>Decrease</td>
<td>1.7</td>
<td>10.3</td>
<td>1.0</td>
<td>9.5</td>
</tr>
<tr>
<td>No. of responses</td>
<td>240</td>
<td>145</td>
<td>199</td>
<td>158</td>
</tr>
<tr>
<td>How much of a change do you expect there to be in the value of your home over the next 12 months?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean response (percent)</td>
<td>15.3</td>
<td>10.5</td>
<td>13.5</td>
<td>9.8</td>
</tr>
<tr>
<td>Standard error</td>
<td>0.8</td>
<td>0.6</td>
<td>0.6</td>
<td>0.6</td>
</tr>
<tr>
<td>No. of responses</td>
<td>217</td>
<td>139</td>
<td>185</td>
<td>147</td>
</tr>
<tr>
<td>On average over the next 10 years, how much do you expect the value of your property to change each year?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean response (percent)</td>
<td>14.3</td>
<td>13.1</td>
<td>14.8</td>
<td>15.7</td>
</tr>
<tr>
<td>Standard error</td>
<td>1.2</td>
<td>1.2</td>
<td>1.4</td>
<td>1.8</td>
</tr>
<tr>
<td>No. of responses</td>
<td>208</td>
<td>137</td>
<td>181</td>
<td>152</td>
</tr>
</tbody>
</table>
Beliefs about social mobility

Source: Alesina et al. (2005)
Belief in markets (WPO 2005)
Belief in markets and size of the state

\[ \beta = -0.22 \]
\[ t\text{ stat} = -2.01 \]

Source: Bénabou (2008)
Do they really believe (act on) it?

- Do so in incentivized experiments, e.g. displaying overconfidence

- Empirical data $\Rightarrow$ evidence that do for health, housing, stocks

- Often incur high costs to defend or “express” beliefs: identity, religion
  - Augenblick et al. (2012) on end-of-world beliefs

- Vote on it:
  - Beliefs about determinants of economic success (luck or effort) are strong explanatory factors of individual attitudes toward redistribution as well as actual national social spending (Alesina et al. 2001)
  - Trust in markets strong (negative) predictor of size of the state/GDP (Bénabou 2008)
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)
(Non) Demand for testing

Figure 1. Testing Behavior and Investigator Evaluation of Risk

- Tested Since Last Visit (Share or Coeff.)
- Investigator Evaluation of Symptoms at Last Visit
- Raw Means
- Coefficients
Behavior consistent with stated beliefs

Figure 5. Behavior Choice Relative to Individuals without HD Expansion
Lecture Plan

1. Motivated beliefs, cognition: why and how?
2. A simple unifying framework
3. Implications and evidence: individual beliefs and behavior
4. Implications and evidence: group beliefs and behavior (organizations, markets, ideology)
5. Religion
6. Concluding remarks
I - Understanding Motivated Cognition

1. Why? (Demand side)

- Standard decision theory: better info $\Rightarrow$ single DM (weakly) better off

- Hedonic value of beliefs: Schelling’s (1984) “mind as a consuming organ”
  - Self-esteem, ego (B & T 2002, Koszegi 2006)

- Functional value of beliefs
  - Self-motivation, self-control: worry about future selves’ actions
  - 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

Fixed heuristics & biases vs. motivated cognition

- Very different from mechanical biases and heuristics ("System I")
  - Critical role of emotions/desires, both current and anticipated, interacting with cognition
  - Responds to incentives and stakes, whether economic or psychological / hedonic. Example: self-serving beliefs vs. confirmation bias
  - More cognitively sophisticated or educated people may be better at maintaining, defending desired beliefs (Kahan 2012)

- Consistent with recent trend in psych. that re-emphasizes role of emotions, especially those evoked by future good and bad prospects
  - 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)
II - Motivated Beliefs: a Simple Unifying Framework

1. Self-Motivation and Belief Distortion

Period 0

\[ \{H \quad \rightarrow \quad H \} \]

Period 1

\[ e^i = 0,1 \]

Period 2

\[ U_2^i \]

signal about project value \( \theta \)

recall (attention, awareness)

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

feels like

\[ \frac{c}{\beta} \cdot e^i \]

hyperbolic discounting, temptation

final payoffs

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

2. Anticipatory Utility and Belief Distortion

Period 0

\[ \{H \quad \rightarrow \quad H \} \]

Period 1

\[ e^i = 0,1 \]

Period 2

\[ U_2^i \]

signal about project value \( \theta \)

recall (attention, awareness)

action choice: cost \( c \cdot e^i \)

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

anticipatory feelings: hope, dread, anxiety...

final payoffs

\[ U_2^i = V(\theta, e^i, c; k_0, \ldots) \]
Preferences and payoffs

- Period 1: makes decisions (if any) to maximize

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

- Period 0: makes cognitive “choices”, aiming to maximize

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

  - Nests anticipatory utility ($\beta = 1, s > 0$) & self-motivation ($\beta < 1, s = 0$)

- Positive results similar. Normative implications potentially different.

- Final payoffs: with $\sigma = H, L$,

\[ U_2^i = \alpha \cdot \theta_\sigma e^i + (1 - \alpha) \cdot \kappa_\sigma^i \]

- $\kappa_\sigma^i$: 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
Information processing

- Signal $\sigma = H$ or $L \Rightarrow$ how much attention to pay, how to interpret, whether to “keep it in mind” or “not think about it”. Also: willingness to pay for $\sigma$

- Wishful thinking: intrapersonal game of communication, via attention, memory, awareness, interpretation, rationalization (Bénabou & Tirole 2002)
  - Realism: acknowledge - encode - recall $H \rightarrow H$ and $L \rightarrow L$
  - Denial: ignore - miscode - misremember $L \rightsquigarrow H$ (or $H \rightsquigarrow L$)
    - Self-deception, selective inattention, rationalization: cost $m \geq 0$
  - Partial awareness: recall rate $0 < \lambda < 1$, when indifferent

- Not wanting to know: : ex-ante information avoidance
  - At $t = 0$, agent chooses whether or not to learn the signal $\sigma$
  - No anticipatory utility nor malleable awareness, but preferences for late resolution of uncertainty (Kreps-Porteus 1978, Bénabou 2013)
  - Tradeoff with decision value of information.
Dealing with unpleasant realities

In state $\sigma = L$, net Incentive to deny, rationalize away red flag is

$$
\Delta U_0^i \equiv U_{0,\text{Denial}}^i - U_{0,\text{Realism}}^i = -\frac{m}{\beta} - \delta \left[ \frac{c}{\gamma} - (\delta + s) \alpha \theta_L \right]
$$

decision impact

$$
+ \delta s \ r(\lambda^i) \ \delta \left[ \alpha (\theta_H - \theta_L) + (1 - \alpha) (\kappa_H - \kappa_L) \right],
$$
gain in anticipatory utility

$$
r(\lambda^i) = \frac{q}{q + \chi (1 - q) (1 - \lambda^i)}
$$

$\lambda^i$: agent $i$’s equilibrium realism (recall rate for $L$ signals)

$\chi \in [0, 1]$: degree of Bayesian sophistication. Benchmark: $\chi = 1$
III – Main Results: Individual Behavior

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

- Ex-ante, information avoidance: willful blindness

- Comparative statics: selective awareness more likely for beliefs relevant to:
  - Tasks for which perseverance in spite of temptation is more of an issue, i.e. $c < (\delta + s) \alpha \theta_L < c/\beta c$
  - Fixed or long-lasting forms of "capital": intelligence, health, attractiveness, honesty, social or cultural capital, ethnic identity, specialized human capital, illiquid assets: higher $s$
  - Issues on which final resolution further into the future
  - Higher initial endowments of durable, illiquid assets –or more salient:

  $$\kappa^i_0 \equiv \theta k^i_0 \Rightarrow \frac{\partial \Delta U^i_0}{\partial k^i_0} = s(1 - \alpha)(\theta_H - \theta_L)$$

  $\Rightarrow$ Stakes-dependent beliefs
Main results: individual behavior

- Decisions for which cost of mistakes is smaller, e.g. because individual less likely to be pivotal: e.g. voting (Caplan 2007, B & T 2006, B. 2008)

- **Endowment effect:** have $k_0^i \Rightarrow$ 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 (self-trap. Pursuit of wealth, fame, “purity”).

- **Normative** consequences of belief distortion depend importantly on whether they arise from “mental consumption” (also: concave / convex) or from functional motives. Positive predictions less so.
Asymmetric updating about oneself

- “The Good News-Bad News Effect” (Eil & Rao 2011); Möbius et al. (2010)
- Link to tradition in psychology: evidence of self-serving / selective / biased use or recall of information

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)
  - At the end: elicit willingness to pay to learn / not learn true rank
### Table 2—Mean Belief Changes by Signal Direction and Condition, Dependent Variable: $\Delta\mu$

<table>
<thead>
<tr>
<th>Condition</th>
<th>Beauty</th>
<th>IQ</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta\mu_{Bayex}$</td>
<td>0.212</td>
<td>0.0256</td>
<td>0.141</td>
</tr>
<tr>
<td></td>
<td>(0.160)</td>
<td>(0.072)</td>
<td>(0.136)</td>
</tr>
<tr>
<td>$\Delta\mu_{Bayex} \times 1{\sigma = 1}$</td>
<td>0.475*</td>
<td>0.540***</td>
<td>0.141</td>
</tr>
<tr>
<td></td>
<td>(0.263)</td>
<td>(0.131)</td>
<td>(0.172)</td>
</tr>
<tr>
<td>$1{\sigma = 1}$</td>
<td>-0.642***</td>
<td>-0.772***</td>
<td>-1.119***</td>
</tr>
<tr>
<td></td>
<td>(0.150)</td>
<td>(0.101)</td>
<td>(0.172)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.437***</td>
<td>0.599***</td>
<td>0.625***</td>
</tr>
<tr>
<td></td>
<td>(0.134)</td>
<td>(0.061)</td>
<td>(0.168)</td>
</tr>
<tr>
<td>Observations</td>
<td>206</td>
<td>183</td>
<td>385</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.49</td>
<td>0.55</td>
<td>0.48</td>
</tr>
</tbody>
</table>

Bootstrapped standard errors in parentheses

* * * p < 0.01, ** p < 0.05, * p < 0.1
**Table 1—Subject’s Mean Belief as a Function of Bayesian Mean**

<table>
<thead>
<tr>
<th></th>
<th>Beauty</th>
<th>IQ</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bayesian $\mu$</td>
<td>0.641***</td>
<td>0.846***</td>
<td>0.589***</td>
</tr>
<tr>
<td></td>
<td>(0.081)</td>
<td>(0.102)</td>
<td>(0.099)</td>
</tr>
<tr>
<td>Bayesian $\mu \cdot 1{\text{All } \sigma = 1}$</td>
<td>0.406***</td>
<td>0.0714</td>
<td>-0.00118</td>
</tr>
<tr>
<td></td>
<td>(0.096)</td>
<td>(0.129)</td>
<td>(0.192)</td>
</tr>
<tr>
<td>$1{\text{All } \sigma = 1}$</td>
<td>-2.311***</td>
<td>-0.580</td>
<td>-0.531</td>
</tr>
<tr>
<td></td>
<td>(0.545)</td>
<td>(0.811)</td>
<td>(0.946)</td>
</tr>
<tr>
<td>Constant</td>
<td>2.054***</td>
<td>1.051</td>
<td>2.291***</td>
</tr>
<tr>
<td></td>
<td>(0.525)</td>
<td>(0.757)</td>
<td>(0.760)</td>
</tr>
<tr>
<td>Observations</td>
<td>163</td>
<td>139</td>
<td>292</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.867</td>
<td>0.869</td>
<td>0.717</td>
</tr>
</tbody>
</table>

Standard errors clustered at subject level in parentheses

*** p<0.01, ** p<0.05, * p<0.1
Actual vs. Bayesian updating

- Now includes all observations

*Figure 3. Linear fitted values of round-to-round changes in mean belief by condition.*
<table>
<thead>
<tr>
<th>Condition</th>
<th>Beauty</th>
<th>IQ</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta \mu_{Bayes}$</td>
<td>0.212</td>
<td>0.0256</td>
<td>0.141</td>
</tr>
<tr>
<td>(0.160)</td>
<td>(0.072)</td>
<td>(0.136)</td>
<td></td>
</tr>
<tr>
<td>$\Delta \mu_{Bayes} \times 1{\sigma = 1}$</td>
<td>0.475*</td>
<td>0.540***</td>
<td>0.141</td>
</tr>
<tr>
<td>(0.263)</td>
<td>(0.131)</td>
<td>(0.172)</td>
<td></td>
</tr>
<tr>
<td>1{\sigma = 1}</td>
<td>-0.642***</td>
<td>-0.772***</td>
<td>-1.119***</td>
</tr>
<tr>
<td>(0.150)</td>
<td>(0.101)</td>
<td>(0.172)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.437***</td>
<td>0.599***</td>
<td>0.625***</td>
</tr>
<tr>
<td>(0.134)</td>
<td>(0.061)</td>
<td>(0.168)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>206</td>
<td>183</td>
<td>385</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.49</td>
<td>0.55</td>
<td>0.48</td>
</tr>
</tbody>
</table>

Bootstrapped standard errors in parentheses

*** $p<0.01$, ** $p<0.05$, * $p<0.1$
Response to (self-relevant) bad news is also more noisy

Figure 4. Noise in round-to-round updating by treatment and signal type. Bars give +/- 1.96 standard errors.
### Table 4 — $\Delta \mu - \Delta \mu_{H_{avg}}$ as a Function of Prior Group and Signal Valence

| CB in words | CB prediction | Good news | | Bad news | |
|-------------|---------------|-----------|-----------|-----------|
|             | Above avg | Below avg | Above avg | Below avg |
| Beauty      |            |           |           |           |
| CB          | —         | +         | Under respond | Over respond |
|             | -0.14     | -0.13     | 0.00       | -0.13     |
|             | (-3.74)   | (-1.20)   | (-0.03)    | (-1.32)   |
| IQ          |            |           |           |           |
|             | —         | +         | Under respond | Over respond |
|             | 0.07      | -0.08     | -0.14      | 0.02      |
|             | (1.46)    | (-0.70)   | (-0.91)    | (0.23)    |
| Control     |            |           |           |           |
|             | —         | +         | Under respond | Over respond |
|             | 0.13      | 0.21      | -0.36      | -0.37     |
|             | (1.69)    | (2.11)    | (-5.16)    | (-3.89)   |

$t$ statistics ($H_0 = 0$) in parentheses
**Informational preferences**

**Table 3—Willingness-to-pay as a function of final round beliefs**

<table>
<thead>
<tr>
<th>Condition</th>
<th>IQ</th>
<th>Beauty</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Final Round $\mu$</td>
<td>-0.325***</td>
<td>-0.206**</td>
<td>0.0112</td>
</tr>
<tr>
<td></td>
<td>(0.107)</td>
<td>(0.0855)</td>
<td>(0.0423)</td>
</tr>
<tr>
<td>Final Round $\sigma$</td>
<td>0.911**</td>
<td>0.917</td>
<td>0.237</td>
</tr>
<tr>
<td></td>
<td>(0.382)</td>
<td>(0.575)</td>
<td>(0.174)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.729</td>
<td>0.298</td>
<td>-0.339</td>
</tr>
<tr>
<td></td>
<td>(0.523)</td>
<td>(0.750)</td>
<td>(0.336)</td>
</tr>
<tr>
<td>Observations</td>
<td>77</td>
<td>65</td>
<td>142</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.099</td>
<td>0.180</td>
<td>0.015</td>
</tr>
</tbody>
</table>

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Robust standard errors in parentheses

**Graph:** Mean posterior belief in final round of signals
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 stake in.

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

3. No evidence of confirmatory bias, **valence of signal matters**!

   - Similar experiment (on IQ only) with even "cleaner" methodology: beliefs elicitation mechanism more robust + subjects state beliefs only about binary outcome (being in top 50%) rather than full posterior distribution, making it much easier to compute what Bayesian updating should be.
   - Find underadjsutment even to good signals, but significantly more in response to negative signals.
Asymmetric updating about educational returns

- “How do Students Respond to Information about Earnings?”
  (Wiswall & Zafar, 2013)

- **Three steps:** (a) Elicit beliefs about own future earnings & average earnings by major; (b) Provide actual population earnings, by major; (c) Elicit updated beliefs about own earnings

### Table 6: Self Earnings Updating and Population Errors

<table>
<thead>
<tr>
<th>Dependent Variable: Revisions in Self Earnings Beliefs (Intermediate – Initial)</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A</td>
<td>Error&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.184***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Error x General T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Error x Specific T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Error x 1(Error&gt;0)</td>
<td>0.347*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.19)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Error x 1(Error&lt;0)</td>
<td>0.159***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Err x Gender Match&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.439***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.06)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Err x Gend No Match</td>
<td>0.284***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Num. Obs</td>
<td>2475</td>
<td>2475</td>
<td>1200</td>
<td>2475</td>
<td>2475</td>
<td>2445</td>
<td>2321</td>
</tr>
</tbody>
</table>
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 ≈ $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” ($\sigma \rightarrow \emptyset$), “Confabulation,” ($\sigma \rightarrow \sigma'$), “Delusion” ($\emptyset \rightarrow \sigma$)

<table>
<thead>
<tr>
<th></th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td>s = G</td>
<td>$a_G : CR$</td>
<td>$b_G : \text{Negative C}$</td>
<td>$c_G : \text{Negative A}$</td>
<td>$d_G : \text{Weak Negative A}$</td>
</tr>
<tr>
<td>s = B</td>
<td>$a_B : \text{Positive C}$</td>
<td>$b_B : CR$</td>
<td>$c_B : \text{Positive A}$</td>
<td>$d_B : \text{Weak Positive A}$</td>
</tr>
<tr>
<td>s = $\emptyset$</td>
<td>$a_\emptyset : \text{Positive D}$</td>
<td>$b_\emptyset : \text{Negative D}$</td>
<td>$c_\emptyset : CR$</td>
<td>$d_\emptyset : \text{Weak CR}$</td>
</tr>
</tbody>
</table>
Memory biases conditional on performance
Asymmetric recall of (un)fairness

- "Asymmetric Memory Recall in Social Interactions" (Li 2012)
- Trust Game: A trusts or not, if trusts B reciprocates or not.
- Strategy method. Then, after 0, 7 and 43 days: incentivized recall

Results:

1. A player whose trust was betrayed is more likely to forget the act than one for whom was reciprocated
2. A player whose trust was betrayed is more likely to forget her trusting decision than one who did not trust
3. A player who committed an unkind act perceives it as less unkind as time elapses


Stakes-dependent beliefs

- Mayraz (2011) “Wishful Thinking”

- 145 subjects, observe chart of “historical wheat prices”, then 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
IV - Social Beliefs / Cognition

- “Groupthink: Collective Delusions in Organizations & Markets (RB 2013)
- What interaction structures lead (mis)beliefs to spread, or on the contrary to dampen across agents?
  - Will do here with anticipatory utility but more general (e.g., KP).

<table>
<thead>
<tr>
<th>Period 0</th>
<th>Period 1</th>
<th>Period 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H \rightarrow H$</td>
<td>$e^i = 0, 1$</td>
<td>$s E_1 \left[ U_2^i \right]$</td>
</tr>
<tr>
<td>$L \rightarrow L$</td>
<td>action choice</td>
<td>anticipatory feelings: hope, dread, anxiety…</td>
</tr>
<tr>
<td>signal about project value $\theta$</td>
<td>recall (attention, awareness)</td>
<td>cost $ce^i$</td>
</tr>
</tbody>
</table>

- Take here simplest interaction / organization structure; can enrich substantially (e.g., asymmetries)
- Stakes now endogenous: $\kappa_\sigma^i = \theta (1 - \alpha)e^{-i}$, $\sigma = H, L \Rightarrow$ cognitive linkages
Dealing with unpleasant realities (state \( L \))

- Incentive to deny \( \sigma = L \):

\[
\Delta U_i^0 \equiv U_{0,\text{Denial}}^i - U_{0,\text{Realism}}^i = -\frac{m}{\beta} - \delta \left[ c - (\delta + s) \alpha \theta_L \right]
\]

\[
+ \delta \sigma \left[ \alpha (\theta_H - \theta_L) + (1 - \alpha) [\theta_H - (1 - \lambda^{-i}) \theta_L] \right],
\]

\[
r(\lambda^i) = \frac{q}{q + \chi (1 - q)(1 - \lambda^i)}
\]

- \( \lambda^i \): agent \( i \)'s equilibrium realism (recall of \( L \) signals)
- \( \lambda^{-i} \): other agents’ equilibrium degree of realism
Mutually Assured Delusion (MAD) principle

Note that:

\[
\frac{\partial \Delta U_i^j}{\partial (1 - \lambda^{-i})} = sr(\lambda^i) \cdot (1 - \alpha)(0 - \theta_L)
\]

- When reality avoidance by others is beneficial (positive externalities), individual cognitive strategies are strategic substitutes
  - Others’ disregard of bad news makes such news less bad, easier to accept
- When reality avoidance by others is detrimental (positive welfare externalities), individual cognitive strategies are strategic complements
  - Others’ reality denial makes future prospects even worse, so bad news more scary, harder to face
- “Psychological multiplier” ⇒ interdependent beliefs and actions
Proposition (groupthink)

1. Both realism ($\lambda = 1$) and collective denial ($\lambda = 0$) are equilibria, for $s$ within some range, iff

$$\text{Prob(state L)} \times (\theta_H - \theta_L) < (1 - \alpha) (0 - \theta_L).$$

2. Groupthink more likely when more “common fate”, few exit options ($\alpha \downarrow$); more tail risk, worse bad news ($1 - q \downarrow \theta_L \downarrow$): “black swans”.

- Culture of denial: all persist in wrong course of action, ignoring the red flags –because others do (thereby making reality worse for everyone)

- Testable implications: e.g., vary payoff structure in experiments
Hierarchies: top-down groupthink

- **Cognitive dependency**: agents $i$’s realism, $\lambda^i$, influenced most by how key contributors to his welfare deal with $L$

- Simple hierarchy: agent 1 = manager, 2 = worker(s)

- Manager delusions (e.g., overinvestment, overborrowing) hurt workers $\gg$ the reverse (stay on the job, should look for another):
  
  $b^{12}_L - a^{12}_L$, large, $b^{21}_L - a^{21}_L$ small $\Rightarrow$ unique equilibrium, with...
Follow the leader...

"Trickle down" of beliefs in a hierarchy
V - “Irrational Exuberance” in Asset Markets

- Continuum of firms, investors. Can produce or invest $k^i \leq K$ at $t = 0$ with cost 0, and additional $e^i \leq E$ at $t = 1$ cost $c$

- All units are sold at $t = 2$. Time to build, limited liquidity, no short sales \sim limits to arbitrage,

- Market price $P_\sigma(\bar{k} + \bar{e})$, reflects
  - total supply: $\bar{k} + \bar{e} \in [0, K + E]$
  - variable market conditions: $\sigma = H, L$

- Unchanged information structure, preferences
Is investor exuberance contagious?

- Does other market participants’ exuberance (denial of bad news) make each individual more or less likely to also be bullish?

- General obliviousness to weak fundamentals will further depress the (expected) final price: \( P_L(K + E) << P_L(K) \)

Glut, market crash \( \Rightarrow \) two effects:

- **Substitutability:** if \( i \) remains bullish, will lose even more money on the extra \( E \) units which will produce / invest at \( t = 1 \),

  \[
  [c - P_L(K + E)] E \quad \text{v.s.} \quad [c - P_L(K)] E
  \]

- **Stakes:** if bearish, even greater capital losses must be immediately acknowledged on outstanding position \( k^i \)

  \[
  \Delta \kappa^i = [P_H(K + E) - P_L(K + E)] k^i \quad \text{vs.} \quad \Delta \kappa^i = [P_H(K + E) - P_L(K)] k^i
  \]
With appropriate conditions:

- **Escalating commitment / sunk cost effects**: the more agent $i$ has invested to date ($k_i$), the more likely he is to continue “blindly” / the less likely to be a realist.

- **Market momentum**: the greater was aggregate prior investment ($K$), the more likely each agent is to continue investing “blindly”

- **Contagious beliefs**:

---

**Proposition (market manias and crashes)**

If prior $q$ is high enough and $P_H(K + E)(1 + E/K) < c/\delta$,

1. There is a range of $s$ in which both realism and blind “exuberance” in the face of adverse news are equilibria.

2. Market mania leads to overinvestment and eventual crash.
“Wall Street and the Housing Bubble”

Cheng, Raina & Xiong (2014)

Figure 1: Home Price Indices
This figure plots the Case-Shiller non-seasonally-adjusted home price indices from January 2000 through July 2012. Values for January 2000 are normalized to 100.
Bad incentives or bad beliefs?

- **Standard account:** poorly designed incentives led Wall Street to take excessive risks in the housing market, leading to disastrous consequences: securitizing mortgages with very lax screening of subprime borrowers, liar loans, etc.
  - Unscrupulous insiders, knowingly deceiving households, banks, investors

- But: what did insiders really believe? Can we tell?

- Identify + track down **own housing transactions** of 400 securitization managers, issuers, investors: “securitization agents” comprising vice presidents, senior vice presidents, managing directors, and other non-executives at major investment houses and boutique firms

- **Control groups:**
  - S&P 500 equity analysts who do not cover homebuilding companies
  - Random sample of lawyers who did not specialize in real estate law.
Second-home purchases

Panel B
Home divestures (sales)
Key findings

- Securitization agents increased rather than decreased, their housing exposure during the boom period, particularly through second home purchases and swaps of existing homes into more expensive homes.

- Were also much slower to sell once prices had started falling.

- Difference is not explained by interest rates or financing, and is more pronounced in relatively bubblier Southern California compared to the New York metro region.

- Accords well with stakes-dependent beliefs.

- As a result, securitization agents' overall home portfolio performance was significantly worse than that of control groups.

- Agents working on the sell-side and for firms which had poor stock price performance through the crisis did particularly poorly themselves.
Further research themes

- Bad incentives and bad beliefs: complements, not substitutes
  - Contract theory, mech. design with wishfully thinking / rationalizing agents
- Long-term dynamics of motivated learning
  - Gottlieb (2011) “Will you Never Learn”? (Answer:... nope)
- Alternative theories / mechanisms for contagious beliefs in organizations, markets, politics, religion.
- Feedback with actors whose actions affect the supply of (hard or soft) information
  - Levy (2012) “Soothing Politics”: complementarities between politician’s actions and voters’ cognitive strategies
  - Propaganda, commercial advertising, etc.
- More experimental work, incl. neuro and field / empirical, on individuals but especially collective / contagious beliefs distortions.
- Next, look at yet another important set of beliefs that
  - Vary considerably across places, and are extremely persistent
  - Turn out to be strikingly associated with important economic outcome
VI - Religion

- **The Big One.** What is (always has been) the number one instance of motivated beliefs, which people choose, value, maintain, defend?

- Large traditional literature on the economics of religion, has mostly emphasized the “club goods” aspect: group formed to provide spiritual services (Iannaccone 1992, 1998) *insurance* (Steve & Statavage 2006), commitment by reducing outside options (Berman 2000)

- **Link with human capita.** Barro & McCleary (2003a), Botticini & Eckstein (2012)

- Important, but nothing specific here about beliefs per se


- More recently, starting to focus on religious beliefs per se - their functions, how maintained, etc.
  - Benabou & Tirole (2006): Belief in a Just [after-] World [or not] as self-sustaining social beliefs (equilibria)
  - Levy & Razin (2012): religious beliefs help screen agents likely to be more trustworthy / cooperative with each other
“Forbidden Fruits: The Political Economy of Science, Religion and Growth’
(Bénabou, Ticchi & Vindigni 2013)

- New, rather striking empirical findings: across countries as well as US states, strong negative relationship between religiosity and innovation (patents/capita)
- Very robust, eg., controlling for income/capita, population, higher education, religious freedom

- More directly fitting for 1st Atkinson lecture + brand new findings...
  - Rough & very preliminary
  - Not worked out model yet
Religiosity and inequality across countries
Controls: GSP/capita, religious freedom
Religiosity and inequality in the US

Inequality

Religiosity (very important)
**Controls:** GSP per capita, Population, Fraction with at least Bachelor's Degree