Unrealistic Expectations and Misguided Learning

Paul Heidhues[1] Botond Kőszegi[2] Philipp Strack[3]

[1]Düsseldorf Institute for Competition Economics

[2]Central European University

[3]University of California, Berkeley

November 7, 2017

Motivation

Stylized fact: people have unrealistically positive views of their traits and prospects.

- Old literature in psychology and experimental economics.
  - Individual behavior (Burks et al. 2013, Benoit et al. 2014, Charness et al. 2014).
  - Entry in experimental market (Camerer and Lovallo 1999).
- Also documented in field settings.
  - CEO behavior (Malmendier and Tate 2005, 2008).
  - Unemployment (Spinnewijn 2014).
  - Work (Hoffman 2014).
  - Health (Oster et al. 2013).

Natural question: how these individuals update their beliefs.

Already explored: updating about traits/prospects.

We study the beliefs of overconfident individuals about other variables affecting optimal behavior.

Main Idea

Main mechanism — in the context of overconfidence and delegation.

- Due to unrealistic expectations, agent is surprised about team output.
- He becomes more pessimistic about teammate, and concludes he must do more himself.
- This adjustment in behavior makes things worse, perpetuating the misdirected learning further.

We characterize the situations in which such misdirected learning occurs, and explore implications.

Related Literature

We're connecting two literatures, that on the implications of **overconfidence** and that on **learning with misspecified models**.

Most related learning papers: Eyster and Rabin (2015), Esponda and Pouzo (2016), Fudenberg, Romanyuk and Strack (2016), Spiegler (2016), Le Yaouanq and Nestermann (2017).

Most ignored literature: that studying the positive consequences of unrealistic self-views.

Setup

Output in period t is $q_t = Q(e_t, a, \Phi) + \epsilon_t = a + \Phi - L(e_t - \Phi) + \epsilon_t$, where L is a symmetric loss function with $|L'(x)| < k < 1$.

- Φ is a fixed fundamental whose prior has full support on $(\underline{\phi}, \overline{\phi})$.
- ϵ_t is iid with log-concave distribution.

Agent chooses myopically optimal action.

Key assumption: agent's true ability is A, but he believes with certainty that it is $\tilde{a} > A$.

- Tractable stand-in for forces that generate overconfidence.
- Can allow for biased learning about a.
  - But then, can only analyze limit beliefs.

Other than being overconfident, agent understands above, and makes correct inferences.

Applications (Beyond Delegation)

Control in organizations.

- $-e_t$: control/punishments/extrinsic incentives.
- More control lowers intrinsic motivation (Benabou and Tirole 2003) or morale (Fang and Moscarini 2005).
- Φ: baseline intrinsic motivation. The higher is Φ, the less control is optimal.

Public policy.

- e_t: extent of drug liberalization.
- $-q_t$: drug-related crime/problems.
- $-\Phi$: social fundamentals affecting drug problems.
- Other example: degree of deregulation.

Working.

- $-e_t$: amount of work.
- q_t: life satisfaction.
- $-\Phi$: social norm for how much one should work.

Fixed Action

Suppose e_t is fixed at e.

- Agent observes iid signals

$$q_t = A + \Phi - L(e - \Phi) + \epsilon_t.$$

  Classical problem in econometrics.
- Because ϵ_t averages out, agent assigns probability one to $\tilde{\phi}_\infty$ given by

$$\tilde{a} + \tilde{\phi}_\infty - L(e - \tilde{\phi}_\infty) = A + \Phi - L(e - \Phi)\,.$$

- Since $\tilde{a} > A$, this yields $\tilde{\phi}_\infty < \Phi$.

A kind of self-serving attributional bias.

- Because agent believes he's able, he attributes bad outcomes to external factors.

Endogenous Action

Now suppose e_t is chosen by the agent each period.

Use heuristic graphical argument when agent can change his action only increasingly rarely.

- Then, agent updates only based on output corresponding to last action.

Graphical Analysis

Intuition: control in organizations.

- Output below expected.

$\Rightarrow$ "I must control worker more."

$\Rightarrow$ Even lower output.

$\Rightarrow$ "Need to control them even more."

Manager believes he is learning about an unmotivated workforce ... but is in fact creating it.

Comparative Static: The Importance of Being Right

Graphical Analysis

More important for action to be close to fundamental $\Rightarrow$ it ends up further!

- Agent is hurting himself more $\Rightarrow$ to explain in a consistent way, he must become more pessimistic.

Jumping to Outside Options

Suppose that the agent has an outside option with true (and perceived) utility $\underline{u}$ satisfying $\tilde{a} > \underline{u} > A - L(\tilde{\phi}_\infty)$.

Then, agent first enters but eventually quits task.

If $\underline{u} > A$, then he shouldn't enter or persist in first place.

- Well-understood implication of overconfidence.

If $\underline{u} < A$, overconfidence leads to suboptimal exit.

- The agent eventually stops performing the task *because* he overestimates his ability to do well in it.
- He's especially prone to exit if he overestimates outside option.
- E.g., getting frustrated with coauthor and looking for new one.

Prediction that overconfident can be too prone to exit or jump between tasks contrasts with received wisdom.

No conclusive evidence, but consistent with

- observation that many documented effects of overconfidence in economic settings, such as mergers or innovations, pertain to *new* directions; and
- Landier and Thesmar's (2009) finding that serial entrepreneurs are more overconfident.

Observe: when the agent exits, his learning doesn't affect future beliefs regarding environments where Φ doesn't apply.

- In fact, he tends to seek out exactly these situations.

Underconfidence

Suppose the agent is underconfident ($\tilde{a} < A$). Graphical Analysis

Asymmetry relative to overconfidence: loss under limiting beliefs is $< \Delta$. Intuition:

- Suppose mean prior is correct.
- Surprisingly large output $\Rightarrow$ revise beliefs about fundamental upwards $\Rightarrow$ increase action.
- Resulting loss $\Rightarrow$ reassess.
- So misdirected learning is self-limiting.

Convergence with Endogenous Actions

Problem: beliefs do not even concentrate for arbitrary sequences of endogenous actions!

- Hence, to control beliefs one needs to control actions.
- But, actions depend on beliefs.
- In general, no convergence (e.g., Nyarko 1991, Fudenberg, Romanyuk and Strack forthcoming).
- Sequence of beliefs is function-valued process (which is not Markov in, e.g., the mean).
- We establish convergence for situations in which there is a unique stable belief.
- Use a novel idea in the literature on learning with misspecified models: look at extremal beliefs.

Graphical Idea of Convergence Proof

When Does Self-Defeating Learning Occur?

"Self-defeating learning" $\equiv$ ability to update action makes agent with approximately correct prior about the fundamental worse off.

Basic assumptions: Q is twice differentiable, $Q_a, Q_\phi > 0$, and Q is strictly quasi-concave in e ($\Rightarrow$ unique myopically optimal action).

Self-defeating learning requires optimal action to be sensitive to ϕ.

- If output takes the form $V(S(e_t, a), \phi)$, agent never changes action.
  - Although incorrect inference occurs, it's not self-reinforcing.
- Overconfidence is often considered beneficial in these settings.
  - E.g., ability and effort are complements, and other factors (if any) are additively separable.

Hence, **assume from now on that** $Q_{e\phi} > 0$.

Insights illustrated so far generalize if $Q_{ea} \leq 0$**.**

What if $Q_{ea} > 0$?

- Then, initial action is too high.
- Revision of beliefs about the fundamental downwards lowers action next time, possibly increasing output.
- Two possibilities:
  - In the limit updating increases output.
  - Updating lowers the action below optimal, at which point further misdirected learning is self-reinforcing.

Optimal Learning

Proposition

The following are equivalent:

I. *For any A, $\tilde{a}$, and Φ the agent's limiting action is optimal.*

II. *The agent's limiting action is identical to that with an output function of the form $V(e_t, S(a, \phi))$.*

That is, learning is indistinguishable from that of a realistic agent if and only if problem is *not* identifiable.

- a and ϕ don't have independent effects on output $\Rightarrow$ misinference about ϕ can exactly compensate overconfidence about a.
- E.g., optimal effort depends on total ability of team.

Conclusion: qualitatively, self-defeating learning occurs if the optimal action depends on the fundamental, and either depends less on ability or does so in the opposite way.

Will the Agent Get a Clue?

Natural question: might agent's observations lead him to conclude that something about his beliefs is awry?

If beliefs and actions converge, we identify a strong sense in which they won't — even after he has observed infinite data.

Specification test:

- Agent's beliefs converge to $\tilde{\phi}_\infty$.
- This allows him to extract perceived noise terms in hindsight: $\tilde{\epsilon}_t = q_t - Q(e_t, \tilde{a}, \tilde{\phi}_\infty)$.
- Empirical distribution of $\tilde{\epsilon}_t$ should match true distribution of ϵ_t.

Proposition

It does.

Intuition: agent settles on beliefs that lead him to predict average output accurately, so that he also extracts the noise terms accurately.

Conclusion

Noise inside the production function: $q_t = Q(e_t, a, \Phi + \epsilon_t)$.

Analyzed in previous version, with normally distributed ϵ_t and prior on Φ.

- Data is subjectively iid, so no experimentation motive.
- Show based on stochastic approximation theory that beliefs always converge, even when multiple consistent beliefs.

Currently working on: effect of overconfidence on opinions about in-group and out-group.

Future questions to ask:

- Learning dynamics when beliefs about ability are not degenerate.
- Implications for authority in organizations.
- How do overconfident agents interact?