Learning with Misspecified Models: the case of overconfidence

Jimena Galindo

September 29, 2023

Overconfidence is Costly

OVERCONFIDENCE: Belief that my type is higher than it truly is ("overestimation" as in Moore and Healy (2008))

Overconfidence is Costly

OVERCONFIDENCE: Belief that my type is higher than it truly is ("overestimation" as in Moore and Healy (2008))

It seems to be persistent in various settings.

- Excess entry of entrepreneurs (Camerer and Lovallo, 1999)
- Suboptimal genetic testing and healthcare (Oster et al. 2013)
- Workers overestimate their productivity (Hoffman and Burks, 2020)

Ultimately it leads to sub-optimal choices

Models of Learning

Focus on setting with 2 parameters:

- An **Ego-Relevant** parameter
- An **Exogenous** parameter

Some of the features that theory has incorporated to explain overconfidence are:

- Dogmatism
- Paradigm shifts
- Motivated beliefs
- Myopic optimiztion

Four Theories of Misspecified Learning

In settings with more than 1 unknown parameter:

1. Self-defeating equilibrium (Heidhues et al. (2018)):
    - Bayesian about exogenous parameters
    - Dogmatic about ego-relevant parameters
2. Bayesian Likelihood Ratio test (Schwarstein and Sunderam (2021), Ba (2022)) :
    - Bayesian about exogenous parameters
    - Paradigm shift for ego-relevant parameters
3. Motivated Beliefs / Self-Attribution Bias (Brunnermeier and Parker (2005), Benjamin (2019)):
    - Optimally biased updating
    - Utility from held beliefs
4. Myopic Bayesian (Hestermann and Le Yaouanq, (2021))
    - Bayesian about Both

Questions

Which of the proposed theories better explains the observed behavior?

- Do we observe heterogeneity in the use of misspecified models?

Is ego-relevance of the parameter a key feature for the misspecification?

- Are ego-relevant misspecifications more likely to persist?
- Can the same theories be used to explain the prevalence of stereotypes?

An Example

A student has **unknown intrinsic ability** θ^* (ego-relevant)

They choose a level of effort $e \geq 0$.

Effort and ability are evaluated by a grading system ω (exogenous)

The student wants to maximize utility

$$y = (\theta^* + e)\omega - \frac{1}{2}e^2 + \varepsilon$$

An Example

A student has **unknown intrinsic ability** θ^* (ego-relevant)

They choose a level of effort $e \geq 0$.

Effort and ability are evaluated by a grading system ω (exogenous)

The student wants to maximize utility

$$y = (\theta^* + e)\omega - \frac{1}{2}e^2 + \varepsilon$$

Regardless of their own type and of their beliefs about it, they should choose $e^*(\omega) = \omega$

Learning is Possible

This exercise is repeated for $t = 0, 1, \ldots$

$$y_t = (\theta^* + e_t)\omega - \frac{1}{2}e_t^2 + \varepsilon_t$$

Note that both parameters are identified in this setting:

- Choosing $\hat{e}$ and $\hat{e} + 1$ over multiple periods allows identification of ω
- Once ω is known, θ can be backed out

How come people don't learn their true type and don't choose the optimal effort?

Road-map

1. Unifying Framework
2. Mechanisms and Predictions
3. Experimental Design
4. The Data
5. Parameter Estimation
6. Results

Framework

A Unifying Framework

Finite type space: $\theta \in \{\theta_H, \theta_M, \theta_L\}$

Finite state space: $\omega \in \{\omega_H, \omega_M, \omega_L\}$ with $p(\omega_k) = 1/3$

Finite action space: $e \in \{e_H, e_M, e_L\}$

Binary signal: $P\left[Success|e, \omega, \theta\right]$ where p is an order-preserving transformation of $u(x)$

The Data Generating Process

The probability of success is given by:

θ_L

	ω_H	ω_M	ω_L
e_H	50	20	2
e_M	45	30	7
e_L	40	25	20

θ_M

	ω_H	ω_M	ω_L
e_H	80	50	5
e_M	69	65	30
e_L	65	45	40

θ_H

	ω_H	ω_M	ω_L
e_H	98	65	25
e_M	80	69	35
e_L	75	55	45

The Data Generating Process

	ω_H	ω_M	ω_L
e_H	blue!2550	20	2
e_M	45	blue!2530	7
e_L	40	25	blue!2520

θ_L

	ω_H	ω_M	ω_L
e_H	blue!2580	50	5
e_M	69	blue!2565	30
e_L	65	45	blue!2540

θ_M

	ω_H	ω_M	ω_L
e_H	blue!2598	65	25
e_M	80	blue!2569	35
e_L	75	55	blue!2545

The Data Generating Process

$\longleftarrow$ $\longleftarrow$ $\longleftarrow$

	ω_H	ω_M	ω_L
e_H	blue!2550	20	2
e_M	45	blue!2530	7
e_L	40	25	blue!2520

θ_L

	ω_H	ω_M	ω_L
e_H	blue!2580	50	5
e_M	69	blue!2565	30
e_L	65	45	blue!2540

θ_M

	ω_H	ω_M	ω_L
e_H	blue!2598	65	25

A Stable Misspecified Belief

	ω_H	ω_M	ω_L
e_H	[HTML]b84f7950	20	2
e_M	45	30	7
e_L	40	25	20

θ_L

	ω_H	ω_M	ω_L
e_H	80	[HTML]b84f7950	5
e_M	69	[HTML]f09ebe65	30
e_L	65	[HTML]f09ebe45	40

θ_M

	ω_H	ω_M	ω_L
e_H	98	65	25
e_M	80	69	35
e_L	75	55	45

θ_H

The Self-Confirming Equilibria

	ω_H	ω_M	ω_L
e_H	[HTML]b84f7950	20	2
e_M	45	[HTML]5f94b830	7
e_L	[HTML]69a35b40	25	20

θ_L

	ω_H	ω_M	ω_L
e_H	80	[HTML]b84f7950	5
e_M	[HTML]fab14369	65	[HTML]5f94b830
e_L	65	[HTML]9662f045	[HTML]69a35b40

θ_M

	ω_H	ω_M	ω_L
e_H	98	65	25
e_M	80	[HTML]fab14369	35
e_L	75	55	[HTML]9662f045

An Example

- True type is θ_M
- True parameter is $\omega_M \rightarrow$ the student believes it is uniformly distributed

θ_L

	ω_H	ω_M	ω_L
e_H	50	20	2
e_M	45	30	7
e_L	40	25	20

θ_M

	ω_H	ω_M	ω_L
e_H	80	blue!2550	5
e_M	69	blue!2565	30
e_L	65	blue!2545	40

θ_H

	ω_H	ω_M	ω_L
e_H	98	65	25
e_M	80	69	35
e_L	75	55	45

The Dogmatic Modeler

Holds a degenerate belief: type is $\hat{\theta}$ with probability 1

Their belief is potentially misspecified:

- Overconfident if $\hat{\theta} > \theta^*$
- Underconfident if $\hat{\theta} < \theta^*$

Updates $p_t(\omega)$ using Bayes Rule

The Dogmatic Modeler: Mechanism

- A student who dogmatically believes he is θ_H

1. Chooses e_H and is disappointed $\rightarrow$ adjust belief about ω downward
2. Eventually chooses e_M and is disappointed as well $\rightarrow$ adjust belief about ω
3. Eventually chooses e_L and falls into a self-confirming equilibrium

Dogmatic Overconfident: Simulated

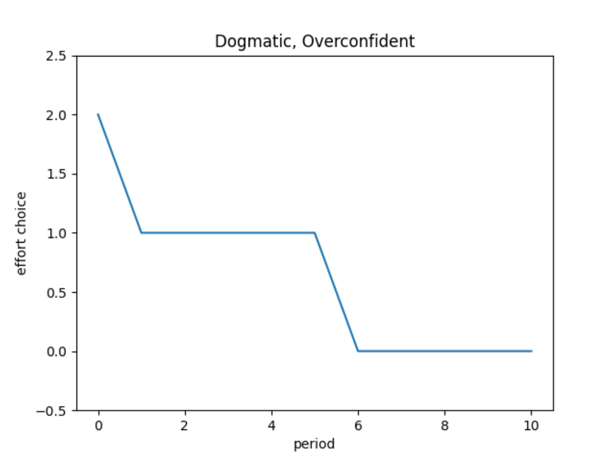


Figure 1: $\theta^* = \theta_M$, $\hat{\theta} = \theta_H$, $\omega^* = \omega_M$

The Switcher (paradigm shifts)

Same initial belief as the Dogmatic, but is willing to consider and alternative paradigm θ'

Keeps track of the likelihoods of the two possible paradigms:

- $p(\cdot|h^t)$ for $\hat{\theta}$ and θ'

They swithch to whichever paradigm is morelikely to have generated the signals

$$\frac{p(\theta'|h^t)}{p(\hat{\theta}|h^t)} > \alpha \geq 1$$

The Switcher: Mechanism

1. Chooses e_H and is disappointed $\rightarrow$ adjust belief about ω downward
2. Eventually chooses e_M and is disappointed as well $\rightarrow$ adjust belief about ω
3. Eventually chooses e_L and falls into a self-confirming equilibrium
4. At some point, the likelihood of θ_M becomes much larger than that of θ_H and the agent updates their belief

Switcher Overconfident: Simulation

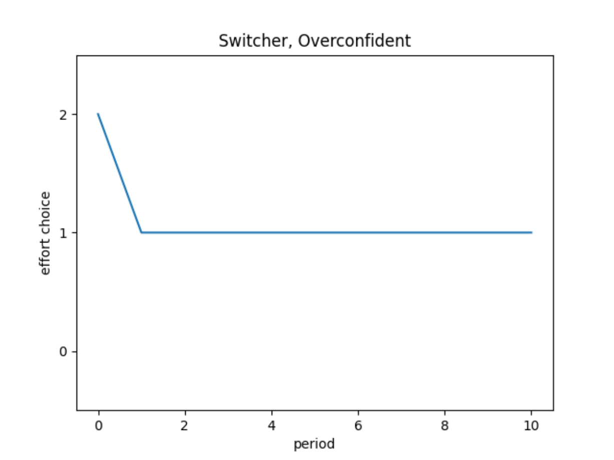


Figure 2: $\theta^* = \theta_M$, $\hat{\theta} = \theta_H$, $\omega^* = \omega_M$, $\alpha = 1.1$

Self-Attribution Bias / Optimal Expectations

Start with a diffused prior over (θ, ω) but updates with a bias

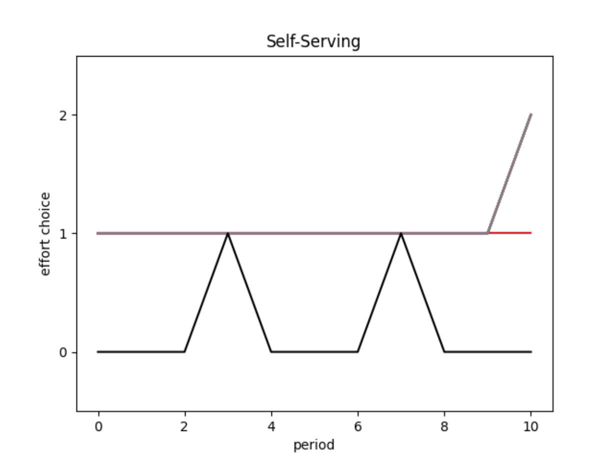
- Success $\to$ overweight parametrizations with $\theta > \omega$
- Failure $\to$ underweight parametrizations with $\theta < \omega$

$$p_{t+1}(\theta, \omega | s_t) =$$

Self-Attribution: Mechanism

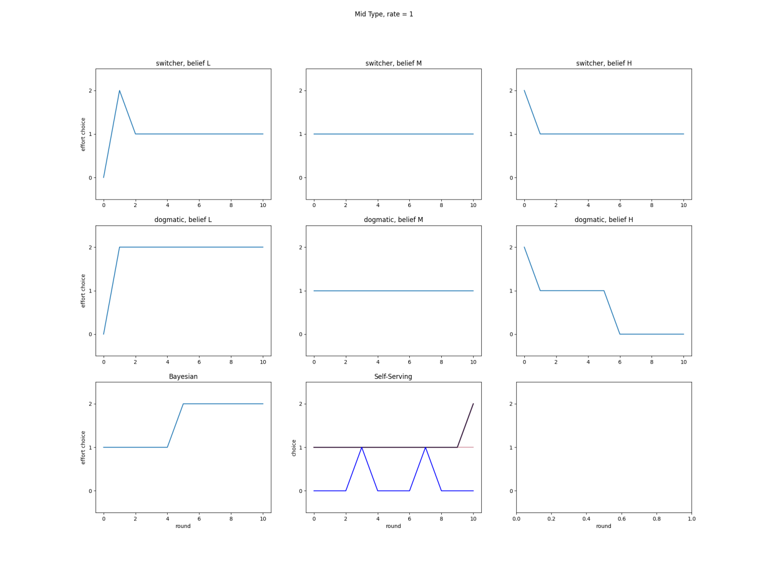
1. Chooses e that maximizes utility according to priors
2. Belief on ω deteriorates a lot after bad news $\rightarrow$ big change in effort
3. Belief on θ increases a lot after good news $\rightarrow$ small positive (or negative) change in effort

Self-Attribution: Simulation



The Myopic Bayesian

All Models



Mid Type, rate = 1
switcher, belief L
switcher, belief M
switcher, belief H
dogmatic, belief L
dogmatic, belief M
dogmatic, belief H
Bayesian
Self-Serving
effort choice
choice
round

Experimental Design

Set the Types

- Quiz: Answer as many questions as you can in 2 minutes
  - Math, Verbal, Pop-Culture, Science, Us Geography, Sports and Video games

- How many questions do you think you answered correctly in each quiz?
  - o 0 to 5
  - 6 to 15
  - 16 or more
- How sure are you about your choice?
  - Random guess $\rightarrow 1/3$
  - Another is equally likely $\rightarrow 1/2$
  - Fairly certain $\rightarrow 3/4$
  - Completely sure $\rightarrow 1$

Choice and Update

Effort choice and feedback (One topic at a time)

- Choose an effort
- Receive a sample of 10 signal realizations

x 11 per topic

Eliciting Beliefs?

- $E[\omega]$ is revealed by their choice of effort
- Eliciting beliefs for θ can incentivize learning in a way that is not consistent with the theory

Allow them to see the success rate matrix for only one type.

- Track the matrices they choose to see in each round

Stereotype condition

Observe the characteristics of a participant

- Gender,
- US National or not

Answer the same questions about slef and other

Belief updating and effort choice:

- The DGP depends on the θ the other participant

x 11 per topic

Screen

Based on the other participant's Science and Technology Quiz results

Which probability matrix would you like to see?

Low Score | Mid Score | High Score

Your Previous Outcomes

Choice	Successes	Failures
You have no data for this task yet		

See History

Next

Screen

Based on the other participant's Science and Technology Quiz results

Which probability matrix would you like to see?

Low Score | Mid Score | High Score

Choose a gamble	:	Rate A	Rate B	Rate C
A	○	40	45	65
B	○	30	65	69
C	○	5	50	80

Your Previous Outcomes

Choice	Successes	Failures
You have no data for this task yet		

See History

Next

The Data

The Data

Subject pool:

- Run at the CESS lab in person
- 45 subjects in Ego
- 33 subjects in Stereotype

The Sessions:

- 8 sessions
- 45 minutes on average
- Average payment: $23
    - $10 show-up fee
    - $0.20 per correct answer
    - $0.20 per success
    - Paid one topic at random

Learning

Initial Misspecifications

The Stereotypes

Transitions

Parameter Estimation

Identification of α

Estimation of Self-Attribution Bias

Results

Model Fit

Heterogeneity

The end

Thank you!