

# Learning from Data Through Models

Jimena Galindo

September 05, 2023

## Abstract

TBW

## 1 Introduction

## 2 Related Literature

## 3 Framework 1

An agent is of type  $\theta \in \Theta$  and faces an unknown exogenous state  $\omega$  drawn from some density  $f$  over  $\Omega$ . The agent knows the distribution of  $\omega$  but not its realized value. His prior belief about his type is  $p_0(\theta)$  and his belief about the state  $p_0(\omega)$  coincides with the true distribution  $f$ . Let the agent's type be  $\theta^*$  and the realized state be  $\omega^*$ .

An agent has a *misspecified* belief if the prior assigns probability zero to their true type. Furthermore, the agent is *dogmatic* if he holds a degenerate belief that places probability one on being a particular type,  $\hat{\theta}$ . A dogmatic agent can be dogmatic and misspecified. That means that  $\hat{\theta} \neq \theta^*$  and  $p_0(\hat{\theta}) = 1$ .

The agent chooses an action  $a \in A$  and observes a noisy outcome  $h \in H$ . The outcome is a function of the agent's type, the state, and the action. In particular  $h = h(\theta^*, \omega^*, a) + \varepsilon$

with  $h(\cdot)$  increasing in both  $\theta^*$  and  $\omega^*$ ; and such that conditional on a pair of parameters  $(\theta, \omega)$ , there is a unique optimal action.  $\varepsilon \sim N(0, \sigma)$  is noise in the output.

After observing the outcome, the agent updates his beliefs about  $\theta$  and  $\omega$  using some algorithm and moves on to the next period. He repeats this process infinitely many times. I make the simplifying assumption that the agent is myopic and chooses the action that maximizes the payoff in each period. This assumption is crucial for the results as the agent will not trade off the flow payoff for learning the truth as a forward-looking agent would.

Within this framework, I consider two nested theories of belief updating. The first one is a dogmatic modeler from Heidhues et al. [2018]. The second one is a switcher, as in Ba [2023]. The dogmatic modeler can be seen as a switcher with an infinitely sticky initial belief and this is the sense in which the two are nested.

A key notion in this setting is that of a self-defeating equilibrium<sup>1</sup>. A *self-defeating equilibrium* is a belief and action pair such that the agent’s belief about their type is misspecified and the outcome generated by the action is consistent with the misspecified belief. The average outcome under the true type and the true state equals the average output the agent expects under the misspecified belief. In addition, the agent’s belief is said to be *stable* when this happens.

### 3.1 The Dogmatic Modeler

A dogmatic agent does not update their beliefs about  $\theta$ ; instead he holds a degenerate belief that places probability one on being a particular type,  $\hat{\theta}$ , which is potentially not his true type. In this case, no matter how much information he gathers against being of type  $\hat{\theta}$ , he will not update his beliefs. Any discrepancies between the observed outcomes and his believed type are incorporated using the Bayes rule to update their beliefs about  $\omega$ .

A key difference between the dogmatic agent and a fully Bayesian agent is that the dogmatic agent does not aggregate across types when updating their beliefs about  $\omega$ . This

---

<sup>1</sup>This notion is an adaptation of the Berk-Nash equilibrium in @Esponda2016

means the dogmatic agent will never learn their true type if they are misspecified.

Heidhues et al. [2018] show that, under certain assumptions on the per-period utility,<sup>2</sup> about the outcome  $h$ , a dogmatic modeler will inevitably fall into a self-defeating equilibrium where the outcomes they observe reinforce their belief on  $\omega$  in such a way that as  $t \rightarrow \infty$  the agent will be sure that the state is some  $\omega'$  consistent with their believed type and the observed data.

The mechanism by which the dogmatic agent falls into the self-defeating equilibrium is the following: Suppose the agent holds the misspecified belief that they are type  $\hat{\theta} > \theta^*$ . For any prior over  $\omega$ , the agent will be disappointed by the outcome; he expected a gain of  $h(\hat{\theta}, \mathbb{E}(\omega), a)$  but instead observes  $h(\hat{\theta}, \omega^*, a)$ . There are two possible sources for the disappointment. The first is that the realized state is lower than the expected state. The second is that the agent is of type  $\theta^*$  and thus for all possible states, his gain will be lower than what he expected. Because the agent is dogmatic, he will not update his beliefs about  $\theta$  and thus will attribute the disappointment to the state being lower than expected. Heidhues et al. [2018] give sufficient conditions for the agent to fall into a self-confirming equilibrium where the agent will continue to choose an action that reinforces an incorrect belief about the state. This force is illustrated in Example 1.

**Example 1:** Set  $A = \Omega$  and  $H = [0, \infty)$  and consider a student with intrinsic ability  $\theta^* \geq 0$  who faces a grading procedure  $\omega^*$  that is unknown to them. However, they know that a higher  $\omega^*$  is more likely to yield a higher grade. In particular, assume the grade is given by  $(\theta^* + a)\omega^*$ .

The student must choose an effort level  $a$ , which determines their grade. For whatever the chosen effort level, the agent must pay a cost  $c(a) = \frac{1}{2}a^2$ . And he repeats this process for infinitely many periods. Assume also that the student's prior is such that  $\mathbb{E}[\omega] = \omega^*$ . Therefore, the student's payoff in period  $t$  is given by

---

<sup>2</sup>The assumptions are that  $u$  is twice continuously differentiable with: (i)  $u_{ee} < 0$  and  $u_e(\underline{e}\theta, \omega) > 0 > u_e(\bar{e}, \theta, \omega)$ , (ii)  $u_\theta, u_\omega > 0$  and (iii)  $u_{e\theta} < 0$  and  $u_{e\omega} > 0$ . The direction of the derivatives is a normalization and the results would hold even when the signs are reversed.

$$u_t(a_t; \theta^*, \omega^*) = (\theta^* + a_t)\omega^* - \frac{1}{2}a_t^2 + \varepsilon_t \quad (1)$$

Under this specification, the myopic optimal effort level is  $a_t^* = \omega^*$ . Nonetheless, because the agent does not know  $\omega^*$ , he will choose  $a_t = \mathbb{E}_t(\omega)$  where the expectation is taken with respect to the agent's belief at the beginning of period  $t$ . If he does not revise his effort choice for multiple periods, he will observe an average grade of  $(\theta^* + a_t^*)\omega^* - \frac{1}{2}a_t^{*2}$  but he was expecting an average grade of  $(\hat{\theta}\theta + a_t^*)\omega^* - \frac{1}{2}a_t^{*2}$

### 3.2 The Bayesian Benchmark

A Bayesian agent simultaneously updates their beliefs about  $\theta$  and  $\omega$  by using Bayes' rule. The posterior odds at period  $t$  about  $\theta$  after observing an outcome are given by:

$$\frac{p_t[\theta_H|\text{outcome}]}{p_t[\theta_M|\text{outcome}]} = \frac{p[\text{outcome}|\theta_H]p_{t-1}[\theta_H]}{p[\text{outcome}|\theta_M]p_{t-1}[\theta_M]} \quad (2)$$

and

$$\frac{p_t[\theta_M|\text{outcome}]}{p_t[\theta_L|\text{outcome}]} = \frac{p[\text{outcome}|\theta_M]p_{t-1}[\theta_M]}{p[\text{outcome}|\theta_L]p_{t-1}[\theta_L]} \quad (3)$$

Where  $p_{t-1}$  is the prior at period  $t$  and  $p[\text{outcome}|\theta] = \sum_{\omega} p[\text{outcome}|\theta, \omega, e]p_{t-1}(\omega)$  is the probability of observing the outcome given the agent's type and the effort chosen. The update is symmetric for  $\omega$ .

Bayesian agents always choose the effort level that maximizes their flow payoff by taking expectations over their prior beliefs about  $\theta$  and  $\omega$ . Since agents are myopic, even though all the parameters could be identified with enough variation in choices,

### 3.3 The Switcher

An agent is a *switcher* if they behave as a dogmatic but is willing to entertain the possibility that they are of a different type. In particular, when they start off as a misspecified dogmatic,

they are willing to switch to a different dogmatic belief if the data is convincing enough.

In order to abandon their initial dogmatic belief, the agent needs to observe a sequence of outcomes that are sufficiently unlikely to have happened if they were of the type they initially believed. They do so by keeping track of the likelihood that each of the possible types generated the data. If the likelihood ratio is sufficiently large, the agent will switch to the alternative and behave as if they are dogmatic about the new type.

In particular, for an agent that starts off with a dogmatic belief that they are of type  $\hat{\theta}$  but is willing to consider the alternative explanation that they are of type  $\tilde{\theta}$ , the agent will switch to the alternative if:

$$\frac{p[h^t|\tilde{\theta}]}{p[h^t|\hat{\theta}]} > \alpha$$

Where  $h^t$  is the history of outcomes up to time  $t$  and  $\alpha \geq 1$  is a threshold that determines how convincing the data needs to be for the agent to change their belief. Notice that if  $\alpha \rightarrow \infty$ , we get the dogmatic agent. In this sense, the switcher is a generalization of the dogmatic type, just as the self-attribution agent is a generalization of the Bayesian agent.

By allowing the agent to keep track of the likelihoods and switching to an alternative type, the switcher can avoid the self-confirming equilibrium the dogmatic agent falls into. However, if the prior belief on *theta* is sufficiently tight around a self-confirming equilibrium, the switcher might look identical to the dogmatic even in a case where  $\alpha$  is not too large.

EXPLAIN HOW THEY MIGHT GET OUT OF THE TRAP.

### 3.4 The Bayesian Agent

a fully Bayesian agent might not learn their true type. This happens because they do not internalize the tradeoff between flow payoff and learning and thus might not experiment enough to learn their type. An alternative to this approach is given by Hestermann and Yaouanq [2021] and is discussed with the results.

### 3.5 The Biased Agent

An agent that updates their beliefs with self-attribution bias will update their beliefs about the state  $\omega$  and their type  $\theta$  by over-attributing successes to a high value of  $\theta$  and under-estimating the role of higher  $\omega$ . Similarly, they will attribute failure to a low state more than an unbiased agent would. To model the self-serving attribution bias, I take the approach of Benjamin [2019], where a generalization of the Bayes rule above gives the update.

$$\frac{p_t[\theta_H|\text{outcome}]}{p_t[\theta_M|\text{outcome}]} = \left( \frac{p[\text{outcome}|\theta_H]}{p[\text{outcome}|\theta_M]} \right)^{c_s^\theta \mathbb{I}\{\text{success}\} + c_f^\theta \mathbb{I}\{\text{failure}\}} \frac{p_{t-1}[\theta_H]}{p_{t-1}[\theta_M]} \quad (4)$$

and

$$\frac{p_t[\theta_M|\text{outcome}]}{p_t[\theta_L|\text{outcome}]} = \left( \frac{p[\text{outcome}|\theta_M]}{p[\text{outcome}|\theta_L]} \right)^{c_s^\theta \mathbb{I}\{\text{success}\} + c_f^\theta \mathbb{I}\{\text{failure}\}} \frac{p_{t-1}[\theta_M]}{p_{t-1}[\theta_L]} \quad (5)$$

$c_s^\theta$  and  $c_f^\theta$  are the self-serving attribution bias parameters for the agent's type  $\theta$ . If  $c_s^\theta = c_f^\theta = 1$ , the agent is unbiased and the update is the same as the Bayesian update. On the other hand, if  $c_s^\theta > c_f^\theta$  the agent over-attributes success to their type and under-attributes failure to their type<sup>3</sup>.

The update for  $\omega$  is analogous but with  $c_f^\omega$  and  $c_s^\omega$  instead of  $c_f^\theta$  and  $c_s^\theta$  and the bias is present whenever  $c_f^\omega > c_s^\omega$ . That is, the agent over-attributes failure to a low state relative the higher states and under-attributes success to a low state relative to the higher states.

## 4 A Simple Example

The agent can be of one of 3 types:  $\theta \in \{\theta_L, \theta_M, \theta_H\}$  with  $\theta_H > \theta_M > \theta_L$ . They face an unknown exogenous success rate  $\omega \in \{\omega_L, \omega_M, \omega_H\}$  with  $\omega_H > \omega_M > \omega_L$ . Each of the values of  $\omega$  is realized with equal probability. The agent knows the distribution of  $\omega$  but not its realized value. Let the true type be  $\theta^*$  and let the true state be  $\omega^*$ .

---

<sup>3</sup>notice that the values of  $c_s^\theta$  and  $c_f^\theta$  are not restricted to be greater than 1. If they are both equal to each other but less (more) than one, then the bias is simply underinference (overinference).

The agent holds some prior belief about  $\theta$ <sup>4</sup> and chooses a binary gamble  $ein\{e_L, e_M, e_H\}$ . The agent observes whether the gamble is a success or a failure and gets a payoff of 1 and if it is a success. They get 0 otherwise.

The probability of success is increasing in both  $\theta$  and  $\omega$  and is fully described in the following table:

	$\omega_H$	$\omega_M$	$\omega_L$		$\omega_H$	$\omega_M$	$\omega_L$		$\omega_H$	$\omega_M$	$\omega_L$
$e_H$	50	20	2	$e_H$	80	50	5	$e_H$	98	65	25
$e_M$	45	30	7	$e_M$	69	65	30	$e_M$	80	69	35
$e_L$	40	25	20	$e_L$	65	45	40	$e_L$	75	55	45
	$\theta_L$				$\theta_M$				$\theta_H$		

Conditional on a type, the agent's flow payoff is maximized by choosing the gamble that matches the state. For example, if the value of  $\omega$  is  $\omega_H$ , the agent's flow payoff is maximized by choosing  $e_H$  and if the state is  $\omega_L$  the flow payoff is maximized by choosing gamble  $e_L$ . The agent myopically chooses gambles every period to maximize the flow payoff for  $T$  periods.

After observing the outcome of each gamble, the agent updates their beliefs using some algorithm and moves on to the next period.

Notice that both  $\theta$  and  $\omega$  can be identified from the outcomes if enough variation in the effort choices exists. This can be seen by confirming there is no pair of  $\theta$  and  $\omega$  such that the probability of success is the same for all effort choices. Therefore, by changing the effort choice, the agent can learn both their type and the state if they observe enough outcomes.

In this example, for an agent with a dogmatic belief about their type, a self-defeating equilibrium is one in which the agent chooses an effort level that, under the true  $\theta$ , yields a frequency of success that is consistent with the agent's misspecified belief. That is, the observed rate of success under the true type and the effort level chosen,  $P[\text{sucess}|\theta^*, e]$  is the same as the probability of success under the agent's belief,  $P[\text{sucess}|\hat{\theta}, e]$ .

In the data-generating process described above, there are five such equilibria. For example, if the agent is of type  $\theta_M$  but mistakenly believes that he is of type  $\hat{\theta} = \theta_H$  and

---

<sup>4</sup>which is potentially misspecified as in the dogmatic and switcher cases discussed

the and  $\omega^* = \omega_M$ , when the effort chosen is  $e_L$ , the agent will observe a success with 45% chance. Because the agent dogmatically believes that their type is high, they will erroneously conclude that the rate is  $\omega_L$ . Under this belief, the optimal action is  $e_L$  which will continue to generate successes with 45 probability, further reinforcing the incorrect belief. If the agent were to learn the true type and rate, they would instead choose  $e_M$ , yielding a success with 65 probability and strictly higher expected payoffs.

From Heidhues et al. [2018], we know that if  $\theta^*$  and  $\omega^*$  are such that there is a self-confirming equilibrium, a dogmatic agent will always fall into the trap and continue to make sub-optimal choices. On the other hand, a switcher will be able to learn the true type and state as long as their prior is diffused enough and their switching threshold is low enough as in Ba [2023].

By including self-confirming equilibria, the example captures the forces from each of the updating mechanisms discussed in the previous section and allows for direct comparison of all the theories. For realizations of  $(\theta, \omega)$  for which there are self-confirming equilibria, the dogmatic agent will fall into the trap whereas the switcher will be able to escape it. Similarly, an agent with self-attribution bias will update their beliefs differently from an unbiased Bayesian, leading them to choose different gambles. I exploit such cases in order to test which model is a better fit for how subjects behave in a laboratory experiment.

ILLUSTRATE THE DIFFERENT PATHS WITH THE SIMULATION.

## 4.1 The Dogmatic Modeler

A dogmatic modeler updates their beliefs about omega by using bayes rule. For the example, that means he ends up with the following posterior odds ratios:

$$\frac{p_t[\omega_H|\text{outcome}]}{p_t[\omega_M|\text{outcome}]} = \frac{p[\text{outcome}|\omega_H, \hat{\theta}, e]p_{t-1}[\omega_H]}{p[\text{outcome}|\omega_M, \hat{\theta}, e]p_{t-1}[\omega_M]} \quad (6)$$



and

$$\frac{p_t[\omega_M|\text{outcome}]}{p_t[\omega_L|\text{outcome}]} = \frac{p[\text{outcome}|\omega_M, \hat{\theta}, e]p_{t-1}[\omega_M]}{p[\text{outcome}|\omega_L, \hat{\theta}, e]p_{t-1}[\omega_L]} \quad (7)$$

## 5 Experimental Design

I recruited XXX undergraduate subjects from the CESS lab at NYU who participated in an in-person experiment. Sessions lasted approximately XXX hours and subjects earned an average payment of XXX. The experiment was programmed using oTree [?].

The experiment consisted of 2 treatments: the ego-relevant condition and the stereotype condition. Subjects participated in only one of them. Treatments were randomly assigned at the session level. The tasks were identical across treatments, but in the ego condition the probability of success was based on the subject’s own performance in a quiz, while in the stereotype condition, the probability of success was based on the performance of a randomly selected subject from another session.

The experiment had 3 parts. In Part 1 subjects had 2 minutes to answer as many multiple-choice questions as they could from a 20 question quiz in a particular topic. They did this for 6 different topics, one after the other. The topics were: Math, Verbal Reasoning, Pop-culture and Art, Science and Technology, US Geography, and Sports and Video Games. The maximum number of questions they could answer correctly was 20. However, they did not know how many questions were available and they were given no feedback.

After taking all 6 quizzes, they proceeded to part 2 where they were asked to guess their score on each of them. In the stereotype treatment they were additionally asked to guess the score of another participant. All they knew about the other participant was their gender identity and whether they were US nationals or not. For each guess they had three score options: 5 or fewer, between 6 and 15, 16 or more. Each of these correspond to  $\theta_L$ ,  $\theta_M$ , and  $\theta_H$  respectively. They were also asked to say how confident they felt about their choices with 4 options: it was a random guess, there is another equally likely score, I am pretty sure, I

am completely sure. These 4 answers are mapped to priors that place probabilities .33, .50, .75, and 1 to the chosen type. The remaining probability is split equally among the other two types. Questions in part 2 were not incentivized, but subjects were told that providing an accurate answer would increase their chances of earning more money in the last part of the experiment.

The answers of part 2 allow me to classify subjects into overconfident, underconfident and correctly specified. If a subject guesses their score to be in a higher category than their true score, they are overconfident. On the contrary, if they guess their score to be in a lower category than their true score, they are underconfident. Finally, if they guess their score to be in the same category as their true score, they are correctly specified. This classification is done for each of the 6 topics separately.

In part 3 subjects completed a belief updating task for each of the quizzes. Before starting the task they were reminded of their guess for the score. In the ego treatment they were reminded of their guess about themselves and in the stereotype treatment they were reminded of their guess about the other participant. In the stereotype treatment, they were also reminded of the characteristics of the other participant.

For one topic at a time and in random order, they were presented with the three gambles from the example and were asked to choose one of them. The probability of success was determined by their own score in the ego-relevant condition and by the score of the other participant in the stereotype condition. Subjects had access to the three probability tables in the printout on the instructions at all times and the meaning of each cell was explained in detail in the instructions. However, on the screen they had to choose which of the 3 tables they wanted to see before entering their choice in it. This was done as an alternative to a belief elicitation in each round. I take their choice of table to be a noisy measure of their beliefs about the underlying type.

Once they have entered their choice, they observe a sample of 10 outcomes from the gamble they chose. After observing the outcomes, they returned to the choice screen and

entered a new choice. In the choice screen subjects had access to the entire history of gambles and outcomes for that task. Once they entered 11 gambles (and observed 110 outcomes), they moved on to the next topic and repeated the same procedure. They all did so for all 6 topics.

At the end of the experiment, one of the 6 topics was randomly selected to determine the payment. They earned \$.20 for each correct answer in the quiz and for each success in part 3.

The randomness is controlled throughout the experiment and sessions by setting a seed drawn randomly before the first session. By doing this I ensure that any two subjects who have the same type and face the same exogenous rate will observe the same outcomes and thus, if they use the same updating procedure, they should be choosing the same gambles. This design feature allows me to identify differences in updating procedures across subjects.

## 6 Analysis

## Conclusion

Cuimin Ba. Robust misspecified models and paradigm shifts. 2023.

Daniel J. Benjamin. *Errors in probabilistic reasoning and judgment biases*, pages 69–186. 2019. doi: 10.1016/bs.hesbe.2018.11.002.

Paul Heidhues, Botond Köszegi, and Philipp Strack. Unrealistic expectations and misguided learning. *Econometrica*, 86:1159–1214, 2018. ISSN 0012-9682. doi: 10.3982/ecta14084.

Nina Hestermann and Yves Le Yaouanq. Experimentation with self-serving attribution biases. *American Economic Journal: Microeconomics*, 13:198–237, 2021. ISSN 19457685. doi: 10.1257/mic.20180326.