

2020

06

Working Paper

INSTITUTO DE POLÍTICAS Y BIENES PÚBLICOS [IPP]

An experimental test of some economic theories of optimism

Adrián Caballero and Raúl López-Pérez

INSTITUTO DE POLÍTICAS Y BIENES PÚBLICOS – CSIC

Copyright ©2020. Caballero, A., López-Pérez, R. All rights reserved.

Instituto de Políticas y Bienes Públicos
Consejo Superior de Investigaciones Científicas
C/ Albasanz, 26-28
28037 Madrid (España)

Tel: +34 91 6022300

Fax: +34 91 3045710

<http://www.ipp.csic.es>

How to quote or cite this document:

Caballero, Adrián and López-Pérez, R. (2020). An experimental test of some economic theories of optimism. Instituto de Políticas y Bienes Públicos (IPP) CSIC, Working Paper. 2020-06

Available at: digital.csic.es

An experimental test of some economic theories of optimism*

Adrián Caballero[†] and Raúl López-Pérez[‡]

Abstract: Economic theories of optimism provide different rationales for the phenomenon of motivated reasoning, and a recent empirical literature has tested some of them, with mixed results. We contribute to this literature with a novel experimental test of two mechanisms, according to which optimism is respectively predicted when (1) the potential material losses due to the bias are relatively small or (2) the cognitive costs of the bias are small enough. In our design, these two accounts predict inflated expectations regarding some future payoff. Contrary to that, the average subject tends to (slightly) underestimate that financial prospect. Although a minority of the subjects overestimate systematically, the size of their errors is rather reduced, and they hardly differ in their personal characteristics from the rest of the subjects. In fact, optimism in our experiment is correlated with the sample observed, in that it is more likely when a subject observes relatively few good signals. This is again at odds with (1) and (2). These mechanisms, we conclude, do not appear to fully capture under which circumstances people fail into a positivity bias. Yet (1) seems to be empirically less relevant, in that we observe a similarly limited level of bias irrespectively of its monetary cost.

Keywords: Belief Updating, Biases, Motivated Beliefs, Optimism, Wishful Thinking.

JEL Classification: C91, D03, D80, D83, D84.

* We are grateful to Kai Barron, Alexander Coutts, Guy Mayraz, Jonathan Parker, Eli Spiegelman, and participants at seminars in the IPP-CSIC for helpful comments and suggestions. Of course, all remaining errors are ours. We also gratefully acknowledge financial support from the Spanish Ministry of Economics, Industry and Competitiveness through the research project ECO2017-82449-P, and helpful research assistance by Fabio Casalegno, Roberto Lucas, Sergio Rubio, and Sara Yuste.

[†] Department of Economic Analysis, Universidad Autónoma de Madrid, Cantoblanco, 28049 Madrid, Spain. e-mail address: adrian.caballero@uam.es

[‡] Institute of Public Goods and Policies (IPP), Spanish National Research Council (CSIC), C/ Albasanz, 26–28, 28037, Madrid, Spain. E-mail address: raul.lopez@csic.es

1. Introduction

Numerous studies show that people *sometimes* have ‘too’ optimistic beliefs about self-relevant events and future material outcomes, as if their beliefs were influenced and aligned with their desires or preferences –for reviews, see Bénabou and Tirole (2016); Epley and Gilovich (2016); Kunda (1990); Wicklund and Brehm (1976). *When such a positivity bias is observed*, further, there is strong evidence that it is caused by asymmetric updating, that is, the under-weighting of undesirable information relative to the desirable one –e.g., Eil and Rao (2011); Möbius et al. (2011); Sharot et al. (2011). Wiswall and Zafar (2015), for instance, report that undergrads update their beliefs about their own future earnings asymmetrically: After learning the actual average earnings for each major, their original beliefs about self-earnings are slightly revised downwards when their prior estimation of average earnings was too positive, whereas the corresponding upward revision is more substantial if they underestimated the actual mean value. On the other hand, it is yet far from clear which environmental and individual factors make people (more) skewed: People fall into positivity biases, but it is not well-known *when* they do so, and *who* is more likely to do so.¹

Understanding which factors or conditions are more propitious for the formation of optimistic beliefs is crucial for at least two reasons. *First*, these beliefs can motivate suboptimal decisions which sometimes may lead to undesirable collective outcomes, as the next examples illustrate. **I**: If people want to believe that they are often right in their presumptions and hence commit few mistakes, a confirmation bias follows. That is, agents underweight disconfirming evidence and overweight confirming evidence, e.g., Eil and Rao (2011). As a result, voters may develop partisan biases and polarized beliefs on issues such as climate change (Sunstein et al., 2016), as well as a higher credulity for fake news in line with their prejudices. **II**: Optimism may lead to overinvestment during economic booms and boost the formation of financial bubbles (Aliber and Kindleberger, 2015; Shiller, 2000). As an illustration, a few years before the subprime mortgage crisis, a survey conducted by Case and Shiller (2003) to new homeowners in different cities in the US showed that 90 percent of them believed that housing prices in their cities would keep increasing for the next 10 years at an average estimated rate between 9 and 15%, depending on the city. **III**: Oster et al. (2013) study testing among individuals at risk for Huntington disease –a degenerative neurological disorder associated to a genetic alteration. One of their main findings is that those who reject testing generally underestimate their actual risk and behave as if they do not have the disease, even when diagnosed individuals behave differently (for example, they are more likely to retire, make major financial changes or change their recreation habits). More speculatively, but also in the public health realm,

¹ A similar point has been stressed by Benjamin (2019) in his comprehensive review about biases in beliefs.

many government's initial reaction (or lack of it) to the Coronavirus threat in January-February 2020 suggest some optimism –e.g., the virus will not arrive, and hence there is no need to buy in advance enough personal protective equipment for the healthcare workers in the public hospitals in case it comes.

Second, there is also a positive side of optimism. Beliefs about the future, the morality of our acts, or about oneself can trigger different emotions like hope, pride, anxiety, guilt, or shame. In this sense, individuals may prefer information and beliefs that contribute to emotional well-being. In fact, optimism has been associated to a good mental and physical health (Rasmussen et al., 2009; Strunk et al., 2006). Also, optimistic beliefs about one's abilities can motivate the individual to undertake difficult tasks and to overcome the obstacles that can arise (Bénabou and Tirole, 2002, 2004). Further, self-confidence can also help to convince others (von Hippel and Trivers, 2011). To sum up, optimism has arguably cons and pros and, depending on the context, we may wish to deter or promote such bias. That requires, however, a precise understanding of the environmental, individual, and institutional conditions leading to more optimism (or more realism).

In this endeavor to find answers for the *when* and *who* questions cited above, models are invaluable tools, as they offer insights, allow precise policy design, and organize the analysis. Several economic theories of optimism have been proposed, and they give different answers to those questions. Our goal here is to test the predictions of two families of theories by experimental means. A common hypothesis in all these models is that individuals have implicit preferences over the possible states of nature and derive utility (disutility) from thinking that their preferred state is (not) true. If people experience anticipatory feelings like excitement, joy, fear or anxiety from thinking on some future uncertain events, for instance, they would rather believe that the future will be favorable, so as to trigger the relatively more positive emotions. Another idea common to these models is that human inference operates *as if* people chose their beliefs, although most likely through a subconscious mechanism. This means that belief acquisition responds to incentives and constraints, as in a usual economic decision like, e.g., buying some good. The families differ mostly in the specific factors that restrain optimism and prevent individuals' beliefs from departing too much from reality. In models like Akerlof and Dickens (1982) and Brunnermeier and Parker (2005) –AD and BP henceforth, respectively–, the 'demand' for a biased belief depends on its material price: If biases lead to sufficiently costly or risky decisions, individuals may be more reflective, cooling down their expectations. In Rabin (1994) and Bracha and Brown (2012), alternatively, self-deception requires the individual to selectively avoid or look for information, to rationalize it or to restrain certain thoughts. To formalize this idea, these models explicitly incorporate cognitive costs associated to belief distortion, which increase the further away from Bayes' rule the beliefs are.

While there is plenty of evidence of situations in which individuals seem to be optimistic, the specific implications of these models have received much less attention. As Coutts (2019b, p. 549) notes, “rigorous tests of existing theory and direct evidence about optimism are scarce”, and the evidence collected so far is not conclusive (see Section 2 below for a review). This paper hence contributes to this literature with a lab experiment. On the adequacy of this methodology, we note that, while field studies offer extremely suggestive evidence from a naturalistic setting, they can rarely offer full control of the agents’ relevant priors and evidence observed. As a result, it is difficult to establish if, say, an optimistic prediction is the result of some bias in the updating procedure or caused by the very positive evidence that the person has received or her skewed priors. More generally, field studies cannot be used to test many of the predictions of the existing theories of motivated inference. Lab studies, in contrast, permit such fine-grained tests.

Any participant in our experiment faces an urn with 100 balls. Each ball has either a boy or a girl name written on it, and the number $F \in [0, 100]$ of ‘female’ balls in the urn is known to be randomly determined for each subject. Further, subjects know that they can get a *state prize*, that is, 0.50 euros per female ball in their own urn, and hence have an incentive to believe that F is as high as possible (observe well that this state prize does not depend on the subject’s choices, as F is randomly determined). Thirty balls are then consecutively drawn out with replacement, and the subject observes the name in each one. After draws 15, 22 and 30, furthermore, the subject is unexpectedly asked to estimate F (estimation rounds 1, 2, and 3 henceforth). While the first two estimations are not incentivized, the subject can get an *estimation prize* of 10 euros if her last estimation is sufficiently accurate –to prevent hedging, in fact, subjects could get either the state or the estimation prize, randomly determined with probability 0.5 at the end of the experiment. After round 3, furthermore, the subject must provide the shortest credible interval for F that contains 95% of the probability mass (this was not incentivized).

From a Bayesian perspective, the optimal estimate of F involves a trivial problem of extrapolation from the observed proportion of ‘female’ balls (see Section 4.1 for a detailed explanation). In contrast, the models by AD and BP predict that (some) individuals will overestimate F , so as to correspondingly ‘inflate’ the state prize. This is particularly true when there are no potential losses associated to inaccurate beliefs, as in the first two estimation rounds, where individuals get no prize for accuracy. Since subjects in our experiment can only get either the state or the estimation prize, further, AD and BP predict a correlation between risk aversion and Bayesianism: very risk averse types would rather have realistic beliefs, so as to maximize the likelihood that they get something in case the estimation prize is selected for payment.

Overall, we find scarce evidence supporting these models, and particularly for models like AD and BP. Specifically for these two models, *first*, we do not observe systematic optimism, i.e., overestimation, in any of the three rounds. Indeed, the average and median subject slightly underestimates F in every round. *Second*, the size of the estimation bias is not reduced by the introduction of incentives for accuracy in the third round. While there are subjects who never underestimate F , *third*, they account to just 26.47 percent of the sample ($N = 68$) and inflate F to a rather limited extent. *Fourth*, optimism is unrelated to any of the individual characteristics that we record, except a relatively higher CRT score. That is, the ‘optimistic’ subjects look scarcely different from the others, particularly in terms of risk aversion. However, *fifth*, we find that overestimation in our experiment is extremely correlated with the sample observed: Optimists (pessimists) tend to be subjects who observe relatively few (many) female extractions. This is again hardly consistent with models like AD and BP. In what regards the confidence intervals, *sixth*, the model by BP says that optimistic subjects will report ‘positively skewed’ intervals, i.e., the subject’s estimate of F is the lower limit of the interval. No evidence supports this prediction, though. Note that models with cognitive costs like Rabin (1994) are also inconsistent with the first, fourth, and fifth findings just cited. They are in line though with the second finding and *perhaps*, conditional on the cognitive cost function assumed, with the limited extent of the biases observed in our experiment, i.e., our third finding (we have not analyzed what these models predict for the confidence intervals). Overall, therefore, this second family appears to fit better with our results, at least in relative terms.²

The rest of the paper is organized as follows. The next section discusses some related experimental literature and our contributions to it. Section 3 introduces the experimental design. Section 4 starts by presenting the predictions of the Bayesian model, as well as applying AD and BP to our setting. This section also reports experimental results afterwards, and discusses the models with cognitive costs. Section 5 concludes by mentioning potential future venues of research.

2. Literature review

To organize this survey, we mention in what follows several predictions of the theories and how they compare with the existing experimental evidence so far.³ The first prediction refers to the very phenomenon of optimism and hence is common to all of the models considered here. Note yet that the models predict this phenomenon under different conditions, to be specified later. In addition, some models like Möbius et al. (2014) explicitly analyze how optimists infer, explaining belief

² The model in Mayraz (2013) assumes no costs to belief distortion and is hence consistent as well with our second finding, but not with our other findings (leaving aside the sixth one, which we have not checked).

³ This review cannot make justice to the whole literature in this respect, but consult Caballero and López-Pérez (2020a) for a fuller review of the literature on motivated inference and optimism. See also Caballero and López-Pérez (2020b) for a review of the existing literature on the relation between selective recall and optimism.

inflation as the result of asymmetric updating. As explained in the introduction, the idea is that signal observations are over-weighted or under-weighted depending on whether they support or contradict, respectively, the decider's desired beliefs.

Prediction 1: Deciders inflate their beliefs. Specifically, the difference between the subjective and the objective probabilities is correlated with the utility payoff that deciders get from having the desired beliefs about the state space. Inflation occurs because individuals process 'good' and 'bad' news asymmetrically, thus reinforcing their favorite beliefs.

Evidence: Several studies report inflated beliefs and find evidence of asymmetric updating as a potential cause, but not all studies do so. In line with our discussion in the introduction, our interpretation is that inflation requires propitious conditions, still not well understood. Two groups of experimental studies can be perhaps distinguished for the sake of the exposition. In a first one, the beliefs analyzed are arguably relevant for self-esteem. When eliciting a subject's posterior distribution about her rank in a group according to some ego-relevant trait (specifically, physical attractiveness and/or IQ score), for instance, Eil and Rao (2011) and Möbius et al. (2014) find evidence of positively skewed updating (see also Heger and Papageorge, 2018). Specifically, people seem to update beliefs according to Bayes' Rule when the signal is good or desirable, and under-update when the signal is negative. In contrast, Ertac (2011) reports negatively skewed updating, while Buser et al. (2018) find no evidence at the aggregate level for asymmetric updating about relative performance. See also Zimmermann (2020), who finds evidence of underweighting of the negative signals when the posteriors are elicited one month after feedback, but not when they are elicited immediately after. Grossman and Owens (2012), in turn, explore learning about *absolute* performance and find no evidence of asymmetric updating.

In a second group of economic experiments, closer to our study, subjects have a financial stake in some specific event E and must report the posterior that E occurs after observing some relevant evidence –Gotthard-Real (2017), the Baseline condition in Barron (2020), Coutts (2019a); and Heger and Papageorge (2018).⁴ Little evidence of a positivity bias has been found in the studies just cited, in spite of the fact that they display several differences in what regards the priors on E , the prize if E occurs, or the randomization mechanism, e.g., mechanical or using a computer program. In addition, Coutts (2019a) studies inference in a “value relevant” treatment, i.e., when subjects have a preference for some event e to be true, and a “neutral” one, and finds no differences in belief updating across treatments. This occurs when e is financially relevant, i.e., the E above, but also when it is ego-relevant, i.e., the ranking in a math or verbal quiz.

⁴ Some recent studies do not fit exactly within any of the two groups considered. Engelmann et al. (2019), for instance, report that subjects under-estimate the probability of receiving an electric shock that is outside of their control

To our knowledge, there are just two lab studies in Economics reporting a motivated bias when subjects have a financial interest for some state. Both involve between-subjects designs, so the “bias” comes from the fact that one role is more positive than another one, given similar information but different preferences. In Mayraz (2013), subjects are shown a chart of historical wheat prices and have to predict afterwards the price at some future time point, getting a bonus for accuracy. In addition, subjects get a payoff that increases (decreases) with that price if her randomly-selected role is “Farmer” (“Baker”). In average, farmers make significantly higher predictions than bakers, consistent with a positivity bias. In the Strategic condition of Charness and Dave (2017), in turn, subjects play a 2x2 game. The payoff matrix is a priori uncertain, as there are two possible payoff constellations. While subjects in the Odd role get the same equilibrium payoff in both matrices, those in the Even role have a preference for one of them. Prior to playing the game, participants observe a sequence of six signals and their incentivized posteriors of each state/matrix reveal that Even players underweight more strongly the negative signals, i.e., those confirming the ‘worst’ matrix.

Neuroscientists and psychologists have also gathered some supportive evidence for inflation and asymmetric updating in beliefs about future outcomes. In Sharot et al. (2011), participants are sequentially presented a total of 80 adverse events, such as being diagnosed with Alzheimer’s disease or suffering a car accident, and have 6 seconds to estimate their chances of facing any such event in the future (without incentives). In a second stage, subjects are shown for 2 seconds the actual frequency with which any such event happens among individuals living in the same socio-cultural environment as them, and must guess their posteriors of encountering that event. Sharot et al. (2011) report evidence for asymmetric updating in favor of good news. Using a similar design, Ma et al. (2016) find that intra-nasally administered oxytocin promotes optimism and asymmetric updating. This is particularly true in individuals with high depression or anxiety traits, who under-weight undesirable feedback more pronouncedly than similar individuals in a placebo treatment –on how depressed individuals update beliefs, possibly in a relatively more balanced manner, see also Alloy and Abramson (1979) and Garrett et al. (2014). ■

The next prediction follows from models like BP. The idea is that people will be *less* biased when there is more risk. Taking into account that having inaccurate beliefs often leads to suboptimal choices, that is, the models propose that belief inflation will be attenuated when acting as an optimist leads to large expected losses (relative to a Bayesian). Intuitively, people think more when there is a lot at stake, and hence their beliefs are dominated relatively less by their “animal spirits”.

Prediction 2: The correlation between beliefs and preferences will get weaker as the expected material loss for holding inaccurate beliefs increases.

Evidence: several of the papers cited above have systematically analyzed whether the size of the expected loss reduces the degree of inflation. Coutts (2019a,b) runs sessions with different accuracy payments, i.e., low (\$3), medium (\$10), or high (\$20). In addition, participants can either get a nil or high prize (\$80) if some target event E occurs. According to Prediction 2, subjects have no incentive to distort their beliefs about the probability of E in case the prize is \$0, provided that they get no other utility from the occurrence of E , e.g., if E is not ego-relevant. In contrast, distortion should be maximal if the prize is high and the accuracy payment low, i.e., \$3. In spite of this, the author concludes that neither prizes nor accuracy payments alter updating. Mayraz (2013) varies the size of the accuracy bonus from £1 to £5 and reports that the magnitude of the bias does not depend on the scale of the bonus (neither on the size of the prize associated to the desirable event). Similarly, Ertac (2011) studies the effect of rewarding accurate beliefs and finds no significant difference across compensated and non-compensated sessions in terms of the distribution of priors and the absolute value of the bias (Mann–Whitney test, $p = 0.89$ and $p = 0.79$, respectively). See also Engelmann et al. (2019) for similar negative results regarding the accuracy payment. If people suffer from a positivity bias, in summary, the cost of such bias does not seem to set limits on its size, at least with the parameterizations that have been considered so far. ■

In some models, belief distortion is assumed to be cognitively challenging and hence involving an explicit cost —Bracha and Brown (2012), Rabin (1994). For instance, Rabin (1994, p. 180) contends that “developing beliefs that differ from this level [of natural, intellectually honest beliefs] is costly because it may intrinsically conflict with other parts of a person’s belief system, and reintegrating it can involve laborious intellectual activity.”

Prediction 3: Less inflation when it is cognitively costly.

Evidence: To our knowledge, the experimental literature has not dealt thoroughly with this question yet. Coutts (2019b) compares and tests some predictions of BP’s model of optimal expectations and the model of affective decision making proposed by Bracha and Brown (2012), which includes explicit mental costs from belief distortion. While Coutts (2019b) finds limited evidence supporting some of the implications of the model by Bracha and Brown (2012), its work focuses on the effects of state-dependent prizes and accuracy prizes on optimistic bias, and not directly on the cognitive aspects of belief distortion.

3. Experimental design

Any subject faces her own virtual urn, with 100 balls inside. Each ball in the urn has either a boy or a girl Spanish name, and the 100 names in the urn are different. Balls with a girl/boy name are called henceforth female/male balls —these terms were not used in the subjects’ instructions; see

Appendix I. The precise rate θ of female balls is a multiple of 0.01 selected by the computer with uniform probability over the interval $[0, 1]$ at the start of the session; the rate of male balls is hence $1 - \theta$. It follows that the number of female balls F is equal to $100 \cdot \theta$; we will make reference generally to θ for consistency, although the instructions were expressed in terms of F . Although the subject does not know θ , the method to determine it is known in advance.⁵ Priors are hence arguably fixed. Each subject then observes the realization, i.e., name, of an a priori undetermined number (in fact, 30) of consecutive random draws with replacement from her/his box. Subjects did not observe others' samples. After the first 15, 22 and 30 extractions, further, the subject is asked to provide a point *estimation* of θ –therefore, she gives estimates in 3 rounds, each one with a progressively enlarged dataset. Subjects were explained each estimation task only immediately after observing the corresponding extractions and did not receive any feedback about prior extractions.

Subjects get either a 'state prize' that depends on the rate/state θ or an 'estimation prize' depending on the accuracy of the participant's last estimation of θ . The prize that a subject finally gets is randomly determined with probability 0.5 at the end of the experiment. As a 'state prize', specifically, the subject gets 0.50 euros for each female ball in the urn, e.g., a maximum of 50 euros if $\theta = 1$. For the 'estimation prize', in turn, let $\hat{\theta} \in [0, 1]$ denote a subject's last, i.e., third, estimation. The subject earns 10 euros if the corresponding error $|\theta - \hat{\theta}|$ is smaller or equal to 0.02, and 0 euros otherwise. The elicitation of the first two estimations of θ , in turn, is not incentivized. Participants are informed about the nature of the 'state prize' before they observe any extractions, whereas the structure of the 'estimation prize' is only revealed just before the last estimation task, i.e., after the 30 extractions. Indeed, the initial instructions only stated that with probability 0.5 they will get either the 'state prize' or an undefined prize whose nature will be specified later (this design choice is irrelevant to test the theories considered here, but relevant for the analysis in Caballero and López-Pérez (2020b)).

Additional tasks and questions are inserted *between* some extractions. After the first 7 extractions, specifically, we included a brief questionnaire where we gathered information on personal and socio-demographic characteristics (age, gender, major, religiosity, and political ideology). A risk aversion index was elicited after the first 19 extractions.⁶ Also, subjects completed an expanded cognitive reflection test or CRT (Frederick, 2005), including the three classical

⁵ To determine the specific names in each urn, we used two lists with the most popular, non-compound female and male names in Spain, respectively. The lists, elaborated by the Spanish National Statistics Institute, order the names according to frequency; see <https://www.ine.es/en/welcome.shtml>. Once θ had been randomly determined for a subject, therefore, we randomly selected $100 \cdot \theta$ girl names and $100 \cdot (1 - \theta)$ boy names in the corresponding lists to 'fill' the urn.

⁶ Subjects faced the choice between lottery A with prizes 2 and 1.5 Euros and lottery B with prizes 4 and 0 Euros, with equal probabilities of the larger and lower prize across lotteries. Letting P denote the probability of the larger prize, they had to indicate the threshold value of P such that they always preferred B to A, on a scale from 0 to 100.

questions and two additional ones, after the first 26 extractions. Furthermore, after the third estimation task, i.e., the incentivized one, subjects had to report the shortest 95% confidence interval they could figure out. In other words, they indicated a lower and an upper bound for θ , such that they believed that the correct θ was ‘almost surely’ in the interval determined by those limits. Confidence intervals were not incentivized; as we discuss later, however, our results do not differ much from those in López-Pérez et al. (2020), where subjects were paid for accuracy. After this interval estimation, additionally, we included an incentivized ‘recall task’ and two questions so as to check whether they expected to recall better female than male extractions, i.e., good than bad news; this data is irrelevant for the test of the theories considered here, but see Caballero and López-Pérez (2020b) for a full description and analysis. Subjects responded, in addition, two questions on statistical knowledge, the LOT-R test on optimism (Scheier and Carver, 1985; Scheier et al., 1994), and a test on disappointment, in this order, thus ending the experiment.

The study consisted of six computerized sessions at Universidad Autónoma de Madrid, with a total of 68 participants. The software used was z-Tree (Fischbacher, 2007). Participants were not students of the experimenters. After being seated at a visually isolated computer terminal, each participant received written instructions that described the decision problem (translated to English in Appendix I). Subjects could read the instructions at their own pace and we answered their questions in private. Understanding of the rules was checked with a computerized control questionnaire that all subjects had to answer correctly before they could start making choices (see the screenshot in Appendix I). At the end of the experiment, subjects were informed of their final payoff and paid in private. Each session lasted approximately 60 minutes, including paying subjects individually, and on average subjects earned 20.50 euros, including a show-up fee of 3 euros.

Discussion

While appropriate for the test of the theories considered in this paper, the elicitation of a subject’s $\hat{\theta}$, i.e., the mode of her posterior beliefs, is a rather unusual feature in the literature on belief updating, specifically on motivated inference, where the subject’s posterior probability distribution is often elicited instead –e.g., using the lottery method as in Coutts (2019a), or the crossover method in Möbius et al. (2014). In a sense, we elicit an ordinal instead of a cardinal measure of probability. We introduced this relatively novel aspect for three reasons. First of all, we found the question of how people compute empirical frequencies when they have a preference for some states/values an interesting one in itself. Second, incentive compatible elicitation procedures are often complex to explain to subjects, e.g., Schlag et al. (2015). In contrast, our estimation prize is rather straightforward. Third, we suspected that the computation of the exact probability of any rate was a substantially more demanding problem than the estimation of $\hat{\theta}$, which requires only extrapolating

from the sample (see 4.1). As noted by Tversky and Kahneman (1983) and Gigerenzer and Hoffrage (1995), posing problems in frequentist (as opposed to probabilistic) terms may mitigate some errors. In summary, our design attempted to reduce any potential noise due to the subjects' misunderstanding of the elicitation procedures or the statistical nature of the problem. The drawback is that we lose rich information on their posteriors, although the confidence interval estimation offers some insights. Note also that, since we do not elicit the precise posteriors, it is not our research goal to analyze whether people update their probabilistic beliefs in a conservative manner, or display asymmetric updating. Yet these are issues that have received attention before, as we have explained in Section 2.

For another remark, note that subjects are never paid for both their beliefs and the actual state θ . Otherwise we might face hedging problems and hence the subjects' potential misreporting of their beliefs in the incentivized elicitation (Blanco et al., 2010). To understand this, suppose for the sake of the exposition that a subject can get both prizes and believes that $\theta = 0.9$ with probability $p > \frac{1}{2}$ and $\theta = 0.4$ with probability $1 - p$. If she reports $\hat{\theta} = 0.9$, therefore, she expects with probability p a state prize of $50 \cdot 0.9 = 45$ Euros plus an estimation prize of 10 Euros, and a payoff of $50 \cdot 0.4 = 20$ Euros with probability $1 - p$. A report of $\hat{\theta} = 0.4$, on the other hand, generates a lottery with payoffs of 45 and $20 + 10$ with respective probabilities p and $1 - p$. It follows that a sufficiently risk averse subject would rather report $\hat{\theta} = 0.4$, so as reduce variability. More generally, the Bayesian prediction would be conditional on the subject's degree of risk aversion if there were hedging problems; our design prevents this kind of complexities.⁷ In any case, we note incidentally that Ertac (2011) does not find strong evidence in support for hedging in her study.

Finally, the beliefs were incentivized only in one round, i.e., the last one. The rationale under this design choice is multiple. On the one hand, the models of optimism described in Section 4 say that subjects will inflate more when the 'price' of inflation, i.e., the potential monetary loss for being inaccurate, is nil (see Proposition III below). The models therefore predict more overestimation in the non-incentivized rounds. A potential objection against not incentivizing some rounds is that subjects could give little thought to the issue. We note however that these are exactly the type of situations in which BP intuitively predict more optimism and, more substantially, we can compare our results across rounds and hence check whether incentives reduce noise (anticipating somehow our results, the evidence is negative in that respect). The reason to incentivize precisely the *last* round, further, was to avoid ambiguous predictions by the models of optimism. If subjects were paid instead in the first round only, for example, in posterior rounds they would simultaneously wish to

⁷ Subjects also get a payoff in the recall task, but this is introduced after the estimation task.

believe that (i) the share of female balls is high, but also that (ii) their estimate in the first round was accurate. In these circumstances, it would not be clear whether, according to models of optimism, we should expect optimistic estimates or some degree of anchoring relative to the first estimation. The incentive structure of our design, in contrast, tries to guarantee that only incentive (i) is present in the first two rounds, while both (i) and (iii) a desire for accuracy are relevant in the last round.

4. A test of several theories

In this section, we first introduce in 4.1 the Bayesian model. Afterwards, we apply AD and BP to our experiment and derive a series of predictions, which guide the posterior analysis of the experimental evidence. The discussion concerning models with cognitive costs comes at the end of this section.

4.1 Models and predictions

We start by introducing some general notation, together with the standard Bayesian theory.

General setup & the Bayesian model

An expected payoff-maximizer called Eve must estimate the frequency/rate $\theta \in [0, 1]$ with which some phenomenon f occurs. Specifically, there is an i.i.d. signal S , taking on value $v \in \{f, m\}$, and such that probability $(S = f) = \theta$ —for expositional purposes, we sometimes refer to f as female, and m as male. Eve does not know the exact value of θ . Let $\Theta \subseteq [0, 1]$ denote the space of potential values of θ (for expositional convenience, we assume that Θ is finite). Eve has prior beliefs over Θ , quantified by a finitely additive probability measure. Let p_k denote Eve’s priors about rate $\theta_k \in \Theta$. In our experiment, $\Theta = \{0, 0.01, \dots, 1\}$, whereas the (uniform) prior of any rate θ_k is $p_k = 1/101$.

Eve has observed some realizations of S and hence can use that evidence to update her priors. The number of female observations is denoted as f and that of male ones as m . Given *data* $D = (f, m)$, Eve’s posterior beliefs about any $\theta_k \in \Theta$ are obtained by means of Bayes’ rule (the last equality is true only if priors are uniform):

$$p_{k|D} = \frac{p_k \cdot \theta_k^f \cdot (1-\theta_k)^m}{\sum_{\theta_i \in \Theta} p_k \cdot \theta_i^f \cdot (1-\theta_i)^m} = \frac{\theta_k^f \cdot (1-\theta_k)^m}{\sum_{\theta_i \in \Theta} \theta_i^f \cdot (1-\theta_i)^m} \quad (1)$$

If Eve were a subject in our experiment, she would face a rather simple problem of inference. Let $f \in [0, 1]$ denote the (rounded) frequency of female balls in the sample observed by Eve, i.e., $f = \frac{f}{m+f}$. Since priors are uniform in our experiment, it follows from a standard Bayesian argument that Eve’s posterior beliefs have a unique mode at $\theta_k = f$ and a concave shape. Given the structure of the estimation prize in the third round, therefore, Eve reports there an estimate $\hat{\theta} = f$ —except when the sample observed is ‘extreme’, i.e., contains 0, 1, 29, or 30 female balls; in these cases, she reports an

estimation slightly different than f , a point that we take into account in our analysis below.⁸ When the point estimations are not incentivized, finally, the argument is analogous except that no distortion is here expected for any value of f . Our first result is hence direct.

Proposition I (Bayesian): In each estimation round, a Bayesian subject who has observed so far a sample where the (rounded) share of female balls equals $f \in [0, 1]$ chooses f as an estimate of θ . The only exception appears in the last estimation round, where some slight distortion is predicted if the sample observed contains extremely few or extremely many female balls. In average, point estimations do not significantly differ from the average empirical frequency.

Choice of beliefs: Applying AD and BP to our experiment

We maintain the notation introduced in the general setup, and consider an agent called Abel, identical in all respects to Eve except Bayesian updating, i.e., equation (1) above. If Abel participates in our experiment, specifically, any estimation round is conceived as divided in two periods; in period 1, Abel chooses his subjective beliefs about θ ,⁹ while he opts for the corresponding estimate $\hat{\theta} \in \Theta$ in period 2, based on those beliefs. Let \hat{p}_k denote the subjective probability of state θ_k as chosen in period 1. Abel can choose any set of subjective probabilities, provided that they satisfy Kolmogorov's probability axioms.

Abel's problem in any round can be solved recursively. In period 2, he chooses the estimate that maximizes his expected monetary payoff, based on his subjective beliefs. If we abstract for simplicity from the payoff in the recall task, Abel's payoff equals either the state prize or the estimation prize; both prizes have equal probability. The state prize, recall, is proportional to the share of female balls, i.e., equal to $M \cdot \theta$, where $M = 50$ in our experiment. With respect to the estimation prize, we simplify matters by assuming that it amounts to zero unless estimation $\hat{\theta}$ exactly matches the actual state of the world (i.e. $\hat{\theta} = \theta$), in which case it equals \bar{v} (implicitly, $\bar{v} = 10$ in the third, incentivized round, but $\bar{v} = 0$ in the first two rounds). Further, let $u(x)$ denote the utility

⁸ To clarify, think of the case in which the sample contains 30 female balls. The most likely value of θ is 1. If Eve reports an estimate of 1, however, she would eventually earn prize 2 only if the true rate is 0.98, 0.99 or 1. On the other hand, if her estimate is 0.98, she earns prize 2 if the true rate is between 0.96 and 1, both included. A further subtlety is that the distribution of posteriors is not symmetric in general, and particularly for the samples considered here. If Eve observes 1 female ball, specifically, the mode is 0.03, but interval $[0.01, 0.05]$ has less aggregate probability than $[0.02, 0.06]$. In this case, therefore, Eve should report $\hat{\theta} = 0.04$. A similar argument applies when there are 29 female balls in the sample. There are no other cases where a rational Bayesian should report an estimate different than the mode. These "distortions" could be prevented if the estimation prize required an absolutely correct estimate of θ . Since this could reduce a subject's incentive to exert attention on this task, however, we tried to achieve an equilibrium. Note also that the optimal estimation of θ depends on the structure of the estimation prize. A different set of incentives could imply that the optimal point estimate is a different statistic than the mode, like the mean, the median, etc.

⁹ To reduce degrees of freedom and for simplicity, we posit that subjects in our experiment trust the experimenter's instructions although, formally speaking, the models here allow subjects to choose their beliefs in this respect as well –e.g., believing that the probabilities of earning either the state or the estimation prize are not the same.

function of money, where x indicates the monetary gain; we posit $u(0) = 0$. In period 2, to sum up, Abel chooses $\hat{\theta}$ so as to maximize expected utility function

$$\frac{1}{2} \sum_{\theta_k \in \Theta} u(M \cdot \theta_k) \cdot \hat{p}_k + \frac{1}{2} u(\bar{v}) \cdot \hat{p}_k(\theta_k = \hat{\theta}) \quad (2)$$

It comes straightforward that (2) is maximized by choosing the $\hat{\theta}$ that matches the most likely state of nature according to the subjective probabilities, i.e. the subjective mode θ_{sm} with subjective probability \hat{p}_{sm} such that $\hat{p}_{sm} \geq \hat{p}_k$ for any $\theta_k \in \Theta$. If the posterior subjective distribution has several modes, Abel is indifferent between them; in this case, θ_{sm} denotes the mode chosen in period 2. Let p_{sm} denote the objective probability of θ_{sm} –note well that p_{sm} does not represent the objective probability of the objective mode θ_{om} , since θ_{sm} is not necessarily equal to θ_{om} . The corresponding probabilities \hat{p}_{om} and p_{om} are analogously defined for the objective mode, which in our problem is unique for any data D observed (see Proposition I).

In period 1, optimal beliefs are chosen. Following AD and BP, we assume that Abel may experience some anticipatory utility at the end of period 1. That is, he gets utility from thinking about his future material payoff. This anticipation utility depends on the specific parameters described above and his current beliefs, i.e. the subjective probabilities. In period 1, that is, Abel chooses the beliefs that maximize

$$a \frac{1}{2} [\sum_{\Theta} u(M\theta_k) \cdot \hat{p}_k + u(\bar{v})\hat{p}_{sm}] + \frac{1}{2} [\sum_{\Theta} u(M\theta_k) \cdot p_k + u(\bar{v})p_{sm}] = \frac{1}{2} \sum_{\Theta} u(M\theta_k)(a\hat{p}_k + p_k) + \frac{1}{2} u(\bar{v})[a\hat{p}_{sm} + p_{sm}] \quad (3)$$

where $a \geq 0$ reflects the intensity of the anticipatory utility. Note that expression (3) has four components. Two of them refer to the anticipatory utility from the state and estimation prizes, i.e., $\sum_{\Theta} u(M\theta_k) \cdot \hat{p}_k$ and $u(\bar{v})\hat{p}_{sm}$, respectively; they depend on the beliefs chosen in Period 1. Intuitively, the first component is maximized when beliefs put the whole probability mass on $\theta_k = 1$, while the second one is maximized when a single rate receives all the probability mass; it follows that anticipatory utility is maximized when rate $\theta_k = 1$ is certain, i.e., $\hat{p}_k = 1$ for $\theta_k = 1$. In turn, a third component of (3) is the objective expectation of the state prize, $\sum_{\Theta} u(M\theta_k) \cdot p_k$, which cannot be altered by Abel. Finally, the objective expectation of the estimation prize, $u(\bar{v})p_{sm}$, depends on Abel's estimate, which in turn depends on his beliefs. The size of this component decreases as θ_{sm} moves further from θ_{om} , as this reduces p_{sm} . Several implications follow from this optimization problem. The proofs can be consulted in Appendix III.

Proposition II. If $a = 0$, then any subjective belief such that θ_{sm} coincides with the objective mode θ_{om} is optimal.

Proposition II implies that, in the absence of anticipation utility, only beliefs that do not alter the optimal action in the second period are optimal. Note the difference with Eve's case: Since Abel can choose his beliefs, the optimal ones when $a = 0$ are indeterminate, except that the mode of Eve's and Abel's must coincide.

Proposition III. If $\bar{v} = 0$, then the optimal beliefs are characterized by $\hat{p}_k = 1$ for $\theta_k = 1$ and $\hat{p}_k = 0$ for any $\theta_k < 1$.

This simply states that, if there is no potential loss in keeping distorted beliefs, then it is optimal for Abel to believe that the only possible state is the most favorable. This prediction is relevant for the first two estimation rounds, where there were no incentives for accuracy. The following results offer insights on the third, incentivized round.

Proposition IV. The optimal subjective mode θ_{sm}^* is at least equal to the objective mode. If $a > 0$, further, optimal beliefs are characterized by $\hat{p}_k = 0$ for any $\theta_k < \theta_{sm}^*$.

In our experiment, Proposition IV implies that Abel will never underestimate the number of female balls in the urn and that posteriors will be extremely skewed about its mode, with clear implications on Abel's (subjective) 95% confidence intervals. The rationale for this prediction is quite intuitive. On one hand, the assignment of non-nil probability $\hat{p}_k < \hat{p}_{sm}$ to any rate $\theta_k < \theta_{sm}^*$ is 'useless': it does not affect the estimate at period 2 and hence the objective probability of getting the estimation prize, and has an opportunity cost, in that \hat{p}_k could be assigned instead to θ_{sm}^* , thus increasing anticipatory utility from both prizes (if $a > 0$). In addition, when Abel chooses beliefs where the subjective mode is different from the objective mode, this comes at the cost of reducing the objective probability of getting the estimation prize. The only incentive for such a choice, therefore, is to sufficiently increase the subjectively expected payoff from the state prize, which requires choosing beliefs with a subjective mode at least equal to θ_{om} .

The remaining propositions study conditions for overestimation, i.e., the subjective mode being higher than the objective one, and for overprecision, which refers in our context to the length of the 95% subjective confidence interval. In very general terms, they express that the position chosen by Abel for the subjective mode has implications in the optimum regarding the level of overprecision. We distinguish two different situations.

Proposition V. Consider beliefs with subjective mode θ_{sm} . If $u(\bar{v}) \geq u(M) - u(M\theta_{sm})$, a necessary (but not sufficient) condition for these beliefs to be optimal is that they assign $\hat{p}_k = 0$ to any $\theta_k \neq \theta_{sm}$.

Proposition V implies that any beliefs with subjective mode ‘close’ to $\theta_k = 1$ cannot be optimal if they do not concentrate all the probability mass in that mode. In these circumstances, intuitively, anticipatory utility is maximized when Abel is certain to get the estimation prize. A problem of Proposition V is that it says little about the optimal value θ_{sm}^* and cannot be tested using directly observable data. However, Propositions IV and V imply the following corollary, which states a sufficient (although not necessary) condition in the optimum for maximal overprecision, i.e., a degenerate belief distribution. This condition is based on the value of the objective mode θ_{om} , which is determined by the evidence available to Abel and hence observable.

Corollary 1: If $u(\bar{v}) \geq u(M) - u(M\theta_{om})$, Abel assigns all the probability mass to one single rate. Its specific value depends on the curvature of the utility function of money, $u(x)$, but also on a . In particular, Abel is Bayesian, i.e., $\theta_{sm} = \theta_{om}$, if a is low enough or if the utility of money increases at a sufficiently lower rate than the posterior beliefs (as we move towards the objective mode).

To clarify how stringent condition $u(\bar{v}) \geq u(M) - u(M\theta_{om})$ is, note that $\bar{v} = 10$ and $M = 50$ imply that the condition is necessarily satisfied if Abel is risk-averse and $\theta_{om} > 0.8$. However, the condition can be also fulfilled for much lower values of θ_{om} if Abel displays sufficiently high levels of risk-aversion. Let θ_A denote the rate such that $u(\bar{v}) = u(M) - u(M\theta_A)$. Corollary 1 says that, if $\theta_{om} \geq \theta_A$, Abel is extremely overprecise, in the sense that he believes that there is only one possible state of nature. When $\theta_{om} < \theta_A$, in contrast, the following prediction shows that Abel’s posteriors can be more spread. In other words, beliefs can be nondegenerate only if $\theta_{om} < \theta_A$. The degree of overprecision, in other words, is conditional on the evidence received. In average, confidence intervals will be larger when θ_{om} is relatively small, in particular when $\theta_{om} < 1/2$.

Proposition VI. Consider beliefs with subjective mode θ_{sm} such that $u(\bar{v}) < u(M) - u(M\theta_{sm})$. Conditional on risk aversion, a necessary (but not sufficient) condition for these beliefs to be optimal is that they assign evenly as much probability as possible to some of the largest state(s), subject to $\hat{p}_{sm} > \hat{p}_k$ for any θ_k .

While Proposition VI is not very specific about the optimal θ_{sm}^* , a similar argument to that in Corollary 1 seems to apply as well for low values of θ_{om} . That is, again, $\theta_{sm}^* = \theta_{om}$ if a is low enough or if the utility of money increases at a sufficiently lower rate than the posterior beliefs (as we move towards the objective mode). In summary, risk aversion correlates with more Bayesian estimations (assuming that Abel’s degree of risk aversion is independent of his a). Intuitively, a subjective mode larger than θ_{om} increases the anticipatory utility from the state prize but reduces the chances of earning anything if the estimation prize is finally the selected one. Obviously, a very risk-

averse Abel strongly dislikes such possibility. Note also that, since θ_A depends on Abel's degree of risk aversion, a relatively risk-averse Abel would also be less likely to be over-precise, given some evidence characterized by θ_{om} , although this prediction is more complex to test and hence will not be considered in our posterior data analysis.

4.2 Data analysis

For starters, we consider the Bayesian model. Proposition I above states that, in any round, subjects should report estimations of θ that track the (rounded) frequency f of female balls in the sample observed by them so far (leaving aside extreme samples in the last estimation round).

Hypothesis I: In average, point estimations do not significantly differ from the average f .

Evidence: In average, the subjects' urns have around 56.7 female balls, i.e., the mean θ equals 0.567. In the samples corresponding to the first 15, 22, and 30 extractions, furthermore, the mean (non-rounded) f is 0.584, 0.577 and 0.578, respectively. Since the subjects' actual estimates of θ have averages equal to 0.517, 0.522 and 0.530, respectively, we observe a systematic (although small) underestimation of the number of female balls in all rounds. In this respect, the differences between the mean Bayesian and subjects' estimates are significant in the first two rounds (paired t-test, p-value = 0.018 and 0.021) but not so much in the final round (paired t-test, p-value = 0.068).

Observed frequency	Estimation round		
	1	2	3
[0, 0.2)	0,010	0,060	0,030
[0.2, 0.4)	0,060	-0,005	-0,025
[0.4, 0.6)	0,000	-0,010	0,000
[0.6, 0.8)	0,005	-0,020	0,000
[0.8, 1)	-0,140	-0,110	-0,020
Aggregated	-0,005	-0,020	-0,010

Table 1: Median deviation from the Bayesian estimation, conditional on the observed frequency

For a more disaggregate analysis, we define a subject's deviation in a round as the difference between her actual estimate of θ and the predicted Bayesian estimate. In Table 1, each column corresponds to one of the three estimation rounds, and there is one row for each of the intervals [0, 0.2); [0.2, 0.4), etc. Each cell indicates the subjects' median deviation, conditional on the estimation round and the value of f observed so far. Intuitively, the table compares the subjects' biases when most or the majority of news are bad, i.e., f low, and when most news are good. The Bayesian model predicts a nil deviation in each cell. In this respect, we see that the median deviation is practically

zero in several cells of the table. Interestingly, overestimation seems to be more systematic when f is low, whereas underestimation tends to occur when f is large. We will return to this point later.¹⁰

When we consider average, not median, deviations, the differences with the Bayesian prediction are a bit more pronounced –see Table A in Appendix II. Some inflation is observed in this case, particularly when f is low. However, it is far from systematic. Indeed, recall, the overall mean deviation is negative in any round. Further, in those contingencies where there is significant inflation or alternatively deflation, it is largely run by the presence of some outliers. This is illustrated by Figure 1, where each dot corresponds to a subject, placed according to her actual estimate in the third round, and her predicted, Bayesian estimate. As we see, the majority of the points are close to the diagonal. That is, most subjects’ estimates track the empirical frequency, except for a few cases.¹¹

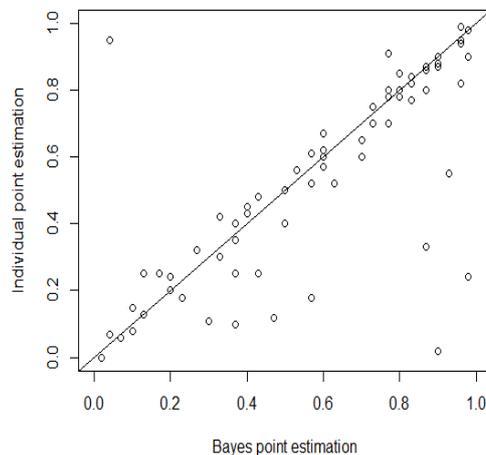


Figure 1: Subjects’ estimates in the last round vs. Bayesian prediction

Result I: The average and median subjects slightly underestimate θ in any round. Conditional on the actual frequency of female balls observed by the individual, f , inflation is observed at a systematic level only when f is low, and is rather small in size.

Consider now the models by AD and BP described in Section 4.1. As we proved there, these models imply that subjects should inflate the estimate of θ , that is, the average/median estimate should be significantly higher than the average/median observed f in any estimation round. Further,

¹⁰ Does this evidence signal a bias “towards 50%”, maybe because people have (inaccurate) priors assigning non-uniform probability to $\theta = 0.5$? We note in this respect that our control questionnaire explicitly asked whether priors were uniform (see Appendix I) and that Figure 1 is hardly consistent with such hypothesis, although we cannot exclude this possibility for a few subjects.

¹¹ As the figure shows, the underestimation observed in the third round is not *exclusively* due to the ‘distortions’ described in Footnote 8. Out of the 68 subjects, 10 of them faced in the last round a sample with 0, 1, 29, or 30 female balls, and just 7 of them observed either 29 or 30 balls, which are the only cases where some underestimation is predicted by the Bayesian model, although never higher than 2 balls. Yet the mean deviation among these subjects was -0.1495, i.e., around 15 balls, while the median one was -0.0267. Note though that one of these 7 subjects deviated in 76 balls from the Bayesian estimation; the mean deviation among the remaining 6 subjects is -0.0478.

inflation should not be conditional on f . While the evidence described in Result I is not very encouraging, Figure 1 above also shows that some people overestimate. This warrants further tests of those models. In what follows, therefore, we consider additional hypotheses based on Propositions II to VI. The next one follows from Propositions II and IV. Intuitively, the level of inflation depends on whether the estimate of θ is incentivized. Without incentives for accuracy, there is no risk of a material loss for having a mistaken belief about θ . Hence, utility is maximized if the belief is as optimistic as possible, which means that Abel should report $\hat{\theta} = 1$ in the first two, non-incentivized estimations. When there is some risk, in contrast, Abel should infer in a more Bayesian manner, so as to reduce the chance of a mistake (this is particularly true if Abel is very risk averse or cares little about anticipatory utility, i.e., has a small a). In the third estimation, that is, the average $\hat{\theta}$ should be strictly lower than 1 but higher than the average f , assuming that subjects display enough heterogeneity in a and risk aversion.

Hypothesis II: The prevalence and extent of inflation is higher in the first two rounds. In the third round, an optimistic subject's accuracy depends on her degree of risk aversion.

Evidence: Several findings speak against the first part of the hypothesis. To start, the share of subjects who give inflated estimates in the first, second, and third rounds equals 36.76, 35.29 and 33.82, respectively. In theory, there should be less people inflating in the third round, but the effect seems negligible (McNemar's test, p-value = 0.6171 and 0.7815 for the comparison of the first and second round with the third one, respectively). Second, people do not become more accurate as a result of the introduction of incentives in the third round. To check this, we consider the *absolute* value of a subject's deviation in a round, which measures the extent of her error. At first sight, we observe increased accuracy, as the mean of the absolute deviation after 15, 22 and 30 extractions is 0.1378, 0.1159 and 0.1060, respectively. An important question, however, is whether this increased accuracy is the result of a learning process or caused by the incentives introduced in the last estimation, as Hypothesis II contends. Thus, we estimate a linear panel data model where the Y variable is a subject's deviation from f (in absolute terms) and the explanatory variables are (i) the round number (1, 2 or 3), in order to measure learning effects; and (ii) a dummy for the third stage, to capture any additional effect due to the incentives provided. Note that learning effects should improve accuracy in the second and third rounds relative to the previous one. In this respect, we find that the coefficient of variable (i) is negative and significant (-0.0219, p-value = 0.023), although quantitatively modest, while that of variable (ii) is not significantly different from zero (p-value = 0.488). There is hence some limited learning effect, while the incentives introduced in the last estimation do not increase accuracy. We have also explored whether incentives affect the direction, if not the extent, of the bias. For this, we run a simple linear regression where the dependent variable is

a subject's deviation from f , i.e., not in absolute value, and the X-variable is a dummy taking value 1 when the estimation is made in the third, incentivized round. The coefficient of this variable is positive but non-significant (p-value = 0.698; results are similar with a panel data model). Observe that the positive sign, although non-significant, means that the deviations tend to move towards the positive side in the last round, something unpredicted by the models (people should inflate less frequently then).

We move now to the second part of Hypothesis II, together with a more thorough study of heterogeneity. As we have said, the overall evidence points out that most subjects do not exhibit a substantial bias. Yet averages can be misleading if some people inflate and others deflate. To check for heterogeneity, we compute the share of subjects who under-estimate θ never, once, twice or thrice across all rounds. The respective figures are 26.47, 20.59, 17.65 and 35.29. It can be worth to analyze what characterizes the subjects in this former group, i.e., the most systematically 'optimistic' ones, who never report an estimate lower than the Bayesian one (see Proposition IV above).¹² Before answering this question, however, it must be noted that these subjects do not exhibit very large deviations from f ; the median deviation is 0.05 and the average one is 0.10.¹³ For the sake of comparison, the median *absolute* deviation is 0.03, 0.08, and 0.13 among the subjects who under-estimate one, twice, and thrice, respectively.

A first thing we observe among the optimistic subjects is that they do not act more Bayesian in the third round, i.e., more accurate. This is indicated by the linear panel data model explored above: if we add a dummy for these subjects, they are not significantly more accurate in the third round (p-value = 0.482). A second thing is that there is a clear correlation between the observed frequency f and the degree of optimism, i.e., the number of rounds in which a subject does not under-estimate. Figure 2 below provides a comprehensive picture. Each box corresponds to a different group of subjects, i.e., those who underestimated in 0, 1, 2, or 3 rounds, and gives information about the mean f observed by those subjects across rounds. Specifically, the length of each box represents the inter-quartile range (IQR) in the corresponding distribution, whereas the vertical lines extend above (below) so as to include all data points within 1.5 IQR of the upper (lower) quartile, stopping at the largest such value. The horizontal line within each box, in turn, indicates the median value of the mean f observed in the associated distribution. As we can see, half of the participants that never underestimate θ observed a mean frequency lower than 0.345 while half of the participants who were consistently pessimistic across all rounds observed a mean f greater than 0.785.

¹² Our conclusions below are similar if we instead define an optimistic subject as one who gives an estimate strictly larger than the Bayesian one in all rounds.

¹³ This average deviation equals 0.06 if we remove one subject from the group whose average deviation was around 0.93.

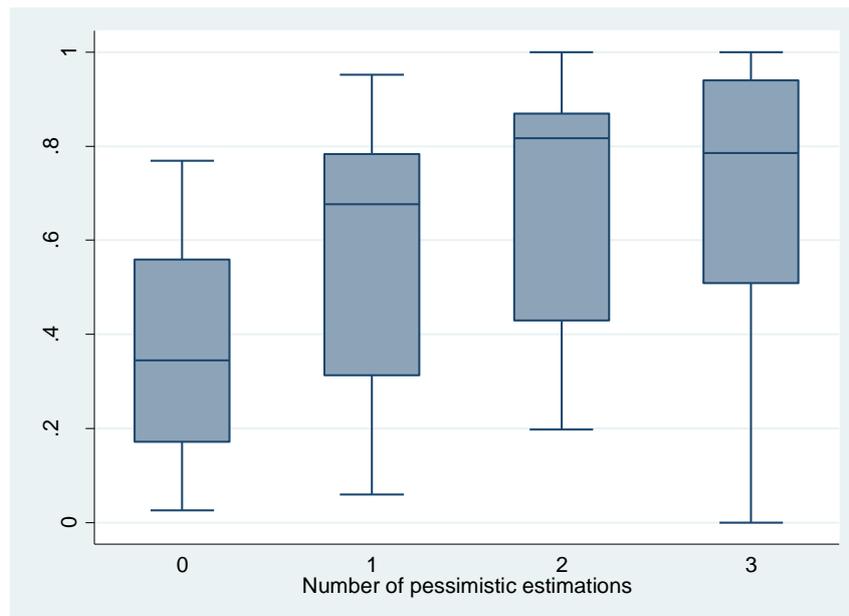


Figure 2: Mean observed frequency conditional on number of underestimations of θ

Therefore the ‘optimistic’ subjects tend to observe samples with relatively few female, i.e., good signals. Are they different from the other subjects in other respects? To study this issue, we run two regressions. To start, a logit regression finds no significant correlation between a binary variable taking value 1 when the subject strictly inflates in all rounds, and (i) any of our socio-demographic variables, (ii) the subject’s degree of risk aversion, (iii) the number of correctly recalled names (net of errors), i.e., with the subject’s memory capacity, and (iv) her knowledge of Statistics.¹⁴ The only exception is the CRT score: More reflexive people are significantly ($p = 0.041$) *more* likely to overestimate in all rounds.¹⁵ Similar results are obtained in a panel data model where the dependent variable is the subject’s deviation in each round: The only significant X-variables are the observed frequency (p -value < 0.001) and the CRT score (p -value = 0.008), which predict a negative and a positive effect, respectively. We stress that our index of risk aversion has no significant predictive power in the econometric models that we have specified, including one focused on the third round and the optimistic subjects: risk aversion does not correlate with a lower deviation, i.e., ‘more’ Bayesian estimates.

¹⁴ As a measure of their statistical knowledge, participants first answered the following question: “In an electoral survey with a sample of 10 voters randomly chosen, 40 percent of them stated they were voting for Party A. From this data and assuming that there are 1000 voters in the country, how many of them do you think will vote for Party A? Provide your best estimate, which must be a number between 0 and 1000.” In addition, they were also asked how many ECTS on Statistics and related subjects (Econometrics, Psychometrics, etc) they had passed in the last five years.

¹⁵ We have also included the amount of time that each subject takes to complete each estimation round, i.e., from the moment that the corresponding screen appears until the subject enters the estimate and proceeds to the next screen. Our hypothesis here is that optimistic subjects might respond relatively fast, without much thought, in these rounds. Yet none of these three variables, i.e., one per round, is significant.

To further explore the relationship between the CRT score and inflation, Figure 3 below represents the average deviation (grey bars) and the average absolute deviation (white bars), conditional on the subject’s CRT score. Two things are worth mentioning here. First, the size of the errors, i.e., the absolute values, tends to be higher for those subjects with low CRT scores, although the effect is not entirely systematic (subjects with a score of 4 have relatively large errors). Second, the CRT score is apparently related with the sign of the errors, as reflexive subjects tend to inflate; note yet that the degree of inflation is in average very small: these subjects tend to be optimistic, but very little. The following result summarizes our key findings so far.

Result II: The size and direction of the deviations does not depend on the round, and hence on the risk of a loss. The share of subjects who overestimate in all rounds is relatively small; moreover, these subjects deviate little from the Bayesian benchmark and do not appear different from other subjects, except in their CRT score and the sample observed (relatively few positive signals). Risk aversion does not predict more Bayesianism among the optimistic subjects in the third round. Overall, the evidence seems hardly consistent with models of optimism like AD and BP.

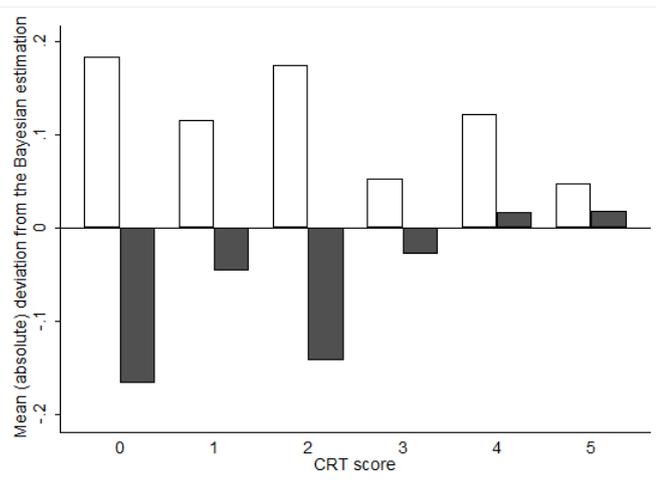


Figure 3: Mean deviation (gray) and mean absolute deviation (white) from the Bayesian estimation, conditional on CRT score

The following hypothesis explores the subjects’ degree of doubt in their inferences, as measured by the 95 percent confidence interval elicited after the third estimation round. If subjects can choose their beliefs, it seems at first sight natural that they should express little doubt, i.e., very narrow confidence intervals, particularly since the interval estimations are never incentivized and hence entail no risk –see Möbius et al. (2014) for an application of these ideas to financial markets. Specifically, optimists might assign all the probability mass to one single rate. As we have shown in Section 4.1, however, this kind of extreme over-precision must be present when $\theta_{om} \geq \theta_A$, but not necessarily when $\theta_{om} < \theta_A$, in which case posteriors might be more spread, conditional on risk

aversion. In addition, Proposition IV says that Abel will never assign positive probability to any rate below the subjective mode, i.e., the point estimation.

Hypothesis III: For an optimistic subject, confidence intervals are asymmetric, assigning in particular nil probability to any rate below $\hat{\theta}$. Further, they are larger when θ_{om} is relatively small, e.g., when $\theta_{om} \leq \frac{1}{2}$.

Evidence: Contrary to the hypothesis, the confidence intervals were generally symmetric around the last point estimation. Specifically, there are no significant differences between the mean last estimate of θ and the mean center of the confidence intervals (paired t-test, $p = 0.1014$). Figure 4 further illustrates this point. Note also that most subjects lie below the diagonal. Hence subjects report intervals whose midpoints tend to be slightly lower than the estimate of θ , contrary to the predictions by the optimism models.

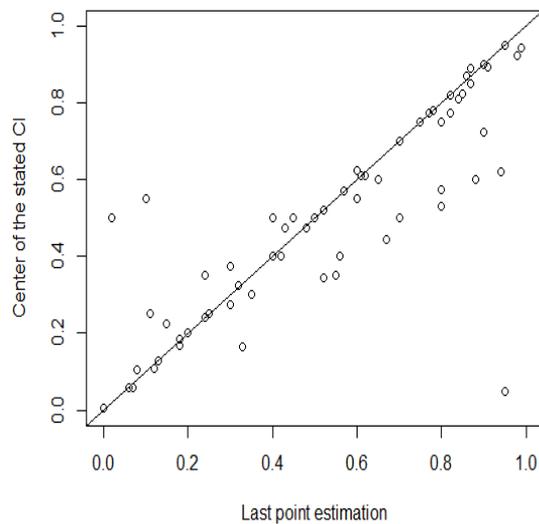


Figure 4: Subject's last estimate of θ vs. midpoint of the stated confidence interval

Note yet that Hypothesis III explicitly refers to the optimistic subjects. Hence, a relevant question is whether the people who never under-estimate θ report also asymmetric intervals. We check this first with a simple linear regression where the dependent variable is the difference between the midpoint of the interval reported by the subject and her last estimate (this difference is called D below), while the X -variable is a binary one taking value 1 when the subject never gives a deflated estimate. The coefficient happens to be negative, although non-significant (p -value = 0.287). Since the estimated constant is negative (and non-significant) as well, midpoints are even lower among the optimistic types, which hardly fits with Hypothesis III.¹⁶

¹⁶ If we control in the regression for the effect of the observed frequency f , closely correlated with optimism in our experiment, the coefficient of the dummy becomes marginally significant (p -value = 0.057), but it is still negative.

For further illustration, Figure 5 represents subjects according to their deviation from the Bayesian prediction in the *last* round (in the X-axis) and variable D. The box is divided in quarters, and the optimistic subjects (in that round at least) are the dots in the right-hand quarters. Those in the lower right-hand quarter, further, indicate an interval such that D is negative, contrary to what Hypothesis III predicts. We can see that many optimistic subjects are placed in such quarter and, in any case, the value of D is rarely large for any of those subjects.

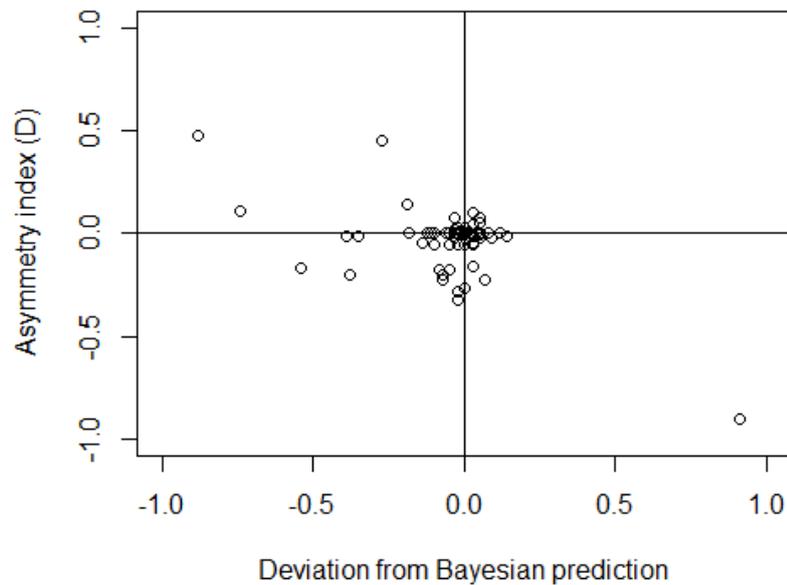


Figure 5: Subject's deviation from Bayesian prediction vs. asymmetry index D

With respect to the second half of Hypothesis III, it basically says that intervals should be of a relatively larger size when $f \leq 0.5$. Note first that the mean and median size of the elicited interval is 0.223 and 0.150, respectively. More specific to our hypothesis, further, the mean and median size of the elicited interval is 0.179 and 0.140 among participants with $f \leq 0.5$; and 0.253 and 0.180 respectively among participants with $f > 0.5$. Contrary to Hypothesis III, therefore, intervals are larger when $f > 0.5$, although this difference is not significant (t-test, p-value = 0.1576). For further detail, the dark circles in Figure 6 show the observed frequency f and the size of the elicited interval for each individual. Again, we observe that Hypothesis III is not satisfied, as the size of the intervals is larger when f is large, although the difference is not statistically significant, as we have observed. For comparison, Figure 6 also depicts the size of the 95 percent confidence interval of a Bayesian agent for each possible value of f (see the hollow circles). From this figure, it seems quite clear that most participants were overprecise, in the sense that their stated confidence interval were too tight relative to the Bayesian confidence interval. On the other hand, most underprecise individuals observed larger proportions of female names. We stress that similar findings have been also obtained

in the study by López-Pérez et al. (2020), where the elicited intervals were incentivized. This suggests that our results here are not an artifact of our experimental design.

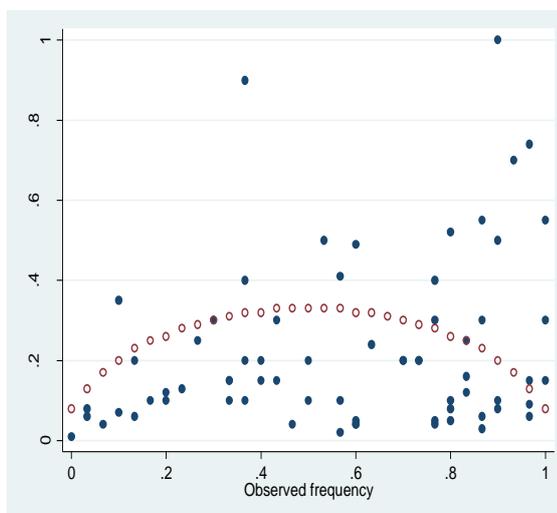


Figure 6: Size of the stated (dark circles) and Bayesian (hollow circles) CIs, by frequency.

Result III: The average subject does not systematically report ‘positively skewed’ intervals in the last round. This is also true in particular for those subjects who overestimate θ . The size of the intervals does not depend on the evidence observed.

Miscellaneous remarks

We briefly consider here three unrelated issues. The first one concerns model of optimism with cognitive costs. According to the model of choice of beliefs that we have considered, based on AD and BP, the only cost associated to optimistic beliefs is monetary. In the last estimation, inaccurate beliefs reduce the chance of winning the estimation prize, which according to the model should alleviate the participants’ estimation bias relative to the previous rounds. Yet, as alternative models suggest –see, e.g. Rabin (1994)–, avoiding evidence that is easily available or repressing unfavorable information are cognitively costly tasks. In our experiment, participants are in fact compelled to observe the evidence. Further, the information provided is easily understandable and the sample size is quite large. Considering these factors altogether, participants may find cognitively hard to ignore the evidence and to get swept up in optimism.

In a nutshell, models of optimism with cognitive costs predict that subjects will overestimate θ in our experiment, but only to a limited extent, similar in all rounds. Although some of our findings go well in line with these predictions, the evidence overall does not seem very supporting. On the negative side, we find that most subjects underestimate at least in one round, i.e., the share of subjects who never underestimate θ is relatively small. Moreover, these optimistic subjects tend to face samples with a reduced share of female balls, and it is not clear why cognitive costs of delusion

should be lower in that case. We have also seen that overestimation is somehow correlated with the subject's CRT score. Is this because more reflective subjects tend to be relatively more successful in repressing or avoiding the negative evidence? We find this conjecture a bit puzzling. On the positive side, it is true that the people who inflate θ always often do it to a reduced extent: As we have seen, in fact, the median deviation from the Bayesian estimate among these participants amounts to 5 balls. Further, the introduction of a payoff for accuracy in round 3 has little or no effect on the magnitude of the bias, which again is favorable to these models. In this sense, the models with cognitive costs seem to fare *relatively* better than models like BP.

A second issue that deserves some exploration concerns the intrinsic utility function that we have considered in Section 4.1, which depends just on the agent's monetary gain. It might be also the case, though, that (some) subjects do not like to be disappointed if their beliefs point too high and the reality happens to be mediocre. This is somehow the opposite of anticipatory utility, in that people deflate their expectations, and could be particularly important in our design, where subjects learn the actual state of the world at the end of the experiment.¹⁷ To check this point, we included a test so as to measure a subject's concern with disappointment and regret. Specifically, participants were asked to think of some experience in which their expectations had not been fulfilled. To illustrate the nature of the problem, we included several examples in which they could think, like their performance in a test, the behavior of a beloved person or the result of some bet. Then, they were asked to assess the following statements: (i) In this kind of situations, I tend towards anger or rage; (ii) In this kind of situations I tend towards sadness and/or to mull over the issue for a long time; (iii) I tend to prevent these situations by adjusting downwards my expectations; and (iv) When I live one of these situations, I find it hard to focus or think on other things. Subjects reported their agreement with each statement in a scale from 1 (totally disagree) to 5 (totally agree). The mean (median) answer to questions (i) to (iv) was, respectively, 3 (3), 3.90 (4), 3.09 (3), and 3.41 (4). When included in the panel data model introduced in the discussion of the evidence on Hypothesis II, these measures of disappointment and regret do not predict the observed bias, except for statement (iii), whose associated coefficient turns out to be positive and significant (0.051, p-value = 0.011). This, however, would mean that participants who claim to avoid disappointment by adjusting downwards their expectations are in fact more prone to provide less pessimistic or even optimistic estimations, other things being equal. Although it is unnecessary for our purposes to find an explanation for this paradoxical result, it might be the case that the participants' answers to statement (iii) give information about their *awareness* of their downward bias. If so, it is plausible that those participants

¹⁷ This is not unusual in the literature; see for instance Gotthard–Real (2017) and Barron (2020).

who are conscious of their bias try to correct them, therefore providing less pessimistic or even optimistic estimations.

A third issue is that, to get a better insight of the participants' heterogeneity in terms of optimism, we also conducted the Revised Life Orientation Test (LOT-R), a widely used instrument in psychology to assess the level of optimism. The LOT-R comprises ten statements and the respondents must indicate their agreement to each one using a scale from 1 (strongly disagree) to 5 (strongly agree). Three of the statements measure optimism directly, while other three statements measure pessimism. The four remaining statements are fillers and they are not considered in the calculation of the LOT-R score, which is computed as the sum of scores in the statements about optimism and pessimism (for the latter, the scores are previously reversed) and it is comprised between 6 (strongly pessimistic) and 30 (strongly optimistic). The mean and median LOT-R score were 20.71 and 22, respectively.¹⁸ When included in the panel data model introduced in the discussion of the evidence on Hypothesis II, neither the aggregate score or the scores of the different statements were found to have a significant effect on the observed bias.

Result IV: Models of optimism with cognitive costs are inconsistent with: (a) most of the participants underestimate θ at least once and (b) overestimation is more prevalent among individuals who observed relatively few female balls. Measures of the participants' degree of optimism, pessimism and disappointment aversion hardly explain the observed biases.

5. Conclusion

Subjects in our experiment face a rather simple problem of inference, that is, estimating the mode of their posterior beliefs by extrapolation from a sample. Moreover, we have a strong control over the subjects' priors and the signals they observe. While most subjects track rather closely the Bayesian prediction, Figure 1 shows that a fraction of them deviate. However, they tend to underestimate the mode, not inflate it. Further, underestimation is equally prevalent when there is no prize for accuracy, it is unrelated to personal characteristics like risk aversion, and subjects rarely report skewed confidence intervals. The preference for the state prize to be high, we conclude, hardly motivates deviations from Bayes' rule in our context. Overall, the evidence is not supportive of the models by AD and BP, whereas models with cognitive costs are empirically more relevant, but only in relative terms.

The two families of models state sufficient conditions for a positivity bias, but our findings in this respect are mostly negative. On one hand, models like AD and BP predict that the prevalence or

¹⁸ The data relative to one of the statements of pessimism were corrupted for participants in the first sessions. The results provided correspond to the remaining subsample ($n=21$).

‘demand’ of the bias should be maximal when its price is nil or low. In our experiment, however, we find always a very small, arguably negligible bias. Relatedly, these models also say that the prevalence of the bias should decrease as its material cost increases. Our findings do not support this idea, at least within our payoff constellation,. Taking into account additional evidence surveyed in Section 2, the relation between the bias and its ‘price’, if any, seems hardly linear. While the demand for the bias could be further explored in a design similar to ours, but with a very large estimation prize, e.g., 100 euros if the estimate is accurate, we are however unconvinced that this possibility is worth the cost given the extremely limited ‘demand’ of optimism in our setting, where the price is low.¹⁹ On the other hand, models with cognitive costs say that the extent of the bias is a function of the proportion of individuals with low costs. The models are not very specific about the determinants of those costs, but the limited evidence for optimism that we find suggests either that (i) our sample was biased towards agents with high costs or (ii) cognitive costs are not so essential for the occurrence of optimism in our setting. Perhaps these costs are negatively related to the complexity of the inference problem which, arguably, was low in our experiment. This suggests a potential line of investigation.

In this line, our plan for future research is to propose alternative sufficient environmental or personal conditions for optimism and extend our experimental design to test them. For instance, the models with cognitive costs are sometimes not specific about the determinants of these costs. One could guess however that the size of the sample and its informativeness are relevant in this regard. In particular, inflation might be more prevalent and acute when the sample size is small, or the posteriors of several beliefs are similar. For a second line of research, one of our conjectures is that optimism failed to appear in our setting because learning was too ‘transparent’ and hence did not lead to a crowd-out of attention, i.e., to a focus on those aspects of the problem that were more beneficial to the decider; Epley and Gilovich (2016). According to this conjecture, therefore, a sufficient condition for optimism is complexity coupled with incentives to pay attention on the stories or details most beneficial. More research is warranted.

¹⁹ Alternatively, one could explore further whether the demand of optimism depends on how desirable the positive beliefs are, e.g., using a state prize of 20 euros per female ball.

Bibliography

- Akerlof, G. A., & Dickens, W. T. (1982). The Economic Consequences of Cognitive Dissonance. *The American Economic Review*, 72(3), 307–319. Retrieved from <http://www.jstor.org/stable/1831534>
- Aliber, R. Z., & Kindleberger, C. P. (2015). *Manias, Panics, and Crashes*. London: Palgrave Macmillan UK. <https://doi.org/10.1007/978-1-137-52574-1>
- Alloy, L. B., & Abramson, L. Y. (1979). Judgment of contingency in depressed and nondepressed students: Sadder but wiser? *Journal of Experimental Psychology: General*. US: American Psychological Association. <https://doi.org/10.1037/0096-3445.108.4.441>
- Barron, K. (2020). Belief updating: does the ‘good-news, bad-news’ asymmetry extend to purely financial domains? *Experimental Economics*. <https://doi.org/10.1007/s10683-020-09653-z>
- Bénabou, R., & Tirole, J. (2002). Self-Confidence and Personal Motivation. *The Quarterly Journal of Economics*, 117(3), 871–915. <https://doi.org/10.1162/003355302760193913>
- Bénabou, R., & Tirole, J. (2004). Willpower and Personal Rules. *Journal of Political Economy*, 112(4), 848–886. <https://doi.org/10.1086/421167>
- Bénabou, R., & Tirole, J. (2016). Mindful Economics: The Production, Consumption, and Value of Beliefs. *Journal of Economic Perspectives*, 30(3), 141–164. <https://doi.org/10.1257/jep.30.3.141>
- Benjamin, D. J. (2019). Errors in probabilistic reasoning and judgment biases. *Handbook of Behavioral Economics: Applications and Foundations 1, 2*, 69–186. <https://doi.org/10.1016/BS.HESBE.2018.11.002>
- Blanco, M., Engelmann, D., Koch, A. K., & Normann, H.-T. (2010). Belief elicitation in experiments: is there a hedging problem? *Experimental Economics*, 13(4), 412–438. <https://doi.org/10.1007/s10683-010-9249-1>
- Bracha, A., & Brown, D. J. (2012). Affective decision making: A theory of optimism bias. *Games and Economic Behavior*, 75(1), 67–80. <https://doi.org/https://doi.org/10.1016/j.geb.2011.11.004>
- Brunnermeier, M. K., & Parker, J. A. (2005). Optimal Expectations. *American Economic Review*, 95(4), 1092–1118. Retrieved from <http://www.aeaweb.org/articles?id=10.1257/0002828054825493>
- Buser, T., Gerhards, L., & van der Weele, J. (2018). Responsiveness to feedback as a personal trait. *Journal of Risk and Uncertainty*, 56(2), 165–192. <https://doi.org/10.1007/s11166-018-9277-3>
- Caballero, A., & López-Pérez, R. (2020a). Economic Models of Optimism: What Does the Evidence Say? *Cuadernos Económicos de ICE*, 1(99). <https://doi.org/10.32796/cice.2020.99.7008>
- Caballero, A., & López-Pérez, R. (2020b). *Self-serving recall is not a sufficient cause of optimism: An experiment*.
- Case, K. E., & Shiller, R. J. (2003). Is There a Bubble in the Housing Market? *Brookings Papers on Economic Activity*, 2003(2), 299–362. <https://doi.org/10.1353/eca.2004.0004>
- Charness, G., & Dave, C. (2017). Confirmation bias with motivated beliefs. *Games and Economic Behavior*, 104, 1–23. <https://doi.org/https://doi.org/10.1016/j.geb.2017.02.015>
- Coutts, A. (2019a). Good news and bad news are still news: experimental evidence on belief updating. *Experimental Economics*, 22(2), 369–395. <https://doi.org/10.1007/s10683-018-9572-5>
- Coutts, A. (2019b). Testing models of belief bias: An experiment. *Games and Economic Behavior*,

113, 549–565. <https://doi.org/https://doi.org/10.1016/j.geb.2018.11.001>

- Eil, D., & Rao, J. M. (2011). The Good News-Bad News Effect: Asymmetric Processing of Objective Information about Yourself. *American Economic Journal: Microeconomics*, 3(2), 114–138. <https://doi.org/10.1257/mic.3.2.114>
- Engelmann, J., Lebreton, M., Schwardmann, P., van der Weele, J. J., & Chang, L.-A. (2019). Anticipatory Anxiety and Wishful Thinking. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3408017>
- Epley, N., & Gilovich, T. (2016). The Mechanics of Motivated Reasoning. *Journal of Economic Perspectives*, 30(3), 133–140. <https://doi.org/10.1257/jep.30.3.133>
- Ertac, S. (2011). Does self-relevance affect information processing? Experimental evidence on the response to performance and non-performance feedback. *Journal of Economic Behavior & Organization*, 80(3), 532–545. <https://doi.org/https://doi.org/10.1016/j.jebo.2011.05.012>
- Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2), 171–178. <https://doi.org/10.1007/s10683-006-9159-4>
- Frederick, S. (2005). Cognitive Reflection and Decision Making. *Journal of Economic Perspectives*, 19(4), 25–42. <https://doi.org/10.1257/089533005775196732>
- Garrett, N., Sharot, T., Faulkner, P., Korn, C. W., Roiser, J. P., & Dolan, R. J. (2014). Losing the rose tinted glasses: neural substrates of unbiased belief updating in depression. *Frontiers in Human Neuroscience*. Retrieved from <https://www.frontiersin.org/article/10.3389/fnhum.2014.00639>
- Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review*, 102(4), 684–704. <https://doi.org/10.1037/0033-295X.102.4.684>
- Gotthard-Real, A. (2017). Desirability and information processing: An experimental study. *Economics Letters*, 152, 96–99. <https://doi.org/https://doi.org/10.1016/j.econlet.2017.01.012>
- Grossman, Z., & Owens, D. (2012). An unlucky feeling: Overconfidence and noisy feedback. *Journal of Economic Behavior & Organization*, 84(2), 510–524. <https://doi.org/https://doi.org/10.1016/j.jebo.2012.08.006>
- Heger, S. A., & Papageorge, N. W. (2018). We should totally open a restaurant: How optimism and overconfidence affect beliefs. *Journal of Economic Psychology*, 67, 177–190. <https://doi.org/https://doi.org/10.1016/j.joep.2018.06.006>
- Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, 108(3), 480–498. <https://doi.org/10.1037/0033-2909.108.3.480>
- López-Pérez, R., Rodríguez-Moral, A., & Vorsatz, M. (2020). *Simplified mental representations as a cause of overprecision (manuscript)*.
- Ma, Y., Li, S., Wang, C., Liu, Y., Li, W., Yan, X., ... Han, S. (2016). Distinct oxytocin effects on belief updating in response to desirable and undesirable feedback. *Proceedings of the National Academy of Sciences*, 113(33), 9256 LP – 9261. <https://doi.org/10.1073/pnas.1604285113>
- Mayraz, G. (2013). *Wishful thinking in predictions of asset prices*.
- Möbius, M. M., Niederle, M., Niehaus, P., & Rosenblat, T. S. (2011). Managing Self-Confidence: Theory and Experimental Evidence. *National Bureau of Economic Research Working Paper Series, No. 17014*. <https://doi.org/10.3386/w17014>

- Oster, E., Shoulson, I., & Dorsey, E. R. (2013). Optimal Expectations and Limited Medical Testing: Evidence from Huntington Disease. *American Economic Review*, *103*(2), 804–830. Retrieved from <http://www.aeaweb.org/articles?id=10.1257/aer.103.2.804>
- Rabin, M. (1994). Cognitive dissonance and social change. *Journal of Economic Behavior & Organization*, *23*(2), 177–194. [https://doi.org/10.1016/0167-2681\(94\)90066-3](https://doi.org/10.1016/0167-2681(94)90066-3)
- Rasmussen, H. N., Scheier, M. F., & Greenhouse, J. B. (2009). Optimism and Physical Health: A Meta-analytic Review. *Annals of Behavioral Medicine*, *37*(3), 239–256. <https://doi.org/10.1007/s12160-009-9111-x>
- Scheier, M. F., & Carver, C. S. (1985). Optimism, coping, and health: Assessment and implications of generalized outcome expectancies. *Health Psychology*. US: Lawrence Erlbaum Associates. <https://doi.org/10.1037/0278-6133.4.3.219>
- Scheier, M. F., Carver, C. S., & Bridges, M. W. (1994). Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): a reevaluation of the Life Orientation Test. *Journal of Personality and Social Psychology*, *67*(6), 1063–1078. <https://doi.org/10.1037//0022-3514.67.6.1063>
- Schlag, K. H., Tremewan, J., & van der Weele, J. J. (2015). A penny for your thoughts: a survey of methods for eliciting beliefs. *Experimental Economics*, *18*(3), 457–490. <https://doi.org/10.1007/s10683-014-9416-x>
- Sharot, T., Korn, C. W., & Dolan, R. J. (2011). How unrealistic optimism is maintained in the face of reality. *Nature Neuroscience*, *14*(11), 1475–1479. <https://doi.org/10.1038/nn.2949>
- Shiller, R. J. (2000). *Irrational Exuberance*. Princeton University Press.
- Strunk, D. R., Lopez, H., & DeRubeis, R. J. (2006). Depressive symptoms are associated with unrealistic negative predictions of future life events. *Behaviour Research and Therapy*, *44*(6), 861–882. <https://doi.org/https://doi.org/10.1016/j.brat.2005.07.001>
- Sunstein, C. R., Bobadilla-Suarez, S., Lazzaro, S. C., & Sharot, T. (2016). How people update beliefs about climate change: Good news and bad news. *Cornell L. Rev.*, *102*, 1431.
- Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, *90*(4), 293–315. <https://doi.org/10.1037/0033-295X.90.4.293>
- von Hippel, W., & Trivers, R. (2011). The evolution and psychology of self-deception. *Behavioral and Brain Sciences*, *34*(1), 1–16. <https://doi.org/DOI: 10.1017/S0140525X10001354>
- Wicklund, R. A., & Brehm, J. W. (1976). *Perspectives on cognitive dissonance*. *Perspectives on cognitive dissonance*. Oxford, England: Lawrence Erlbaum.
- Wiswall, M., & Zafar, B. (2015). How Do College Students Respond to Public Information about Earnings? *Journal of Human Capital*, *9*(2), 117–169. <https://doi.org/10.1086/681542>
- Zimmermann, F. (2020). The Dynamics of Motivated Beliefs. *American Economic Review*, *110*(2), 337–361. <https://doi.org/10.1257/aer.20180728>

Appendix I: Instructions for the control treatment

Thank you for participating in this experiment on Behavioral and Experimental Economics. You will be paid some money at its end; the precise amount will depend on chance and your decisions. All your decisions will be confidential, that is, the other participants will not get any information about your decisions, nor do you get any information about the others' decisions. In addition, your decisions will be anonymous: during the experiment, you will not have to enter your name at any time.

Decisions are made via the keyboard of your computer terminal. Read the on-screen instructions carefully before making any decision; there is no hurry to decide. These instructions meet the basic standards in Experimental Economics; in particular, all the information that appears in them is true and therefore there is no deception.

Please, do not talk to any other participant. If you do not follow this rule, we will have to exclude you from the experiment without payment. If you have questions, raise your hand and we will assist you. The use of calculators and writing tools is not permitted. Please, switch off your cell phone.

Description of the experiment

There is a 'virtual urn' with 100 balls. Each ball has written a different name of girl or boy; any of these names is used with a relatively high frequency in Spain. Let us call F the actual number of balls with a female name in your urn. You do not know either F or the number of balls in your urn with boy name (that is, $100 - F$). You only know that the value of F has been randomly selected by the computer from among all integers between 0 and 100, both included (this means a total of 101 numbers, as 0 is included as well). Therefore, the probability that one of these potential values of F has been chosen is a priori of $1/101$, that is, slightly less than 1%. Important: The value of F will not change throughout the experiment; the urn has always the same content.

During the experiment, the computer will perform several extractions from the urn, randomly and with replacement –in other words: each draw is reintroduced into the urn and can therefore be drawn in the next extraction. Each of the 100 balls has the same chance in each extraction. The computer will show you the name written in each extraction, one by one. Between some of the extractions, you will receive instructions to complete some questionnaire or perform some task.

Once you have completed all questionnaires and tasks, you will be paid in private and in cash. In this regard, you will receive 3 Euros for participating in the experiment, plus an additional payment that will depend of three 'prizes'. **Prize 1:** you receive 0.50 Euros for each ball in your urn with a girl

name. In other words, if there are F balls with female name in the urn, this prize equals $0.5 \times F$. **Prize 2** will be explained later, but will depend on one of the tasks to be performed. The same can be said about **prize 3**. Important: You can only win either prize 1 or prize 2. You do not know now which of them you will win; this will be determined randomly at the end of the experiment, choosing then one of the two prizes with a 50% probability. On the contrary, winning prize 3 is compatible with winning either prize 1 or 2. Observe finally that the prizes are always independent of each other. For example, what you win with prize 3 will not depend on how you have performed in the task corresponding to prize 2, and vice versa.

If you have any questions, please raise your hand and we will attend you.

Examples of screenshots

We have finished with the ball draws and proceed to another task.

The task consists in estimating again the number of balls with a girl name that you believe there are in the urn, based on all previous draws. Note: the estimation must be an integer between 0 and 100, both included.

IMPORTANT: Remember that you will get 3 Euros for participating, plus an additional payment that depends on three prizes (1, 2 and 3). **Prize 1** consists in 0.50 Euros for each ball with a girl name in the urn. **Prize 2** depends on your estimate in **this task** and it is equal to (I) 10 Euros if you guess the exact number of balls with a girl name in the urn, with a maximum error allowed of plus or minus 2 balls, or (II) 0 Euros if your error is larger than 2 balls. **Prize 3** will be explained later.

Once you have made your estimation, click the "Continue" button.

My estimate of F is:

Screenshot for the third estimation task, i.e., the incentivized one

Please, answer the following questions about the content of the instructions. In order to proceed to the next stage, you must answer all the questions correctly. If you have any doubt, please raise your hand and we will assist you.

Once you have finished, click the "Continue" button.

How many balls are there in your urn?

Do you know exactly how many balls in your urn have a girl name? Yes No

Could there be zero balls with a girl name in your urn? Yes No

Which is the maximum number of balls with boy names that there could be in your urn?

There are F balls with female names in your urn. F has been determined randomly, choosing an integer between 0 and 100. Do any of these 101 numbers between 0 and 100 have a larger chance of being chosen than any other? Yes No

Can you get both prize 1 and prize 2 simultaneously? Yes No

If you get prize 1 and there are 90 balls with girl names in your urn. Which would be the exact amount of that prize in Euros?

If you get prize 1 and there are 60 balls with boy names in your urn. Which would be the exact amount of that prize in Euros?

Screenshot for the control questions

Note: In the last two questions in this screenshot, the respective numbers of balls with girl and boy name in the urn were determined randomly for each subject; the selected numbers could be any multiple of 10 between 0 and 100, so as to facilitate mental computations.

Appendix II: Additional data

Observed frequency	First estimation	Second estimation	Third estimation
[0, 0.2)	0,154	0,138	0,127
[0.2, 0.4)	0,024	-0,055	-0,042
[0.4, 0.6)	-0,085	-0,061	-0,079
[0.6, 0.8)	-0,015	-0,018	-0,008
[0.8, 1)	-0,174	-0,146	-0,127
Aggregated	-0,066	-0,054	-0,048

Table A: Mean deviation from the Bayesian estimation, conditional on the observed frequency

Appendix III: Proofs

Proposition II. If $a = 0$, then any subjective belief such that θ_{sm} coincides with the objective mode θ_{om} is optimal.

Proof: if $a = 0$, Abel chooses beliefs that maximize $\frac{1}{2} [\sum_{\theta} u(M\theta_k) \cdot p_k + u(\bar{v})p_{sm}]$. Since p_{sm} is the objective probability of the subjective mode, and it is maximized when the subjective mode coincides with the objective mode, any distribution of subjective probabilities such that $\theta_{sm} = \theta_{om}$ is optimal. ■

Proposition III. If $\bar{v} = 0$, then the optimal beliefs are characterized by $\hat{p}_k = 1$ for $\theta_k = 1$ and $\hat{p}_k = 0$ for any $\theta_k < 1$.

Proof: if $\bar{v} = 0$, Abel chooses beliefs that maximize $\frac{1}{2} \sum_{\theta} u(M\theta_k)(a\hat{p}_k + p_k)$. It is straightforward that this expression is maximized for beliefs such that $\hat{p}_k = 1$ for $\theta_k = 1$ and $\hat{p}_k = 0$ for any $\theta_k < 1$. ■

Proposition IV. The optimal subjective mode θ_{sm}^* is at least equal to the objective mode. If $a > 0$, further, optimal beliefs are characterized by $\hat{p}_k = 0$ for any $\theta_k < \theta_{sm}^*$.

Proof: consider subjective beliefs (A) such that $\theta_{sm} = \theta_k < \theta_{om}$, $\hat{p}_k = x$ and $\hat{p}_{om} = y$ ($x \geq y$). Now consider beliefs (B), identical to (A) except by the fact that $\hat{p}_k = y$ and $\hat{p}_{om} = x$ so that $\theta_{sm} = \theta_{om}$ (if $x = y$ and there are hence multiple modes, condition $\theta_{sm} = \theta_{om}$ means that the estimate chosen in period 2 is θ_{om} when beliefs are (B)). The expected utility with beliefs (B) is larger than with (A) if

$$a \frac{1}{2} [(x - y)(u(M\theta_{om}) - u(M\theta_k))] + \frac{1}{2} u(\bar{v})(p_{om} - p_k) > 0$$

which holds for any $\theta_k < \theta_{om}$, since $p_{om} > p_k$ and $x \geq y$. Therefore, the optimal subjective mode cannot be lower than θ_{om} . For the second part of the proposition, consider beliefs such that $\hat{p}_k > 0$ for some state $\theta_k < \theta_{sm}$. If Abel transfers all the subjective probability from state θ_k to θ_{sm} , his expected utility increases if

$$a \frac{1}{2} [\hat{p}_k (u(M\theta_{sm}) + u(\bar{v}) - u(M\theta_k))] > 0$$

which holds necessarily since $\theta_{sm} > \theta_k$ and $a > 0$. Therefore, any beliefs such that $\hat{p}_k > 0$ for any state $\theta_k < \theta_{sm}$ are suboptimal. ■

Proposition V. Consider beliefs with subjective mode θ_{sm} . If $u(\bar{v}) \geq u(M) - u(M\theta_{sm})$, a necessary (but not sufficient) condition for these beliefs to be optimal is that they assign $\hat{p}_k > 0$ to any $\theta_k \neq \theta_{sm}$.

Proof: the proof of Proposition IV above shows that it is not optimal to assign strictly positive probability to any rate lower than θ_{sm} . Assume now, without loss of generality, that the posteriors

assign strictly positive probability to some states larger than θ_{sm} (if any) and in particular to the largest state, i.e., $\theta_k = 1$ in our experiment. Consider also posteriors (B), identical to (A) except that some probability mass $\varepsilon \geq 0$ is transferred from $\theta_k = 1$ to θ_{sm} . Taking (3) into account, the expected utility of (B) will be higher than that of (A) if

$$\varepsilon \cdot a \cdot u(M\theta_{sm}) - \varepsilon \cdot a \cdot u(M) + \varepsilon \cdot a \cdot u(\bar{v}) \geq 0 \leftrightarrow u(\bar{v}) \geq u(M) - u(M\theta_{sm}) \quad (4)$$

While condition (4) is obtained given a transfer from $\theta_k = 1$ to θ_{sm} , it also ensures that a transfer of probability from any state between 1 and θ_{sm} (if any) to θ_{sm} improves Abel's utility. If (4) holds, therefore, Abel is better if he concentrates the probability mass in θ_{sm} , instead of spreading part of it among some larger states (keeping at the same time the mode in θ_{sm}). ■

Corollary 1: If $u(\bar{v}) \geq u(M) - u(M\theta_{om})$, Abel assigns all the probability mass to one single rate. Its specific value depends on the curvature of the utility function of money, $u(x)$, but also on a . In particular, Abel is Bayesian, i.e., $\theta_{sm} = \theta_{om}$, if a is low enough or if the utility of money increases at a sufficiently lower rate than the posterior beliefs (as we move towards the objective mode).

Proof: Proposition IV says that, in the optimum, the subjective mode must be some rate between θ_{om} and 1. When condition (4) is satisfied for $\theta_{sm} = \theta_{om}$, therefore, Abel's optimal beliefs must necessarily concentrate all the mass in some rate. It follows that the optimal beliefs, i.e., the value of θ_{sm}^* , can be determined by comparing the utility of the potential degenerate distributions. For example, it turns out that a probability distribution with the whole mass in θ_{om} is optimal if

$$au(M\theta_{om}) + p_{om} \cdot u(\bar{v}) \geq au(M\theta_k) + p_k \cdot u(\bar{v}) \quad \forall \theta_k > \theta_{om}$$

Rearranging terms, this can be expressed as

$$p_{om} - p_k \geq \frac{a[u(M\theta_k) - u(M\theta_{om})]}{u(\bar{v})} \quad \forall \theta_k > \theta_{om} \quad (5)$$

which holds true as far as a is low enough and/or the utility of money increases at a sufficiently lower rate than the posterior beliefs (as we move towards the objective mode). ■

Proposition VI. Consider beliefs with subjective mode θ_{sm} such that $u(\bar{v}) < u(M) - u(M\theta_{sm})$. Conditional on risk aversion, a necessary (but not sufficient) condition for these beliefs to be optimal is that they assign evenly as much probability as possible to some of the largest state(s), subject to $\hat{p}_{sm} > \hat{p}_k$ for any θ_k .

Proof: consider again beliefs (A) and (B), described in the proof of Proposition V. For the sake of the exposition, assume $\theta_{sm} < 1$, so that the probability mass is not entirely assigned to $\theta_k = 1$. If $u(\bar{v}) < u(M) - u(M\theta_{sm})$, we have shown that Abel is better by transferring some probability from θ_{sm} to $\theta_k = 1$ or in fact from any θ_k such that $\theta_{sm} < \theta_k < 1$, at least as far as θ_{sm} remains the only

subjective mode (or the one selected as an estimate in period 2 if there are several modes). A relevant question is therefore if it can be optimal to transfer all probability to $\theta_k = 1$, and the answer is not affirmative in general. To check this point, it suffices to compare the expected utility of two degenerate beliefs: one where all the probability mass is in $\theta_{sm} < 1$, and another where the mass is concentrated in $\theta_k = 1$, with objective probability p_1 . The latter beliefs give higher expected utility if

$$a \frac{1}{2} [u(M) + u(\bar{v})] + \frac{1}{2} u(\bar{v}) p_1 > a \frac{1}{2} [u(M\theta_{sm}) + u(\bar{v})] + \frac{1}{2} [u(\bar{v}) p_{sm}]$$

that is, if

$$a [u(M) - u(M\theta_{sm})] > u(\bar{v}) (p_{sm} - p_1)$$

which is not necessarily true under our conditions, because a can be low and $p_{sm} - p_1$ large relative to $u(M) - u(M\theta_{sm})$, which depends on the curvature of the utility function. Although we cannot find a closed-form solution for θ_{sm}^* , hence, we can at least say that it can be lower than 1 in some circumstances, and the probability mass will be in that case spread. One possibility is that optimal beliefs entail subjective probabilities $\hat{p}_{sm} = 0.5$ and $\hat{p}_1 = 0.5$ (if Abel expects $\theta_k = 1$ to be chosen in period 2, he should assign infinitesimally less probability to $\theta_k = 1$). Alternatively, Abel might find optimal to assign some positive probability ε also to the second largest state $\theta = 0.99$. In fact, this is better than assigning positive probability only to the subjective mode and the largest state if

$$\begin{aligned} & a \frac{1}{2} \left[\left(\frac{1}{2} - \frac{\varepsilon}{2} \right) u(M\theta_{sm}) + \left(\frac{1}{2} - \frac{\varepsilon}{2} \right) u(M) + \varepsilon u(0.99M) + \left(\frac{1}{2} - \frac{\varepsilon}{2} \right) u(\bar{v}) \right] \\ & > a \frac{1}{2} \left[\frac{1}{2} u(M\theta_{sm}) + \frac{1}{2} u(M) + \frac{1}{2} u(\bar{v}) \right] \end{aligned}$$

That is, if

$$u(0.99M) > \frac{1}{2} [u(M\theta_{sm}) + u(M) + u(\bar{v})]$$

We are assuming that Abel transfers the same mass of probability from the subjective mode and the largest state to the second largest state. Note that Abel can never be better off by transferring probability from the largest state only. However, given the previous probabilities $\hat{p}_{sm} = 0.5$ and $\hat{p}_1 = 0.5$, the restriction that the subjective mode is unchanged implies that the probability must be transferred in the same amount from θ_{sm} and the largest state. We can generalize the last condition so that Abel will assign positive probability to the n -th largest state θ_k if

$$M\theta_k > \frac{M\theta_{sm} + v + M \sum_{k+1}^{100} \theta_i}{n}$$