Overconfidence as a social bias: Experimental evidence

Till Proeger 1, Lukas Meub ∗,1
Faculty of Economic Sciences, Chair of Economic Policy and SME Research, University of Goettingen, Platz der Goettinger Sieben 3, 37073, Goettingen, Germany

HIGHLIGHTS

- We compare the overconfidence for an individual and a social, observational experiment.
- We find realistic confidence levels in the individual setting on average.
- Introducing the observation of other subjects considerably increases the overconfidence.
- The social setting reduces underconfidence compared to the individual setting.
- Our results suggest that overconfidence is a social rather than an individual bias.

ARTICLE INFO

Article history:
Received 27 August 2013
Received in revised form 8 November 2013
Accepted 22 November 2013
Available online 4 December 2013

JEL classification:
C91
D03
D83

Keywords:
Laboratory experiment
Overconfidence bias
Real effort
Self-assessment
Social interaction

ABSTRACT

The overconfidence bias is discussed extensively in economic studies, yet fails to hold experimentally once monetary incentives and feedback are implemented. We consider overconfidence as a social bias. For a simple real effort task, we show that, individually, economic conditions effectively prevent overconfidence. By contrast, the introduction of a very basic, purely observational social setting fosters overconfident self-assessments. Additionally, observing others’ actions effectively eliminates underconfidence compared to the individual setting.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

The overconfidence bias is a firmly established feature of individual behavior in psychological research. In economics, it is put forth extensively to explain inefficient market outcomes connected to a biased sense of self-confidence. Recent studies on overconfidence cover, among others, the role of investor experience (Menkhoff et al., 2013), CEOs’ overconfidence and investment (Malmendier and Tate, 2005), insurance behavior (Sandroni and Squintani, 2013), gender and investment (Barber and Odean, 2001), monetary policy decisions (Claussen et al., 2012), financial trading (Biais et al., 2005), employee incentive schemes (Larkin and Leider, 2012) and pricing decisions for consumer goods (Grubb, 2009). The majority of economic studies invoking the dismal effects of overconfidence either build on overconfidence as a stylized fact or reproduce the psychological experiments with non-incentivized individual tasks and self-assessments.

However, there is substantiated criticism concerning the applications of overconfidence within economics as being too far-reaching. Clark and Friesen (2009) point to the uncritical adoption of overconfidence by showing that the individual overconfidence for both relative and absolute self-assessments is easily avoided in a real effort task once monetary incentives and timely feedback are provided. In line with Hoelzl and Rustichini (2005), they show...
that the alleged persistence of overconfidence is mostly due to a lack of economic conditions and the ambiguity of the traits evaluated. Their doubts over the indiscriminate applicability of overconfidence as a stylized fact in economics connect well to more general considerations regarding the robustness of biases under economic conditions, with monetary incentives and learning effects correcting behavioral anomalies (Levitt and List, 2007; List and Millimet, 2008). Persistent biases subsequently result from laboratory conditions and thus hold little external validity (Zizzo, 2010).

While these results challenge the notion of a pervasive individual overconfidence in economic contexts, Burks et al. (2013) provide a different angle by explaining overconfidence as a social signaling bias, as initially suggested by Bénabou and Tirole (2002). The signaling of confidence is thus a means of individually gaining advantages by appearing more competent than others. This connects well with Johnson and Fowler’s (2011) model, which suggests that overconfidence might be an evolutionarily rational strategy. Therefore, claiming resources through biased self-assessments in individuals and populations leads to net benefits at the aggregate level. This helps to explain how a swift spread of overconfidence in social settings can be favorable, given that the potential costs of failure are smaller than the average gains of successfully claiming resources.

However, while overconfidence as a social bias is convincing, it date no experimental study applying economic conditions has investigated overconfidence in a social context. Although monetary incentives and feedback are shown to eliminate the overconfidence for individuals, implementing a social context might in turn foster overconfidence. This could explain how individuals show a reasonable confidence when deciding autonomously in experiments yet display overconfidence in numerous economic contexts, such as finance, trade or investment. Consequently, this paper aims to identify the effect of introducing a social setting on individuals’ reported levels of confidence. We replicate the basic experimental features of Clark and Friesen (2009) and implement a simple mechanism of observing other players. By running two experiments, we show that even a very basic social context triggers overconfidence. We thus find that — while realistic individually — subjects display overconfidence once the observation of others is enabled, despite monetary incentives.

2. Experimental design

Of the two experiments, experiment 1 (Individual) features individual decisions, while experiment 2 (Observation) implements a social context by randomly matching three players who observe the respective decisions of the other two players. There are no pay-off externalities, yet full information on the other players’ actions. However, observing others neither provides any additional information, nor does it alter the rational strategy. Thus, no incentives for the active signaling of competence are implemented to determine exclusively the increase in confidence due to a social context. All other parameters are identical across experiments. Subjects are asked to complete a real effort task of ten simple calculations and an ensuing self-assessment of their success, which is repeated three times. For the real effort task itself, we draw on Meub et al. (2013). Participants predict future values using a formula that comprises four determinants and remains constant throughout the experiment. The formula is $x_t = a_t + b_t - c_t + d_t$, with $x_t$ being the value participants are asked to predict in round $t$, while $a_t$, $b_t$, $c_t$, and $d_t$ are known determinant values changing in each round. $d_t$ is a random variable, which is uniformly distributed over the interval $[-15, 15]$. Its realizations are unknown to the participants and change every round. The expected value is zero; therefore, the rational strategy to maximize payoffs is given by the calculation of the respective expected values, consequently ignoring $d_t$.

Following every block of ten calculations, participants are asked to assess how many of their predictions were correct. We thus draw on Clark and Friesen (2009), who argue that this type of frequency-based method for eliciting confidence levels is more intuitively understood by participants. After every such self-assessment, feedback concerning the actual performance is provided to participants in order to enable learning effects. Participants have one minute for every calculation and self-assessment. To identify over- or underconfidence, we define a prediction as correct if it is three points over or under the correct value for the respective round. This is explained to participants using a simple example in the instructions. For the example of the third round, $E(x_t)$ equals 114, $d_t$ is $-11$ and thus the correct value is 103. All predicted values within the interval $[100, 106]$ are evaluated as being correct. Given that all values for $d_t$ are equally likely, the probability of a correct prediction is $7/31 = 0.225$ if the prediction does not deviate more than 12 points from the expected value. Assuming this is true for all ten predictions of one block, the rational self-assessment or confidence level equals two correct answers out of the ten calculations. A confidence level greater than 2 is never optimal and can unambiguously be characterized as overconfident, whereas values smaller than two indicate underconfidence.

The payoff for both experiments in each round is fifty cents minus double the absolute difference between the respective prediction and the actual value; therefore, payoffs cannot become negative. Subjects earn one euro for each correct self-assessment. Given the realizations for all determinants, following the rational strategy of predicting the expected value of $x_t$ yields on average 0.28€ per prediction in both treatments. When choosing the rational level of confidence, subjects are correct in two out of three blocks, thus earning an additional 2€. A completely rational subject would receive 10.34€, plus the 1.5€ show-up fee. For Individual (Observation), we have 2 (4) sessions with a total of 30 (90) subjects (using zTree, Fischbacher, 2007, and ORSEE, Greiner, 2004). Experiments were conducted at the Göttingen Laboratory of Behavioral Economics using a standard subject pool across all disciplines. 48% of participants in our sample were female. Including show-up fees, subjects earned 11.29€ on average.

---

4 Note that the calculations are intentionally relatively easy to complete. For instance, the calculation in the third round for the expected value of $x$ is $95 + 55 - 36 = 114$; a task that every participant should be able to complete. Additionally, subjects were able to use the Windows calculator implemented in the 2-Tree program.

5 See Cesaroni et al. (2006) for an investigation into the effects of different modes of self-assessment. They support the view that confidence interval estimations account for a large share of the observable bias and opt for frequency-based self-assessments as a more reliable measurement of overconfidence.

6 Clearly, this criterion changes for fewer optimal predictions in a given block. However, as we will show in Section 3.1, subjects manage to predict very accurately, such that the threshold of two correct answers holds as the optimal level of confidence.

7 The actual average payoff per prediction for a rational player is thus slightly lower than the expected average payoff given the determinant $d$‘s standard deviation of $8.94 = \sqrt{8(1 - 1/12)}$. Plugging the standard deviation in our payoff function yields an expected average payoff of $0.32e \approx 0.2 + (8.94)$. Instructions and screenshots are available from the authors upon request.
3. Results

3.1. Prediction accuracy

We compare our experiments in terms of the accuracy of predictions in the real effort task and the self-assessments. The performance in the real effort task is measured by the absolute deviations of predictions from the payoff maximizing expected values. We find around the same prediction accuracy in Observation compared to Individual, whereby the average absolute deviation amounts to 4.641 points and 4.697, respectively. There are learning effects, with individuals provided the opportunity to observe the actions of others tending to learn faster while starting off less accurate in the first block. Learning leads to a reduction of outlier predictions; accordingly, the standard deviation from the expected values decreases from 12.03 (16.61) in block 1 of Individual (Observation) to 8.40 (7.19) in block 2 to 6.04 (6.03) in block 3.

In sum, subjects perform rather well in our numerical predictions task; indeed, we find 40.7% (44.7%) of predictions in Individual (Observation) to be optimal and 95.4% (95.6%) in the range of potentially correct predictions opened up by the random variable $d_i$ of $[E(x_i) − 15, E(x_i) + 15]$. These small differences in prediction accuracy hold no relevance for the derivation of the optimal level of confidence. This can be seen when deriving the expected number of correct predictions conditional on the subjects’ performance, i.e. the optimal levels of confidence given the predictions made. In Individual (Observation), this optimal reported level of confidence amounts to 2.02 (2.01) in block 1, 2.19 (2.21) in block 2 and 2.21 (2.22) in block 3. Recall that the payoff-maximizing reported level of confidence for an individual is never greater than two. Thus, there should be no differences between the experiments in this regard.

However, as can be seen in Fig. 1, there is a higher level of confidence in Observation for all blocks ($z = 1.961$, $p = 0.0499$ for block 1; $z = 1.614$, $p = 0.1066$ for block 2; $z = 1.986$, $p = 0.0490$ for block 3; Mann–Whitney U-Test). Thus, the average level of confidence is around 30% higher for Observation. For blocks 2 and 3, this gap is reduced to around 20%, which is still remarkably large. We conclude that the level of confidence depends on the social context of the decision situation, which in our case is constituted by the opportunity of observing the actions of others.

Result 1. Observing other players’ decisions triggers overconfidence.

3.2. Distribution of reported confidence

When considering the distribution of the reported levels of confidence, we can show that underconfidence occurs more frequently in Individual. As Fig. 2 shows, there are almost no reported levels of confidence smaller than 2 in Observation, with only 6.3% equal to 1 or 0. In Individual, this value amounts to 17.78% overall, which is even higher for blocks 2 (23.33%) and 3 (20%). Accordingly, the distribution of reported levels of confidence is significantly different across treatments ($p = 0.014$; Fisher’s exact).

Result 2. When decisions are taken individually, over- and underconfident decisions tend to even out on average. Observing others eliminates most underconfident decisions.

3.3. Homogeneity of decisions

Finally, we are interested in the effect of the homogeneity of decisions within the three randomly matched players (“group”) on...
the level of confidence. We measure homogeneity by the sum of absolute deviations of a player’s predictions from the two observed ones.

Fig. 3 again shows the learning effects in terms of prediction accuracy, which makes predictions more homogeneous in blocks 2 and 3, as well as decreasing levels of confidence. However, there is no systematic relationship between the homogeneity of decisions (smaller deviations) and confidence. Reported levels of confidence vary arbitrarily over the deviations realized within a group.\(^{11}\) Note, however, that average deviations within groups are fairly small overall: the highest deviation of around 25 in block 3 implies a deviation of only 1.25 points per subject and prediction. Thus, within the overall rather homogeneous decisions, less deviation does not lead to higher self-assessments.

Result 3. The reported level of confidence does not depend on the homogeneity of group members’ decisions.

4. Conclusion

The economic literature on overconfidence seemingly portrays a contradiction: while the bias is shown for numerous real-world situations, it can be easily eliminated in experiments using monetary incentives and feedback, both of which doubtlessly prevail in the contexts discussed in economic studies on overconfidence (e.g. in investment or financial trading).

We resolve this ambiguity by pointing to overconfidence as a social bias. Individually, subjects display a rational level
of confidence, which supports the results of Clark and Friesen (2009) and Hoelzl and Rustichini (2005). As one of the most basic social settings, the observation of others strongly fosters overconfidence by effectively eliminating the underconfidence in one’s own performance. This result holds regardless of the actual homogeneity of the decisions within a group.

When interpreting the remarkable increase in confidence caused by the reduction of underconfident self-assessments, two complementary explanations seem plausible. Given that the actual intra-group homogeneity is high, underconfident subjects may be encouraged by the perception of group conformity. Based on the “meta-preference for conformity” (Klick and Parisi, 2008, p. 1319), subjects have been shown to become more confident when seeing their compliance with group members. Alternatively, when underconfident subjects are confronted with rather heterogeneous group decisions, they might still perceive homogeneity as being high due to the false consensus effect. This anomaly prompts individuals to falsely assume that their decisions are in line with the majority of others. The subsequent feeling of compliance is shown to boost individual confidence regardless of the actual conformity of decisions (for the seminal contribution, see Ross et al., 1977). Engelmann and Strobel (2012) confirm this anomaly under economic conditions, given that the information on the actual consensus of subjects “is not handed to them on a silver platter” (ibid. p. 679). This would also hold true in our case, as the levels of deviation within groups (i.e. the actual consensus) take effortful computing. While both approaches appear valid, a comprehensive explanation as to why underconfident subjects are encouraged by observing others should be subject to further studies.

Acknowledgments

We are grateful to Tonia Sauerwald and Asri Özgümüş for the excellent research assistance. We also thank an anonymous referee for very helpful comments.

References


