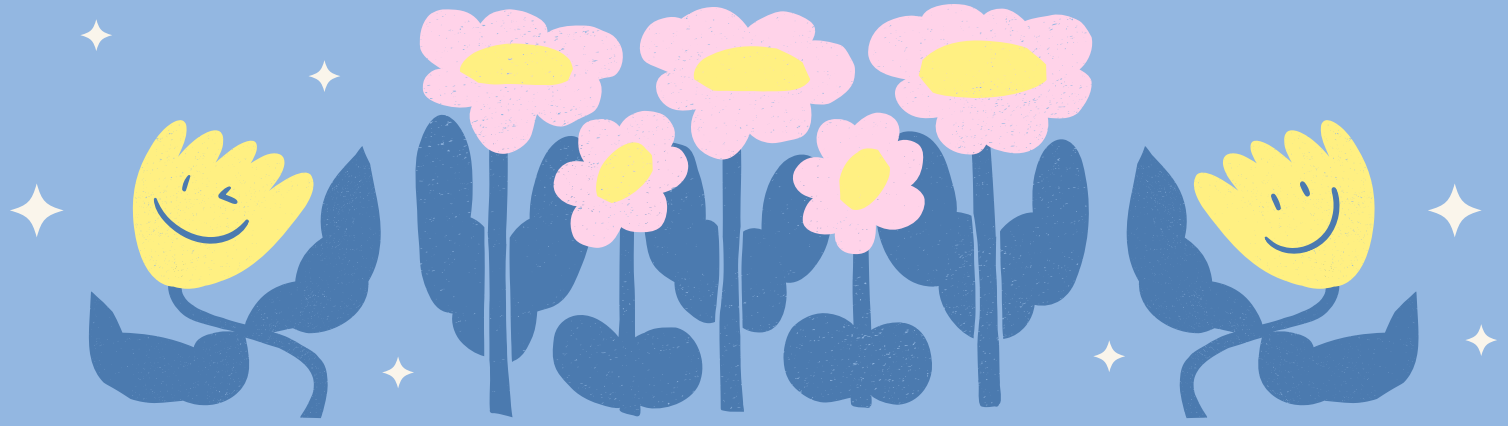
BRIBING THE SELF



URI GNEEZY , SILVIA SACCARDO , MARTA SERRA-GARCIA , ROEL VAN VELDHUIZEN

Introduction

Decision making normally rely on variety of domains.

Informational asymmetry between expert and consumer

Expert

More knowledgeable about the quality of the goods and services

Consumers

Cannot fully assess whether the advice they received was in their best interest

- Generated incentives for expert to deliver advise that is in their own best interests, not the consumer
- However, being dishonest towards consumer may lead to a conflict between advisors' material goals and their desire to maintain a self-image as honest

Objective : We study the constraints to advisors' ability to self-deceive, and test the effect of progressively reducing the scope for justifying biased advice as ethical.

Introduction

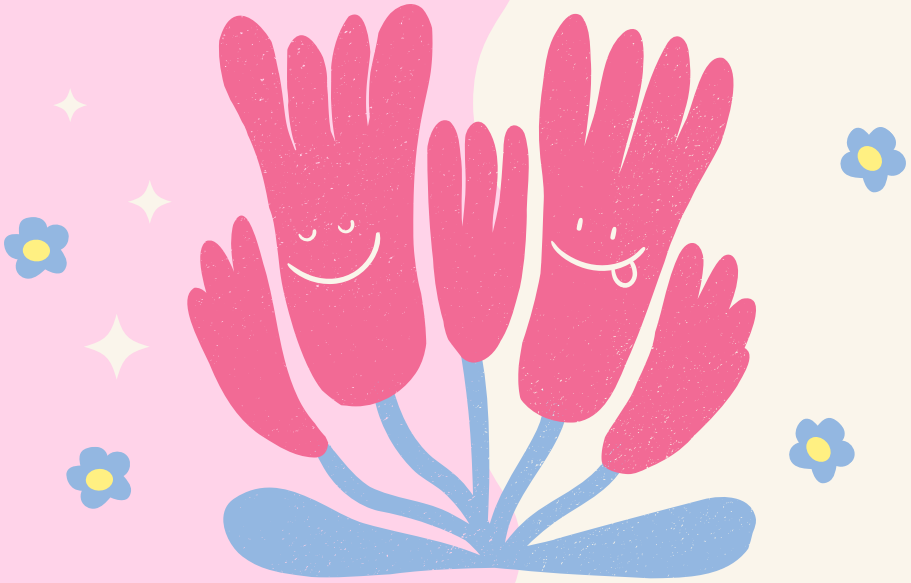
Social-image (Positive reputation)	Some advisors knowingly bias their recommendations in order to increase their own profits and conceal this behavior to preserve a **positive reputation or social-image**
Self-image	Advisors convincing themselves to believe that the **choices that benefit their financial benefit** are actually in the client's best interest, thereby preserving their positive self -image

Learning about the incentives before and after evaluating the investment options constrains advisors' ability to self-deceive

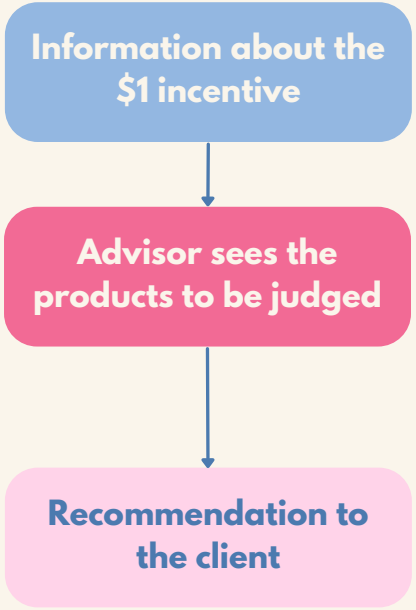
- If the advisor is **informed about the incentives before having a chance of evaluating the investments**, she might inadvertently distort her beliefs and recommendation what she is incentivized to to be the one the client would prefer.

- But, if she instead **privately evaluates the investments before learning about her incentives**, engaging in self-deception becomes harder.

Experimental Design



The Advice Game



Control Treatment

Before Treatment

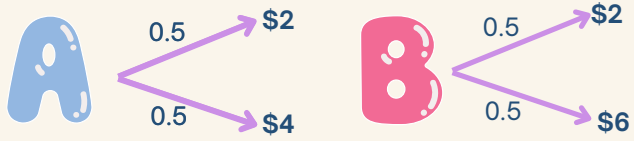
After Treatment

The Advice Game

RiskRuturn Experiment

The expected payoff of B ($4) is higher than that of A ($3).
However, B has a higher variance.

Dominance Experiment



The expected payoff of B unchanged ($4), but B now yields the same payoff as investment A in the low state, and a higher payoff in the high state.

ObviousDominance Experiment

The expected payoff of B yields a strictly higher than A in both state of nature.

Hypotheses

- The experimental design allows us to test whether advisors' ability to self-deceive increases their propensity to recommend the incentivized investment, and to distinguish the self-deception mechanism from other potential drivers of advice.

If advisors are purely guided by their own self-interest, they should recommend A

Whenever the advisors have
- **Incentive**
- **Timing to learn about their own incentive**

→

those things will Not Affect ❌ recommendations

In all experiments, the fraction of A recommendations should be the same in the Before and After treatments, and higher than in Control.

Hypotheses

If advisors care about their self-image

Advisors who do not engage in self-deception

Believe client would prefer A

Advisors unable to recommend A
while preserving a positive self-image

Advisors who care about their self-deception

Believe client would prefer A

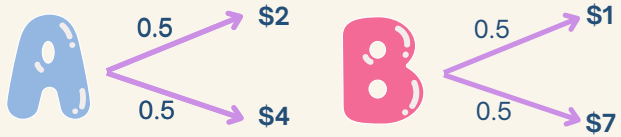
Advisors recommend A

- **The rate of A recommendations in the presence of incentives should be similar to the rate of recommendations in the Control treatment if advisors not engage in self-deception**
- **By contrast, if advisors who care about their self-image engage in self-deception, the timing of information may have an important effect on recommendations.**

Hypothesis

- In the Risk Return,

  Reasons that could make A the preferred option



The advisor could believe the client is risk averse and thus favor A

A may be seen as the fair recommendation by both parties, since the advisor receives a $1 commission for recommending investment A

Hypothesis 1 : In the Risk Return experiment, advisors recommend A more frequently in Before than in After treatment

Hypothesis

- In the Dominance experiment,

- Investment B yields weakly higher payoffs than A in both states, hence the risk in B is no longer a reason to favor A.
- However, in this experiment the client's costs (in expectation) of following an A recommendation are the same as in the RiskReturn experiment.
- This might provide advisors with an excuse for recommending the incentivized investment.

Hypothesis 2A : If advisors self-deceive about the client's risk preference, in the Dominance experiment advisors recommend A equally frequently in the Before as in the After treatment

Hypothesis 2B : If advisors self-deceive about the client's fairness concerns, in the Dominance experiment advisors recommend A more frequently in the Before than in the After treatment

Hypothesis

- In the ObviousDominance experiment,

- The cost of recommending *A* to the client increases substantially. This greatly diminishes the scope for arguing that the client would prefer *A*.
- We hence hypothesize that there is no room for self- deception in this experiment.
- This should be reflected in the gap between the rate of *A* recommendations between the Before and After treatment, which should decrease.

Hypothesis 3 : In the ObviousDominance experiment, advisors recommend A equally frequently in the Before as in the After treatment.

Hypothesis

Finally, across all three experiments, if there is no scope for self-deception in the After treatment, we would expect no difference in recommendations between After and Control. By contrast, if advisors still recommend the incentivized option without self-deception because they mainly care about their own monetary incentives, we would observe a difference in recommendations between the After and Control treatments.

Procedures

University of California San Diego in the Spring and Fall of 2015

- The three experiments were conducted sequentially and randomized assignment to each treatment occurred at the advisor level.
- Clients were recruited , and participated in separate sessions.

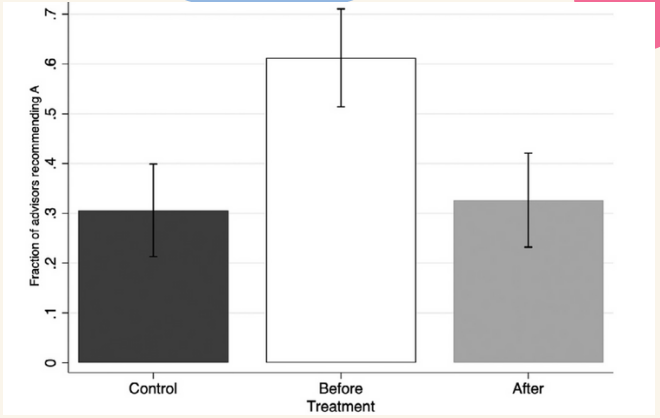
The procedures were identical for all experiments. The instructions were presented to the advisor on four separate pages on the computer screen

- First screen, we introduced the advisor to a study on economic decision making.
- Second screen, the advisor was informed her role in the experiment and **given a fixed payment of $1 for her participation**. She was told to recommend one of two investments (A and B) to another participant, but do not know what A and B exactly were. She also learned that the other participant received no information about A or B except her recommendation.
  - In the **Before treatment**, $1 commission has been informed
- Third screen, advisors were presented with the details of A and B.
- Fourth screen, the assistant give her a piece of paper on which she could write her recommendation.

- To receiving information about the lotteries, the advisor was asked to think about her recommendation and continue to the forth screen once she was ready to provide it.
- All participants must move to the next screen before writing anything on paper.
  - In the **After treatment**, the advisor learned about the $1 commission, and was then asked to provide her recommendation both on paper and on screen.
- To introduce only one change across treatments, the information on the commission was also presented on the fifth screen in the **Before treatment.**
- This procedure had the advisor send a message in her own handwriting direct electronic record of recommendations.

Result : The RiskReturn

- The percentage of advisors recommending A in the **Control treatment is 30.6%.**
- When information about the incentive tied to *A* is provided before reading about *A*, advisors are significantly more likely to recommend it
- **Before treatment (self-deception is easy)** : They recommend *A* in 61.2%
- **After treatment (self-deception is harder)** : The rate drops to 32.7%



Fraction of A recommendations in the RiskReturn experiment

Table 1
Treatment Effects on the Likelihood that *A* is Recommended across the Three Experiments.

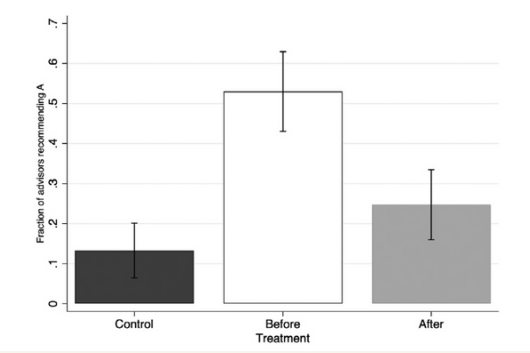
	(1)	(2)	(3)	(4)
	Pr(A is recommended)			
	RiskReturn	Dominance	ObviousDominance	All
Before	0.286***	0.283***	0.011	0.286***
	(0.069)	(0.067)	(0.063)	(0.069)
Control	−0.020	−0.115**	−0.133**	−0.020
	(0.067)	(0.056)	(0.057)	(0.067)

- In the **RiskReturn experiment**, the recommending A increases by 28.6 percentage points in the Before treatment, relative to the After treatment.
- The **difference** between the **Before and Control** treatments is also significant.

The results of the RiskReturn experiment could also be consistent with two alternative explanations.

1. Participants in the Before treatment avoid evaluation altogether and simply recommend the incentivized investment (Signal the better product)
2. The smaller bias could also result from preferences for consistency

Result : The Dominance



Fraction of A recommendations in the Dominance experiment

- **Control treatment**, participants recommended A in **13.3%** of the cases.
- **Before treatment:** Advisors recommended A at a higher rate, in **53%**
- **After treatment:** Advisors recommended A in **24.7%**
- The advisors are **less likely to recommend A in the Control treatment than in the After treatment** in the Dominance experiment

- The distortion in recommendations in the Before treatment is **similar to the RiskReturn experiment**
- So we can compare the effect of the Before and Control treatments on recommendations in the **Dominance and RiskReturn experiment**.

Table 1
Treatment Effects on the Likelihood that *A* is Recommended across the Three Experiments.

	(1)	(2)	(3)	(4)
	Pr(A is recommended)			
	RiskReturn	Dominance	ObviousDominance	All
Before	0.286***	0.283***	0.011	0.286***
	(0.069)	(0.067)	(0.063)	(0.069)
Control	−0.020	−0.115**	−0.133**	−0.020
	(0.067)	(0.056)	(0.057)	(0.067)

Dominance	−0.079
	(0.065)
Before X Dominance	−0.003
	(0.096)
Control X Dominance	−0.094
	(0.087)

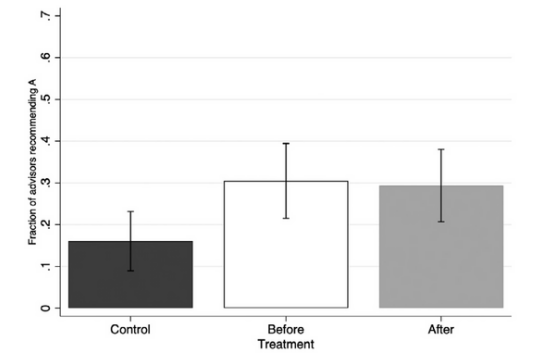
- **If there is less scope for self-deception , the coefficient for the Before treatment should be significantly smaller**

>>> We find no evidence in this case.<<<

  - The point estimate for the coefficient Before treatment X Dominance is close to zero and not significant.
  - These results therefore suggest that advisors are able to **engage in self-deception** even when they have only “minimal” reasons to recommend A

Result : The ObviousDominance

- Investment A is recommended 16% of the time in the Control treatment.
- Before treatment recommendation (30.5%) is **no longer significantly** different from the After treatment **(the timing of information about the incentive no longer affects)**
- Compared to the Control treatment, A is recommended more frequently in the **Before and After treatment** .



Fraction of A recommendations in the ObviousDominance experiment

Table 1
Treatment Effects on the Likelihood that *A* is Recommended across the Three Experiments.

	(1)	(2)	(3)	(4)
	Pr(A is recommended)			
	RiskReturn	Dominance	ObviousDominance	All
Before	0.286***	0.283***	0.011	0.286***
	(0.069)	(0.067)	(0.063)	(0.069)
Control	−0.020	−0.115**	−0.133**	−0.020
	(0.067)	(0.056)	(0.057)	(0.067)

- The Before-After effect is **significantly weaker in the ObviousDominance** experiment than in the RiskReturn and the Dominance experiment.

- The results of the **ObviousDominance** experiment rule out the alternative mechanisms we discussed above.
  - Specifically, both consistency and evaluation avoidance would predict a treatment difference between the **Before and After treatment**
- **The significantly difference** between the **Before and After treatment in the ObviousDominance** experiment, **suggesting that these alternative explanations cannot explain the larger difference observed in the RiskReturn and Dominance experiment.**

Mechanisms of self-deception

Mechanisms of self-deception: belief distortion and choice for self

- **Since we did not measure beliefs, therefore these three laboratory experiments provide indirect evidence of self-deception**

  - So, additional online experiments was launched. By follow the design of the laboratory experiments and collect two additional measures:

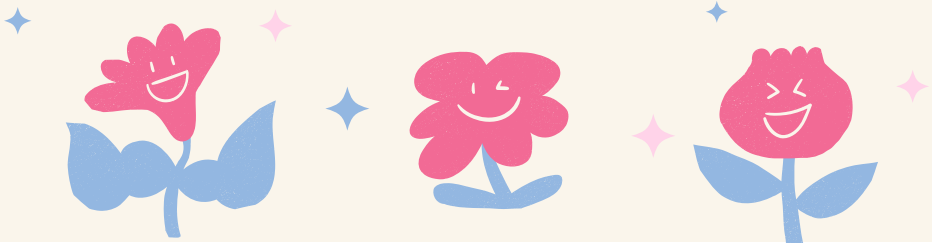
    1. Advisors' beliefs about the investment recommendation the client would prefer
    2. Advisors' choices for themselves between the two investments.

Procedures

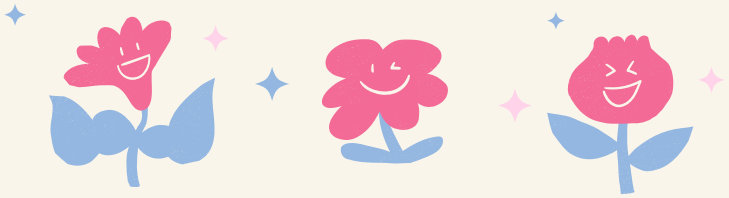
We conducted three experiments (RiskReturn, Dominance and ObviousDominance) on Amazon Mechanical Turk in the Spring of 2019. Focusing on the Before and After treatments only (excluding Control).

We recruited participants to a study for a fixed payment of $0.25.

- In all experiments, investment A was a 50-50 lottery between $0.50 and $1.
- Investment B was a 50-50 lottery
  - In the **RiskReturn** experiment, between $0.25 and $1.75
  - In the **Dominance** experiment, between $0.50 and $1.50
  - in the **ObviousDominance** experiment, between $1.25 and $1.75
- The advisor's commission for recommending A was always $0.25.
- And in this online experiments we implemented a 1 to 1 matching between the advisor and the client.



Procedures



- At the beginning of the experiment, the advisor learned about her role in the experiment and her task of recommending one of two investment options (A and B) to a client.

Before Treatment

- the advisors also received information about her own commission.

- Before asking the advisor to make a recommendation, we elicited the advisor's beliefs.
  - Advisors indicated which recommendation, A or B, they thought the client would prefer.
  - Measuring beliefs before asking subjects to make their recommendation could perhaps distort their recommendation.
- After the belief elicitation task, the advisor moved to the final screen, which asked her to provide her recommendation to the client.

After Treatment

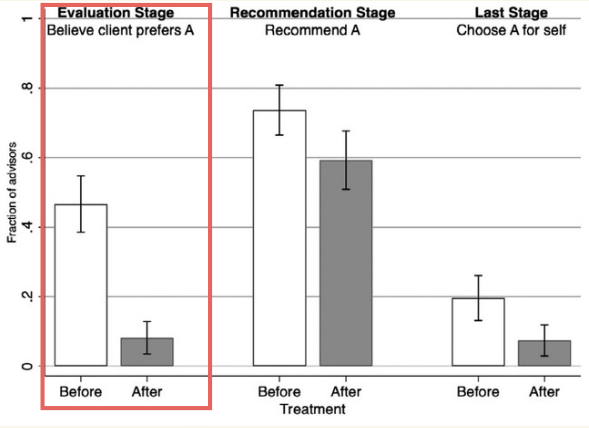
- the advisors learned about their commission prior to being asked to provide their recommendation.

- At the end of the experiment, we asked advisors to choose between the two investments, A and B, as a reward for themselves.
  - Measuring advisors' choices for themselves follows Chen and Gesche (2018), who show that advisors are more likely to choose an investment product for themselves and for a client after having been previously incentivized to recommended it to another client.

✦ The result of this experiment is a total of 88.7% of clients followed the advisors recommendation (90.4% for product A and 85.4% for product B).

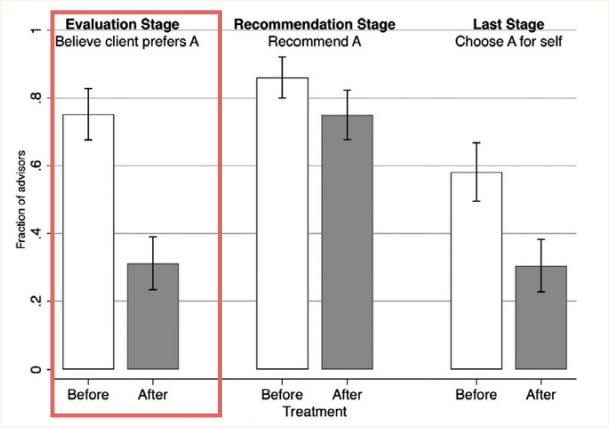
Results (Beliefs)

RiskReturn experiment



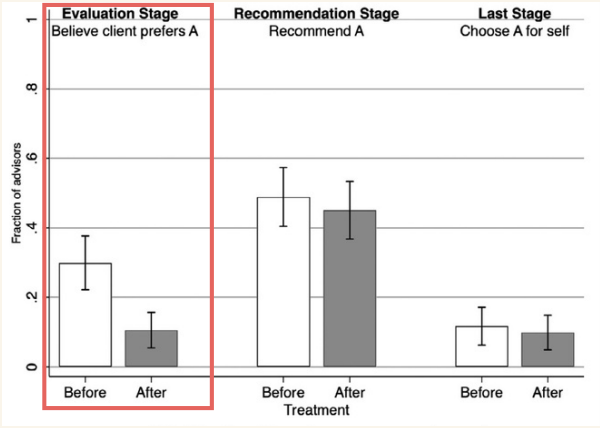
- 75.2% of advisors consider *A* the client's preferred recommendation in the **Before treatment**.
- By contrast, this fraction is only 31.4% after advisors evaluate the lotteries in the **After treatment**.

Dominance experiment



- 46.6% of participants consider *A* the recommendation
- the client would prefer whereas this fraction is only 8.1% in the **After treatment**.

ObviousDominance experiment



- the difference in beliefs becomes smaller
- 29.9% of participants consider *A* the investment recommendation the client would prefer
- this fraction is 10.6% in the **After treatment**.

Results (Beliefs)

Table 2
Recommendations, beliefs and choices for self in the online experiments.

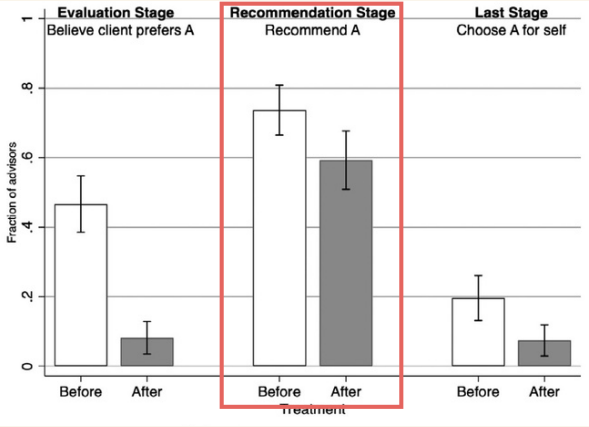
	(1) RiskReturn	(2) Dominance	(3) ObviousDominance
Panel A: Pr(*A* is recommended)			
Before	0.110**	0.144**	0.038
	(0.048)	(0.056)	(0.060)
Constant	0.750***	0.593***	0.451***
	(0.037)	(0.042)	(0.042)
Observations	269	283	279
R-squared	0.019	0.023	0.001
Difference-in-difference			
Before relative to RiskReturn (p-value)	–	0.650	0.347
Panel B: Pr(Advisor believes client prefers *A*)			
Before	0.438***	0.385***	0.194***
	(0.055)	(0.047)	(0.047)
Constant	0.314***	0.081**	0.106***
	(0.039)	(0.024)	(0.026)
Observations	269	283	279
R-squared	0.192	0.182	0.058
Difference-in-difference			
Before relative to RiskReturn (p-value)	–	0.446	0.001
Panel C: Pr(Advisor chooses *A* for herself)			
Before	0.274***	0.122***	0.018
	(0.059)	(0.040)	(0.037)
Constant	0.307***	0.074***	0.099***
	(0.039)	(0.023)	(0.025)
Observations	269	283	279
R-squared	0.076	0.031	0.001
Difference-in-difference			
Before relative to RiskReturn (p-value)	–	0.029	0.000

Panel B of Table 2 shows that the gap between the **Before and After treatment** which becomes significantly smaller in the **ObviousDominance** experiment.

Hence, learning about one's incentive prior to evaluating the investments distorts advisors' beliefs toward the investment that allows them to earn a commission. In line with the idea that reducing ambiguity over the better recommendation limits self-deception, advisors are progressively less likely to distort their beliefs when we move from the RiskReturn to the Dominance, and ObviousDominance Experiment.

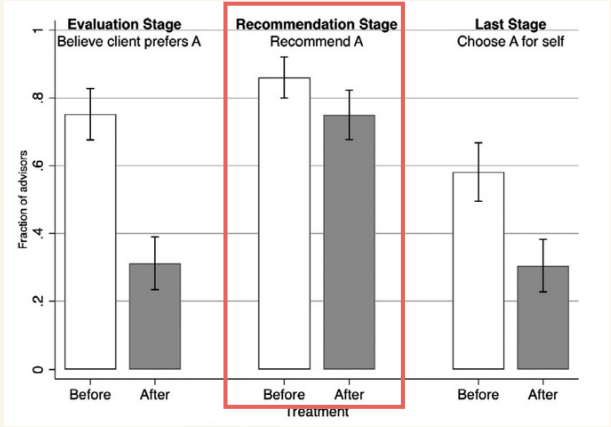
Results (Recommendations)

RiskReturn experiment



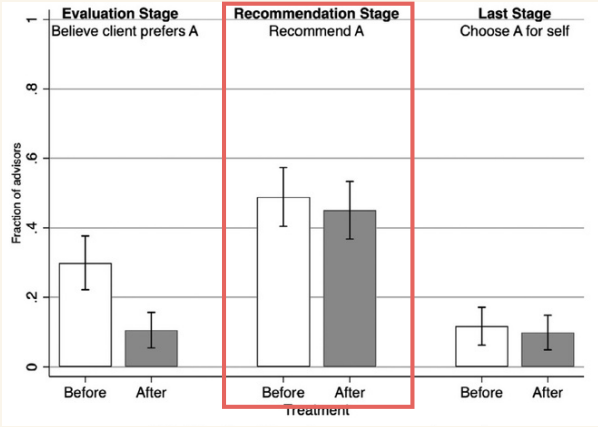
- 86.0% of advisors recommend *A* when they learn about their own incentives before evaluating the two investment options
- 75.0% of advisors recommend *A* when they learn about their incentives afterwards

Dominance experiment



- 73.6% of advisors recommend *A* in the **Before treatment**
- this fraction is only 59.3% in the **After treatment**

ObviousDominance experiment



- there is no significant treatment difference in *A* recommendations, with 48.9% of the advisors recommending *A* in **Before treatment**
- 45.1% of advisors recommending *A* in **After treatment**

Results (Recommendations)

Table 2
Recommendations, beliefs and choices for self in the online experiments.

	(1) RiskReturn	(2) Dominance	(3) ObviousDominance
Panel A: Pr(*A* is recommended)			
Before	0.110**	0.144**	0.038
	(0.048)	(0.056)	(0.060)
Constant	0.750***	0.593***	0.451***
	(0.037)	(0.042)	(0.042)
Observations	269	283	279
R-squared	0.019	0.023	0.001
Difference-in-difference			
Before relative to RiskReturn (p-value)	–	0.650	0.347
Panel B: Pr(Advisor believes client prefers *A*)			
Before	0.438***	0.385***	0.194***
	(0.055)	(0.047)	(0.047)
Constant	0.314***	0.081**	0.106***
	(0.039)	(0.024)	(0.026)
Observations	269	283	279
R-squared	0.192	0.182	0.058
Difference-in-difference			
Before relative to RiskReturn (p-value)	–	0.446	0.001
Panel C: Pr(Advisor chooses *A* for herself)			
Before	0.274***	0.122***	0.018
	(0.059)	(0.040)	(0.037)
Constant	0.307***	0.074***	0.099***
	(0.039)	(0.023)	(0.025)
Observations	269	283	279
R-squared	0.076	0.031	0.001
Difference-in-difference			
Before relative to RiskReturn (p-value)	–	0.029	0.000

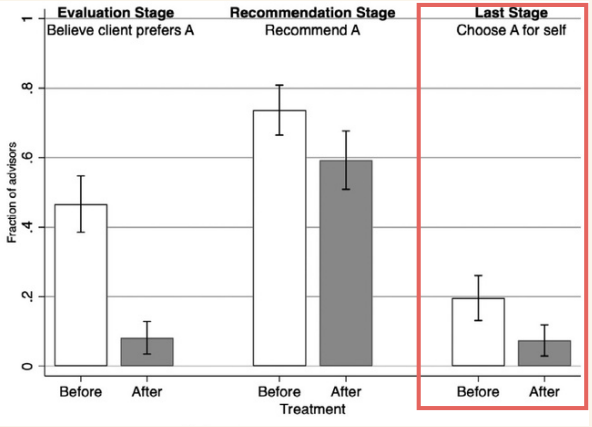
Panel A of Table 2 confirms these results.

- In the **RiskReturn** experiment, the likelihood of recommending *A* increases by 11% in the **Before treatment**.
- In the **Dominance** experiment it increases by 14%.
- In the **ObviousDominance** experiment the difference in *A* recommendations is small and no longer significant.
- However, given the smaller Before-After gap in the **RiskReturn** experiment, we are unable to detect a significant decrease in the size of the gap between the **Before and After treatments**.

Learning about the incentive prior to evaluating the two investments only led to a significantly higher fraction of *A* recommendations in presence of some scope for self-deception.

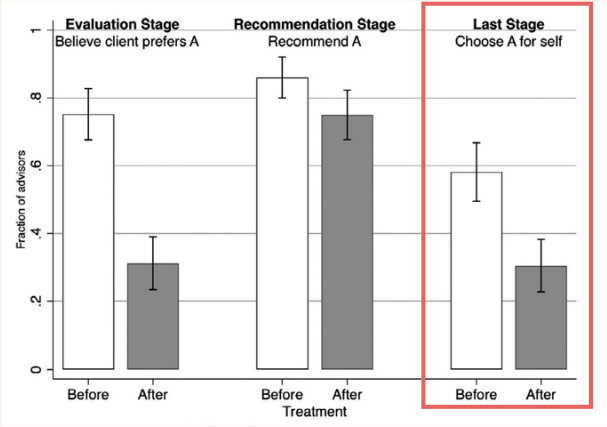
Results (Choice of Self)

RiskReturn experiment



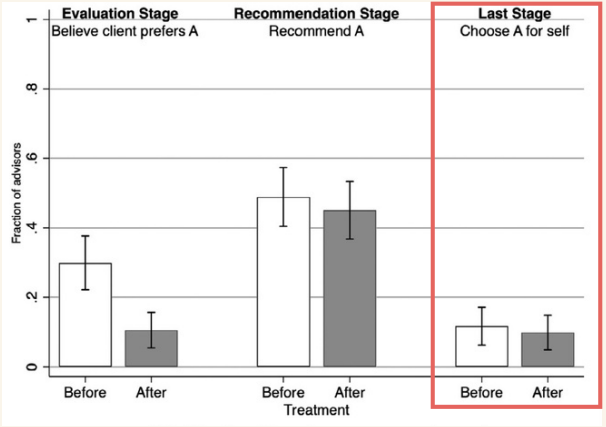
- we find that 58.1% of the advisors choose investment *A* for themselves
- This fraction is only 30.7% in the **After treatment**.

Dominance experiment



- the fraction of advisors choosing *A* in the **Before** treatment is smaller, 19.6%.
- a gap between the **Before** and **After** treatment persists: only 7.4% of advisors chooses *A* in the **After treatment**.

ObviousDominance experiment



- no gap in *A* choices between treatments: 11.7% of advisors choose it in Before and 9.9% of advisors choose it in **After treatment.**

Table 2
Recommendations, beliefs and choices for self in the online experiments.

	(1) RiskReturn	(2) Dominance	(3) ObviousDominance
Panel A: Pr(*A* is recommended)			
Before	0.110**	0.144**	0.038
	(0.048)	(0.056)	(0.060)
Constant	0.750***	0.593***	0.451***
	(0.037)	(0.042)	(0.042)
Observations	269	283	279
R-squared	0.019	0.023	0.001
Difference-in-difference			
Before relative to RiskReturn (p-value)	–	0.650	0.347
Panel B: Pr(Advisor believes client prefers *A*)			
Before	0.438***	0.385***	0.194***
	(0.055)	(0.047)	(0.047)
Constant	0.314***	0.081**	0.106***
	(0.039)	(0.024)	(0.026)
Observations	269	283	279
R-squared	0.192	0.182	0.058
Difference-in-difference			
Before relative to RiskReturn (p-value)	–	0.446	0.001
Panel C: Pr(Advisor chooses *A* for herself)			
Before	0.274***	0.122***	0.018
	(0.059)	(0.040)	(0.037)
Constant	0.307***	0.074***	0.099***
	(0.039)	(0.023)	(0.025)
Observations	269	283	279
R-squared	0.076	0.031	0.001
Difference-in-difference			
Before relative to RiskReturn (p-value)	–	0.029	0.000

Panel C of Table 2 confirms these results using regressions.

Overall, the results from the online experiment confirm the results of the lab and provide additional evidence for the self-deception mechanism. As in the laboratory, we find that a minimal justification is enough to enable advisors to recommend A.

Discussion

How can the findings in our experiments in the laboratory and online be reconciled within one frame?

In Before treatment, advisors have information about their incentives from the start, and can form motivated beliefs about whether the client would prefer an A recommendation.

Advisors derive self-image utility from providing advice that they believe is in line with what the client would prefer.

In After treatment, advisor are initially unaware of the commission, and most advisors come at the cost of losing their initial self-image, which decreases the share of advisors recommending A.

What determines the cost of self deception?

The extent of self-deception is the same in the RiskReturn and Dominance experiments, although fewer individuals actually believe A to be the client's preferred recommendation.

Advisors primarily care about the expected payoffs of the client and that the costs of self-deception can be modeled as a constant.

Conclusion

Our finding illustrate how people have psychological costs associated with distorting advice. Creating procedures that reinforce the role of self-image costs can reduce unethical behavior by ethical-but-biased individuals.

- Physicians may believe incentives do not influence their judgement
  - These beliefs allow them to receive incentives while maintaining their self-image as unbiased professional.
  - The evidence suggest these expert are **wrong**, as incentives do distort their judgment

- Belief - based Utility
  - People have implications for the literature on preferences for truth-telling.
  - People are averse to lying and cheating.
  - People do not lie much even in presence of strong incentives.

- Expert advice is biased by incentives are plentiful and have a huge impact on efficiency and fairness.
  - To limit such incentives when possible
  - However, it may be hard to implement due to pushback from lobbyists and high cost of monitoring and enforcement
  - Additional approaches may reduce the effectiveness of incentives in distorting judgement
    i. Timing of decisions by having experts first evaluate the options and only then receive information about the incentives
    ii. Providing experts with as much information about the client's preferences as possible