When Good Advice is Ignored: The Role of Envy and Stubbornness*

David Ronayne[†] and Daniel Sgroi[‡]

17 January 2018

Abstract

We present results from an experiment involving 1,500 participants on whether, when and why good advice is ignored, focusing on envy and stubbornness. Participants performance in skill-based and luck-based tasks generated a probability of winning a bonus. About a quarter ignored advice that would have increased their chance of winning. Good advice was followed less often when the adviser was relatively highly remunerated or the task was skill-based. More envious advisees took good advice more often in the skill-based task, but higher adviser remuneration significantly reduced this effect. Susceptibility to the sunk cost fallacy reduced the uptake of good advice. *JEL: C91, C99, D91*

Keywords: advice, skill, remuneration, envy, sunk cost fallacy

*The authors are grateful to Nuffield College for funding this research and to Johannes Abeler, Gordon Brown, Vince Crawford, Andrew Dilnot, Andrea Isoni, Nattavudh Powdthavee, Margaret Meyer, Neil Stewart and Bill Zame for helpful comments. The experiment described in this paper was pre-registered at the AEA RCT Registry on 28 June 2017: `https://www.socialscienceregistry.org/trials/2022`. Ethical approval for the experimental work was obtained from the Department Research Ethics Committee on 20 January 2017 under reference ECONCIA17-18-001 and from the Centre for Experimental Social Sciences on 14 February 2017 under reference ETH-170118191-1.

[†]Economics Dept. and Nuffield College, University of Oxford, david.ronayne@economics.ox.ac.uk

[‡]Economics Dept. and CAGE, University of Warwick and Nuffield College, University of Oxford, daniel.sgroi@warwick.ac.uk

"America has always been a country of amateurs where the professional, that is to say, the man who claims authority as a member of an elite which knows the law in some field or other, is an object of distrust and resentment"

W. H. Auden, Faber Book of Modern American Verse, 1956.

"People in this country have had enough of experts"

Michael Gove MP, UK Justice Secretary, 2016.

1 Introduction

Recent years have certainly been tumultuous for experts. Both the debate in the UK surrounding "Brexit" (the national referendum on Britain's EU membership held in 2016) and the 2016 US presidential election explicitly indicated increasing distrust of expert opinion among voters. Reflecting the same trend, the US President, Donald Trump, reportedly refuses to read intelligence briefings arguing that as a "smart person" he finds them unnecessary.[1] One of the UK's most respected scientists Stephen Hawking described Brexit as "the moment when the forgotten spoke, finding their voices to reject the advice and guidance of experts and the elite everywhere."[2] British Member of Parliament Michael Gove stressed that people did not think well of experts in the run-up to the referendum (see the quote above). More generally, the marketing firm Edelman has run large-scale international surveys for 17 years and recently reported distrust of a variety of different expert bodies across the world.[3] Their survey of over 33,000 individuals across 28 countries documents that 60% of respondents found peers, defined as "people like you", to be as useful as academic or technical experts. However, while perceptions of experts may have worsened recently,[4] signs that expert advice may often go ignored stretch back to well before the present day. For example, W. H. Auden, the noted British poet who moved to the USA in 1939 seemed to believe that the distrust of experts is something of a national characteristic of the USA (see the quote above).

There are many reasons why advice may be ignored. For example, there is nearly always uncertainty over whether advice from a given source is good. We define advice from one entity (an "adviser") to another (an "advisee") as "good" if accepting the advice improves the expected utility of the advisee, or "bad" if it reduces it. In reality, if advice is perceived as bad e.g., as a result of coming from a biased or corrupt source, it can of

[1] Newsweek, 30th May 2017, `http://www.newsweek.com/donald-trump-does-not-read-white-house-top-secret-intelligence-briefings-big-617515`

[2] The Express, 3rd December 2016, `http://www.express.co.uk/news/politics/739320/Brexit-Stephen-Hawking-Remain-MPs-subvert-EU-referendum-result`.

[3] The Atlantic, January 20th 2017, `https://www.theatlantic.com/international/archive/2017/01/trump-edelman-trust-crisis/513350/`

[4] Edelman's survey recorded falls in trust for NGOs, government, business and media in 2017; something they have not recorded before in the history of their survey. See `https://www.edelman.com/executive-summary/`

course be rational to ignore it (see Cain et al., 2005). In contrast, if *good* advice is ignored, it is a concern for society at large as well as for experts themselves in many disciplines, including economics. In this paper, we study whether, when and why individuals may irrationally ignore *good* advice.

In order to address these questions, we conducted a pre-registered online experiment.[5] We recruited 1,578 participants through Amazon's Mechanical Turk, a relatively well-explored (see e.g., Paolacci and Chandler, 2014; Paolacci et al., 2010) and commonly-used participant pool in social science, including economics (e.g., DellaVigna and Pope, 2017; Kuziemko et al., 2015). We designed our experiment to unearth the determinants of when we may expect good advice to be ignored. In addition, we used relevant psychological scales to form an underlying explanation. Running an experiment allowed us to control for some of the complicating issues that could arise with field data including any perverse or corrupt incentives, uncertainty over the quality and implications of advice, dynamic issues such as history or reputation, and any confounding issues inherent to a particular context e.g., politics, science, medicine, finance, religion etc. At the same time, we aimed to construct an experiment possessing the main characteristics of advice-taking scenarios: i) in the absence of any advice, an individual will achieve some level of expected utility as a result of their own reasoning, position, or opinion; ii) in the presence of advice, ignoring it leaves this expected utility unchanged whereas accepting good (bad) advice raises (lowers) the individual's expected utility. In our experiment, individuals completed tasks where their responses generated their probability of winning a bonus payment. They were then presented with the opportunity to change their probability of winning a bonus payment to be the same as another participant's (an adviser's) by using the other's responses instead of their own. Therefore, the advice offered to a participant was good (bad) when the adviser's probability of winning the bonus was higher (lower) than the participant's.[6] How accepting the advice would alter their expected payoff was explained to participants: raising it in the case of good advice; lowering it in the case of bad advice. Our first main finding is that in our experiment, while $< 3\%$ of bad advice was accepted, good advice was ignored 25% of the time.

Our conceptualization of advice-taking contexts suggests some behavioral explanations for such findings. We suppose that every individual's position or opinion about the best course of action is a probability distribution over payoff-relevant states of the world. In turn, this distribution generates a level of expected utility. In addition, some individuals (advisers) are in the position of being able to offer their opinion (probability distribution) to others (advisees). When deciding whether to take the advice, a rational individual compares the expected utility achieved under the adviser's distribution to that generated from their own distribution; where the former is higher (lower), the advice is "good"

[5]See the AEA RCT Registry entry: `https://www.socialscienceregistry.org/trials/2022`

[6]Note that neutral language was used in place of words such as "advice" or "adviser".

("bad") and is accepted (ignored).[7] Through this lens, we argue that there is the scope for particular behavioral forces to systematically affect the decision to accept good advice. Crucially, the two distributions compared by an advisee are inextricably linked to their sources: one is the advisee's and one is the adviser's. This is a fundamental feature of contexts where advice is offered and without this, our experiment would simplify to one of choices over lotteries. Specifically, in the case where good advice is offered, an advisee compares their inferior distribution to the adviser's superior distribution. Therefore, we hypothesize that there may be a role for the fundamental human trait of envy to arise during an advisee's comparison of their distribution and perhaps other attributes of themselves, against those of the adviser. More generally, envy is a particularly pertinent feature of the human experience. The philosopher Bertrand Russell argued that envy was quite possibly the root cause of unhappiness in Western society, but at the same time can be harnessed as a force for good (see e.g., Russell, 1930). However, envy has received surprisingly little attention in economics as a possible determinant of behavior (we discuss some exceptions in the next section). For our study, we measured envy by requiring participants to complete the eight-question "dispositional envy scale" of Smith et al. (1999). To our knowledge, we are the first to employ such a scale in economics.

In our experiment, we found that good advice was ignored on average about a quarter of the time. Because our design produced variation in the gain in expected payoff from accepting advice, we could see how the propensity to accept good advice varied with the value of advice. Here we found the more valuable the advice, the more likely participants were to accept it. This suggests that participants traded-off rationality with other factors.

To better understand when good advice is ignored, we conducted novel treatments both within-subject and between-subject. Our within-subject treatment was designed to test whether the nature of the superiority of the adviser played a role. All participants, including those who were chosen as advisers, each completed a luck-based task and a skill-based task to generate their respective probabilities of winning a bonus payment. We found that good advice was accepted more often when the adviser's superior position had been generated through better luck rather than higher skill. Furthermore, when the adviser was more skillful, we found that envy was a significant determinant of whether good advice was followed (but not when the adviser was luckier). Here, we found a positive role of envy: a higher dispositional level of envy was positively associated with the acceptance rate of good advice.

Our between-subject treatment was designed to test whether attributes of the adviser affected the frequency with which good advice was accepted. Specifically, we focused on income inequality by varying the remuneration of the adviser. Survey evidence suggests that this attribute in particular may have an impact on whether advice is followed and

[7] In our experiment, the distributions employed were particularly simple: there were only two payoff-relevant states of the world, and the two distributions available were ordered by stochastic dominance.

has not, to our knowledge, been tested before in a controlled environment. The recent and indicative survey by Edelman suggests a link between resources and expert credibility, reporting that 75% of people surveyed feel the "system" is biased towards the rich. However, this measure does not allow us to know whether that sentiment has affected the decision to accept advice, and if it did, whether the advice was rationally ignored because it was perceived to be bad, or ignored for other reasons. To make further progress, we put some questions to a separate panel of 3,096 voters in the UK's referendum on membership of the European Union.[8] Table 1 presents the questions along with output from a simple linear regression of the responses. The estimates suggest that there may be a negative effect of expert remuneration on the decision to follow advice even when some of the reasons why advice may be rationally ignored are taken into account.

Table 1: Results from a Brexit survey

y = The advice of experts influenced my decision about how to vote.	
Experts earn too much money.	-0.107 (0.028)
I feel that expert advice was generally objective and unbiased.	0.436 (0.023)
Too often, experts have their own agenda.	-0.177 (0.034)
Observations	3,096

OLS regression. A constant was included but not reported. Robust standard errors are shown in parentheses. All four statements (including the dependent variable) were measured on a 0-100 scale, where 0 = "Completely disagree" and 100 = "Completely agree".

In our control condition, the participants who were chosen to be advisers were paid \$0.50 (the same bonus that was available to the advisee participants). In our first between-subject treatment, advisers were instead paid \$100. Overall, this high remuneration treatment reduced the propensity to accept good advice by approximately 5.5 percentage points. Moreover, when the adviser had achieved their status via skill, a negative effect of envy emerged: while in the control condition, a higher level of dispositional envy was associated with a higher propensity to accept advice, in the high-remuneration treatment, the effect of envy was significantly reduced, becoming insignificantly different from zero. Furthermore, the positive effect of envy (and hence the effect of high-remuneration to undo it) was large: we predict that the least and most envious in our sample follow good advice about 71% and 93% of the time respectively, a difference of 22 percentage points. In our second between-subject treatment, advisers were remunerated as in the control condition with the addition that the advisee could directly affect the adviser's payoff: the adviser was paid an additional \$0.50 for each participant that accepted their advice.

[8]Data was collected in November 2017. The questions were included in larger survey run by the Centre for Social Investigation at Nuffield College, University of Oxford.

In contrast to the first remuneration treatment, we found no evidence of an effect here. Together, the results suggest that the decision to ignore good advice is not affected by a desire to affect the other individual's pay, but rather by the scale of the remuneration.

Our results suggest that envy may play a major and varied role in the decision of whether to accept good advice. Naturally, we do not claim that envy is the sole determinant of such behavior. Indeed, our conceptualization of advice-taking scenarios suggests that a second major determinant may be an individual's susceptibility to the sunk-cost fallacy. In the real world (and our experiment) individuals spend resources e.g., time, energy etc., in forming their positions on issues. When presented with a superior viewpoint, a rational individual would rather adopt it than ignore it. However, where individuals suffer from the sunk-cost fallacy, there may be some resistance to moving away from their existing position. In order to measure this potential effect we built the first (to our knowledge) scale designed to measure susceptibility to the sunk cost fallacy, inspired by the work of Thaler (1999) and Arkes and Blumer (1985). We found that our measure of susceptibility to the fallacy was robustly negatively associated with the propensity to take good advice. The measure had a large effect: we predict the least susceptible in our sample to follow good advice 21 and 16 percentage points more often than the most susceptible, in the luck and skill tasks respectively. Also, to separate the role of the sunk costs from a more general notion of stubbornness, we included a third (and final) psychological measure, taken from Wilkins (2015) which we did not find to have predictive power.

The paper proceeds as follows. Section 2 gives a discussion of various related literatures. Section 3 provides a model in order to formalize and clarify our conceptualization of advice, and derive the econometric specification that we use. Section 4 details the experimental design. Section 5 provides the results. Section 6 concludes.

2 Literature

Our work draws and expands on several important areas of literature both within and outside economics.

Whether or not people take advice has been studied with regards to many specific contexts and factors, using various methodologies. In finance, Önkal et al. (2009) find that stock-price forecasts can be favored when they are made by humans rather than machines. Mullainathan et al. (2012) document that financial advisers do not de-bias their clients, which can often lead to higher profits for advisers. On the perceived credibility of experts on policy issues, Lachapelle et al. (2014) show that the way in which experts frame information matters, while Doberstein (2016, 2017) finds large differences depending on the type of institution publishing the work (academic, think tank or advocacy group). Hilger (2016) shows theoretically how providers of credence goods will make use of their informational advantage to overcharge consumers in equilibrium. Medical surveys have long documented a significant proportion of patients not following advice (see e.g., Davis, 1968; for a review of patient non-adherence see Kardas et al., 2013). The face-to-face experiments of the "judge-adviser system" have been used to identify various determinants of advice-taking related to sense and appearance e.g., adviser confidence (Swol and Sniezek, 2005). When there is uncertainty over the quality of advice, many relevant issues arise, including the discounting of advice relative to one's own opinion (Weizsäcker, 2010; Yaniv and Kleinberger, 2000) which has been shown to vary with task difficulty (Gino and Moore, 2007) and many others e.g., overconfidence, risk attitudes etc. Also related, Cook and Lewandowsky (2016) document a polarization of the beliefs of US participants in response to being presented with information concerning the consensus in the scientific community regarding anthropogenic global warming.

In contrast, our experiment provides the first abstracted setting with neutral, context-free language to study the questions of whether, when and why *good* advice is ignored. To do so, we conceptualize the advisee's position and the adviser's advice as probability distributions, the general idea of which dates back to papers such as Morris (1974). We also control for the quality of advice and remove uncertainty over whether advice is good. More precisely, when advice is good (or bad) there is a simple stochastic dominance ordering of the advisee's position and the adviser's advice. This makes the rational action unambiguous and removes the scope for overconfidence and risk attitudes to play a role. We conduct novel treatments to determine the effects of the adviser being superior in skill vs. luck, and the remuneration of the adviser. We utilize our experimental setting to identify these effects, which would be difficult to isolate in the real world: whether an adviser reached their position through relative luck or skill, and in what combination, is difficult to estimate and likely endogenous to the value of their advice; the remuneration of an adviser is potentially observable, but is likely endogenous to the quality of the adviser, as well as to any corruption or bias. Finally, we provide at least a partial explanation of why good advice may not be followed in terms of underlying psychological measures, which we discuss next.

According to Parrott and Smith (1993), envy "occurs when a person lacks another's superior quality, achievement, or possession and either desires it or wishes that the other lacked it." This definition presents the essential dichotomy arising from envy itself: it can lead to an active attempt to diminish others or an active attempt to improve oneself. The philosopher Bertrand Russell argued that envy was quite possibly the root cause of unhappiness in Western society, but he too highlighted the dual effects arising from envy: arguing that while it brings unhappiness, it can also be harnessed to positive ends. He cited the rise of democracy as a tangible good that has come from envy (see Russell, 1930). Van de Ven et al. (2009) take this a step further, pushing the idea that envy is really two separate concepts: a positive force for self-betterment, and a negative force of self-destruction. If envy really is one of the key driving forces of Western society and the primary cause of (un)happiness, then it might seem odd that it has received surprisingly little attention from economists. However, this may relate to the ambiguity over the implications of the concept. For example, in our context, if individuals are envious of the income or status of experts they might wish to take good advice in order to raise their own payoff and narrow the gap. On the other hand, it might be that envy results in an emotional response to shun advice, even if it harms an individual's own payoff.

Within economics, envy has been studied in a variety of contexts though nothing directly related to our own. Following Brenner (1987) and Kuziemko et al. (2015) we know that individuals care about their relative economic status, and Elster (1987) even makes the further refinement that some may be motivated by the desire to avoid generating envy in others. Mujcic and Oswald (2017) show that envy is a powerful predictor of falling levels of mental well-being, especially among young adults in their analysis of a panel of 18,000 adults. Related, Winkelmann (2012) finds that the prevalence of luxury cars in a municipality has negative consequences for life satisfaction which he attributes to envy. Within welfare economics, a number of papers have examined the normative significance of envy, for instance Baumol (1986) as part of his broader examination of fairness, and Foley (1967) and Varian (1974) who consider the implications of an envy-free welfare equilibrium. Banerjee (1990) examines how the distortions caused by envy might be removed through progressive income tax. Brennan (1973) makes the point that envy can motivate support for redistribution from those who want to see the rich made poorer, and Mui (1995) examines the envy that followers may feel towards an innovator. Within experimental economics, envy is also linked to the documented "money-burning" phenomenon (e.g., Zizzo and Oswald, 2001) where players in a game are willing to damage their own utility in order to punish others, and Kirchsteiger (1994) attempts to use envy as a possible explanation for behavior in the ultimatum game. Money-burning and envy more generally are also linked to the literature on fairness, for example, Fehr and Schmidt (1999) suggest that individuals may possess a "willingness to sacrifice potential gain to block another individual from receiving a superior reward", though care needs to be taken not to confuse envy with a general desire for fairness, as noted in Kirchsteiger (1994). Leibbrandt and López-Pérez (2012) examine the motivation for punishments (including envy) inflicted by affected second and unaffected third parties in a set of games, and Blanco et al. (2011) investigate inequality aversion (from the model developed in Fehr and Schmidt, 1999) in four games, relating the propensity to block the rewards of others to envy. Although related to "money-burning" in a loose sense, the effect of rejecting expert advice is different: by ignoring good advice, individuals can harm their own payoff without directly affecting the adviser's payoff. More generally, while none of these papers within economics have a focus on the take-up of advice (good or otherwise), they share the common theme that envy is an important and powerful psychological concept that needs to be better incorporated and understood.

There are also important practical issues on the measurement of envy: if we measure envy through behavior or outputs then issues of endogeneity are likely to hamper any attempt to understand the effects of the concept further. It is for this reason that we deploy a dispositional envy scale (Smith et al., 1999) which we believe we are the first to use in economics.

The literature on stubbornness has largely been developed within psychology and management with the main focus being on understanding and measuring the phenomenon, typically through surveys. Wilkins (2015) provides a summary and a scale which we modify and use to measure general stubbornness. We are the first to use such a scale within economics and to link stubbornness in this way to the avoidance of good advice. More common within economics is the study of a related concept: the susceptibility of individuals to the sunk cost fallacy. There are numerous papers that try to evaluate the extent of this fallacy, through which individuals fail to realize that sunk resources (e.g., time, money, energy etc.) should be ignored when making a decision. We believe we are the first to develop a scale to measure susceptibility to the sunk cost fallacy, for which we draw on Thaler (1999) and Arkes and Blumer (1985) to devise the questions we put to our experimental participants.

3 A Model of Advice

This section serves three roles. Firstly, it provides a more formal set of definitions of the terms used throughout the paper. Second, it maps the terms of the underlying model to the fundamental elements of the experimental design. Finally, it derives the econometric specification used later in the paper.

To formalize the vocabulary of the paper, suppose there is a true state of the world s and that individuals receive a payoff of one for correctly estimating the state, otherwise they receive zero. Let each individual i have a probability of estimating the state correctly, $p_i \in [0, 1]$, heterogeneous across individuals. Given this, we define the terms "advisee", "adviser", "advice", "good advice" and "value of advice".

Definition (Advisee). *Individual i is an advisee if they have the option to replace p_i with p_j for some other individual $j \neq i$.*

Definition (Adviser). *Individual j is an adviser if there exists an individual $i \neq j$ who has the option to replace p_i with p_j.*

Definition (Advice). *If j is an adviser, p_j is advice.*

Definition (Good and bad advice). *For advisee i and adviser j, p_j is good advice when $p_j > p_i$ and bad advice when $p_j < p_i$.*

Definition (Value of advice). *For advisee i and adviser j, the value of advice is $p_j - p_i$.*

The conceptualization of advice as a probability distribution goes back to works such as Morris (1974). In his words, "conceptually, consulting an expert is like performing an experiment where the observed data is a function (probability distribution)": it is as if the decision-maker is provided with a second choice represented by a probability of being correct. This is exactly the notion of advice that we carry through into our experimental design. In our experiment, we give advisee participants binary decisions of whether to accept advice regarding two possible states of the world: winning a bonus payment or not. In reality, depending on the domain of the advice, it is often possible to partially accept advice or to accept it in some dimensions rather than all. There may also be multiple payoff-relevant states of the world. The idea that advice can be simplified down to a binary decision (yes or no, accept or ignore, stick or switch) is a widely used convention, see for instance, as Calvert (1985) puts it (p. 534): "This feature represents the basic nature of advice, a distillation of complex reality into a simple recommendation."

Our experiment consists of two waves. In the preliminary wave, we ask a small number of participants to complete tasks. Based on their performance, for each task we select some participants to act as advisers. In the main wave, all participants are advisees.

We now map the model onto our experimental design. In our experiment, there are two tasks. In the luck task, participants guess whether a coin lands "heads" or "tails" for each of ten tosses. In the skill task, participants select an answer to non-verbal reasoning IQ questions from eight possibilities. For each task, the state of the world, s, is the correct answer to a (uniformly) randomly chosen question. For example, for the luck task, say the fifth toss is selected and it was H, then $s = \{5^{\text{th}} \text{ toss is H}\}$.[9] If advisee i obtains say 4/10 in the coins task, then i's probability of being correct using their own answers is $p_i = 0.4$. However, advisees also have the option to use the answers of an adviser. Suppose the adviser, j, obtained 7/10 i.e., $p_j = 0.7$. The advisee, i, is then faced with a choice between two probabilities of being correct: $p_i = 0.4$ or $p_j = 0.7$. We suppose that when an individual makes the decision of whether to accept advice (i.e., choose p_j) or ignore it (i.e., choose p_i), they compare their expected utility in each case and take the action that yields the higher expected utility. Therefore, absent any other, "non-rational" forces, when participant i is faced with the decision between p_i and p_j, i chooses p_i if and only if $p_i \geq p_j$. Where this is how i acts, we say that i is rational, or, that i has taken a rational action.

However, individuals may depart from rationality by ignoring good advice in systematic ways. In other words, where $p_i - p_j > 0$ i.e., the value of advice is positive, advisee i may not always accept advice. To account for this, we allow for a more general effect of $p_i - p_j$ as well as for additional factors which may determine the probability that individ-

[9] Similarly for the skill task, say the third IQ question is selected and the correct answer is option 8, then $s = \{3^{\text{rd}} \text{ question is } 8\}$.

uals follow good advice. We denote these other factors as variables $x_k, k = 1, \ldots, K$, and allow for an error term, u. In our analysis, these variables of interest include our between-subject treatments, psychometric measures such as envy, susceptibility to the sunk-cost fallacy, stubbornness, interactions of these variables, and participant demographics.

We now derive a probit specification for estimation from a latent variable model in a standard way. Firstly, fix a common adviser with advice $\bar{p}$ which is good advice for any advisee i i.e., $\bar{p} > p_i$ for all i.[10] For advisee i, we suppose that i's utility function is such that the difference in i's expected utility from accepting and ignoring advice takes the following form:

$$E\left[U_i\left(\text{accept}\right)\right] - E\left[U_i\left(\text{ignore}\right)\right] = \sum_v \delta_v \phi_i + \sum_{k=1}^{K} \beta_k x_{i,k} + u_i,$$

where $\phi_{i,v}$ is an indicator variable, equal to one if the value of advice for i, $\bar{p} - p_i > 0$, is v, and equal to zero otherwise. Assuming i chooses $y_i \in \{\text{accept}, \text{ignore}\}$ to maximize expected utility implies:

$$y_i = \text{accept} \iff \sum_v \delta_v \phi_i + \sum_{k=1}^{K} \beta_k x_{i,k} + u_i \geq 0.$$

We assume that $u_i \sim \text{iid}N(0,1)$[11] and are therefore left with the following probit specification for estimation:

$$\begin{aligned} pr(y_i = \text{accept}) &= pr\left(u_i \geq -\sum_v \delta_v \phi_i - \sum_{k=1}^{K} \beta_k x_{i,k}\right) = pr\left(u_i \leq \sum_v \delta_v \phi_i + \sum_{k=1}^{K} \beta_k x_{i,k}\right) \\ &= \Phi\left(\sum_v \delta_v \phi_i + \sum_{k=1}^{K} \beta_k x_{i,k}\right). \end{aligned}$$

Given this model, our empirical strategy will be to first establish whether good advice is indeed ignored and then, if so, to assess the importance of our x_k variables on the decision to accept or ignore good advice using the probit specification.

4 Experimental Design

The participants in our study were from the Amazon Mechanical Turk online pool of subjects. The software used to perform the experiment was Qualtrics. The experiment was registered in advance in the AEA RCT Registry (for details see Ronayne and Sgroi, 2017). Here we provide a summary of the overall experimental design before going into

[10] In the experiment, we set $\bar{p} = 0.7$ in the luck task and $\bar{p} = 0.9$ in the skill task.

[11] The choice of $\sigma = 1$ is without loss of generality: $\sigma \neq 1$ could also be assumed but as σ cannot be identified, convention is to set $\sigma = 1$.

detail on each part.

A key feature of the design was to make clear to participants when advice was good or bad. In order to achieve this, we needed to fix the quality of advice and make this quality known to participants. To this end we define good advice in terms of the probability of success (winning a bonus payment). Without taking advice, an advisee has a known probability of success, but the advisee is offered the chance to change this probability by accepting advice. A rational advisee would accept good advice (which we can think of as "switching" from the initial probability of success to a new probability generated by the adviser), but ignore any advice that has a lower probability of success than she had initially (which we can think of a "sticking" with her original probability of success).[12]

Our design involved two waves of data collection. In the first wave, 75 subjects undertook two incentivized tasks: the first involved guessing the number of heads in a series of coin flips, the second involved undertaking a short Ravens visual IQ test. Some subjects in this first wave were then the advisers who featured in the second wave. In the second wave, 1,503 subjects undertook the same two tasks. The addition over wave 1 is that subjects in wave 2 had the option to switch their own answers for that of someone who achieved a high score in wave 1. For each task, they were told both their score and the score of the adviser from wave 1 and they were then offered the chance to submit their own answers ("stick") or the answers of the wave 1 adviser ("switch"). For ten randomly chosen subjects, a single problem was chosen from their submitted answers and a bonus payment made if the correct answer was provided. These choices were followed by three sets of questions designed to measure stubbornness and envy. The subsections below provide more detail.

4.1 Wave 1: Choosing Advice-givers

The 75 subjects in this wave were paid \$2.00 for completing the experiment which took an average of 7 minutes 34 seconds to complete, corresponding to an hourly wage of \$15.86. Participants were first asked to guess the outcome of a series of ten coin flips, one at a time. They were told that doing "especially well" in this task would result in a bonus payment of at least \$0.50. They undertook two practice questions (flips) to acclimatize them to the software and were given feedback on their performance in the practice questions prior to starting the main questions. They were then asked to undertake a set of ten Ravens visual IQ questions. Again, they were again told that doing especially well in this task would result in a bonus payment of at least \$0.50 and were first tasked with a practice question and received feedback on their performance before undertaking the full test. They were then asked to complete a questionnaire that included questions on the difficulty of the tasks before receiving feedback on their scores followed

[12] In the instructions we use the wording "Keep my answers" or "Use the other worker's answers": see the transcript in the Appendix.

by demographic questions including gender, age, race, household income and information on their education and political affiliation.

This wave was specifically designed to provide advisers for the main, second wave. The performance of subjects in wave 1 was catalogued and advisers were chosen to be those with scores that placed them towards the top of the distribution for each task: the adviser for luck (coin-toss) task was a participant with a score of 7/10 and for the skill (Raven's visual IQ) task, a participant with 9/10. Crucially, their scores were chosen so that they could be beaten by only a small minority of wave 2 subjects. This meant that we would have some data on how subjects with higher scores than the advisers behaved.[13] Another important feature was the wording of the bonus payment instructions: we carefully informed wave 1 subjects that they would receive "at least $0.50" which gave us freedom to allocate different bonus payments and left performance incentives identical across subjects who ended up receiving different bonus payments.

4.2 Wave 2: Stick or Switch?

We recruited a further 1,503 Mturk participants to take part in wave 2.[14] Participants in this wave were paid $2.00 for completing the experiment which took an average of 12 minutes 11 seconds to complete, corresponding to an hourly wage of $9.85. Wave 2 began identically to wave 1. Once again, subjects guessed the outcome of a series of ten coin flips, one at a time. They were told that a good performance in this task would result in a bonus payment of at least $0.50, and undertook two practice questions to acclimatize them to the software (and were once again given feedback on their performance in the practice questions prior to facing the main task). Again, they were next asked to undertake a set of ten Ravens visual IQ questions and told that if they did especially well in this task it would result in a bonus payment of at least $0.50. Once again they were first tasked with a practice question and received feedback on their performance before undertaking the main test. The next part of the experiment marked the first difference between wave 1 and 2: subjects in wave 2 were now informed about the existence of the advisers from wave 1 as follows:

> "... we here describe the experiences of two other workers who, some time ago, completed the exact same tasks you have just tried for the same $2.00 HIT reward. Please pay attention as we will be asking you some comprehension questions about them on the next screen.
>
> These two workers both saw the same instructions you did. This means they were both told that they could receive at least $0.50 for doing especially well

[13] We ran pilots using 501 subjects in total to give us prior information on the likely distribution of scores for the tasks in waves 1 and 2, and to test our other measures. More detail on the pilot studies are available on request.

[14] Participants from wave 1 were excluded from participating in wave 2.

in a task. In addition, they each did especially well in one of the tasks: one scored 7/10 in the coin task, and got a bonus of X; the other scored 9/10 in the logic puzzles task, and got a bonus of X."

The X in the text they were given was varied by treatment. We set $X = \$0.50$ for all participants except for those allocated to the "high-remuneration" treatment where we set $X = \$100.00$. Note from the wording that it was made clear to subjects in wave 2 that the earlier workers who did well also undertook the same tasks and with the same instructions. The next page involved a series of comprehension questions designed to increase concentration and to have an indication of whether the subjects understood everything so far.

Next was the crucial accept ("switch") or ignore ("stick") decision. There was a careful explanation of how decisions would translate into possible bonus payments. Subjects were informed that ten participants would be chosen at random and that if they were chosen, one of their answers (from either task) would be selected, and if correct, they win a bonus. Before proceeding, they were asked to either keep their answers or use the other worker's answers where the "other worker" was the adviser just described on the previous page.[15] Participants were asked this for each of the two tasks i.e., they could independently choose to keep their own answers or use the adviser's in each task. On the same page, we showed the subjects their own score and reminded them of the score of the adviser, and the adviser's received bonus payment. We also explained the probabilities of winning the bonus payments conditional on being selected, both in the case where they chose to keep their own answers and in the case where they chose to use the other participant's answers. Additionally, for those in the second, "per-follower" remuneration treatment, participants were informed that for every wave 2 subject who accepted the answers of the wave 1 participant's answers, the wave 1 subject's bonus would rise by \$0.50.

The wave 2 subjects across all three treatments then answered questions concerning their gender, age, race, household income, education and political affiliation before moving on to the final page of the experiment.

4.3 Wave 2 continued: Testing for Envy and Stubbornness

Wave 2 concluded with three sort questionnaires designed to generate three behavioral measures: dispositional envy, susceptibility to the sunk-cost fallacy, and (general) stubbornness.[16]

The first set of questions was the dispositional envy scale (DES) of Smith et al. (1999), which yielded a measure with a reasonable degree of variation.[17] The subjects were asked

[15] Note that we refrained from using words like "adviser" or "expert" in the text.

[16] The decision to focus on these three behavioral traits was taken before the experiment was initiated, as discussed in the AEA RCT pre-registry entry, Ronayne and Sgroi (2017).

[17] A histogram of the data generated by the DES is given in Figure A1 in the Appendix.

to evaluate the extent to which they agreed with eight statements following a simple 5-point Likert scale (strongly disagree, moderately disagree, neither agree nor disagree, moderately agree, strongly agree). Some example statements are: "I feel envy every day" or "The success of my neighbors does not make me resent them". The scale was generated for each participant by adding the scores from each question where responses indicating a higher level of dispositional envy received a higher score. With eight questions generating a score of between 1 and 5, the total score for each individual was between 8 and 40 with a higher score representing a higher level of dispositional envy.

We restrict ourselves to the study of situations when potential advisees have a prior opinion which forms their initial probability of success in the decision-making task. In our experiment this prior opinion is formed through time and effort and the take-up of good advice which suggests there may be a role for the "sunk cost fallacy": earlier work has documented the widespread role of the sunk cost fallacy which leads individuals to undertake utility-reducing actions because they have paid in advance with resources such as time, money or effort. This trait may also affect the uptake of advice if individuals have already sunk time and effort to form beliefs which lead them in one direction even if those beliefs are likely to generate lower payoffs than simply following expert advice. In this paper we build the first stubbornness scale designed to measure susceptibility to the sunk cost fallacy, inspired by the work of Thaler (1999) and Arkes and Blumer (1985). Our scale is based on answers to five scenarios, for example:

> "Imagine that you have spent \$20 on a ticket to a concert. The day of the concert comes and unfortunately it is snowing heavily, and you feel tired after a tough day. You know you would not have decided to go to the concert if you hadn't already bought the ticket, but you also know that you cannot get a refund. On balance you decide to go to the concert."

Subjects were asked to note their agreement on a 5-point Likert scale. Summing across the five questions created a scale with a range from 5 (minimum susceptibility) to 25 (maximum susceptibility). Arkes and Blumer (1985) note that individuals commit the sunk cost fallacy when they continue a behavior or endeavor as a result of previously invested resources (e.g., time, money or effort) across a variety of dimensions: we try to capture some of these dimensions in our set of five scenarios, with one of our scenarios based on Thaler (1999). To our knowledge we are the first to form a scale to measure susceptibility to the sunk-cost fallacy within economics, though the general method follows the same principle as the DES or other similar scales in psychology e.g., perhaps most famous of all, the Big Five Inventory used to measure various personality traits (see John and Srivastava, 1999).

Finally, the subjects faced a set of questions designed to generate a general stubbornness score to provide a direct alternative to our measure of susceptibility to the sunk cost fallacy. Once again participants were asked to indicate agreement using the same 5-point Likert scale ranging from strongly agree to strongly disagree. Five questions were asked yielding an overall score between 5 and 25. An example is "I do something I want to do even if no one else wants to do it." Again, to the best of our knowledge we are the first to form and use such a scale in economics, but tests of this type are common in the management literature e.g., Wilkins (2015) which provided the five questions used in our scale. For each scale, questions were in some cases "inverted" so that agreement indicated a lack of envy, susceptibility to the sunk-cost fallacy or stubbornness or in some cases agreement indicated the opposite. This was done to prompt subjects to read the questions carefully and to discourage them from answering all questions identically.

4.4 Features of the Design

Here we examine some special features of the design which may not be apparent at first glance, but which highlight the importance of the level of control we have through running an experiment.

First, let us turn to the use of two different tasks. This allows us to not only consider the role of task-type, but also to make within-subject comparisons without fear of learning between tasks. The tasks themselves were chosen as they reflect differing levels of luck and skill. The coin task is entirely luck-based, while the Ravens visual IQ test is largely a skill-based task (albeit with luck playing some role because participants may select the correct answer by chance). The differing tasks also allows us to explore an additional issue: alongside our primary focus on envy and natural exploration of the role of stubbornness (or susceptibility to the sunk cost fallacy) we might ask whether concerns about fairness play a role. Specifically, it may be that individuals are likely to consider higher rewards fair for experts who seem to have earned their payoffs (e.g., status and pay) through hard work or ability and this might weaken the role of envy, but not for those who seem to be rewarded for being lucky or simply for being born in the right family or situation. This suggests it is important to control for perceptions about luck or skill when advice is considered (something known to vary a great deal among individuals; Alesina et al., 2001). Accordingly, each individual in our experiment decides whether to accept advice separately in both the coins task (entirely luck-based) and the Ravens visual IQ test (largely skill-based) allowing us to directly test whether context and perceptions around luck/skill really matter. Note that fairness can also be examined by considering our second "per-follower" remuneration treatment which directly rewards the adviser when a participant uses their answers: if participants feel any reluctance to use the answers of others without the advice-giver being rewarded directly then we should see good advice being taken more often by those in this treatment compared to those in the control.

A second important feature is our ability to isolate the roles of adviser remuneration, envy and stubbornness. Surveys such as the Edelman Trust Barometer or our Brexit survey reported in Section 1 can be indicative but do not drill into what factors might cause expert advice to be ignored, nor can they establish a causal link even if any plausible factors are identified. The level of control we need is only likely to come about in an experimental setting in which we can control the environment and incentives. For example, in practice, higher quality experts are likely to be more highly remunerated so expert pay can be viewed as a signal of quality. This can act as a confound when attempting to study the relationship between the propensity to follow advice and adviser remuneration in a real-world setting, whereas our design fixes expert quality independently of adviser remuneration. In each condition, different bonuses were awarded, but all participants had the same information throughout the experiment. This was made clear to wave 2 participants, meaning they knew adviser performance was not driven by any extra effort that may have arisen from the higher incentives. In turn, this means that any acceptance of their advice was not in order to reward the advisers for higher effort, enabling us to isolate the effect of remuneration per se.

Our design also includes the endogenous production of advice, which brings several advantages. If advice was disseminated by the experimenter we would have to contend with possible reciprocity from participants (the so-called "demand effect", see Zizzo, 2010). If participants were selected as advisers by the experimenter there might also be issues of fairness and an additional channel for envy. Recall also the finding from the Edelman survey that many people were as likely to listen to advice from people like themselves as from academic or technical experts. By having advisers endogenously emerge from a prior wave of the experiment and drawn from the same MTurk pool of workers, we provide a measure of control against this bias.

Finally, note that both the participant's and the adviser's score are presented to the participant so there is no doubt about the value of the advice, and incentives are such that individuals who perform worse than an adviser should opt to accept the answers chosen by the adviser in order to submit a higher number of correct answers. This allows us to focus on a direct test of rationality in the context of advice-taking, something all but impossible to examine outside of a controlled experimental setting. While it is rare to have such certainty over the quality of advice in the real-world, it is only by fixing this quality and making it known that we can pursue causal explanations for any failure to follow *good* advice.

5 Results

Wave 2 consisted of a total of 1,503 participants.[18] Each of these participants made choices regarding both the luck and the skill task. In total, this makes 3,006 decisions of whether to follow advice or not. Table 2 displays a breakdown of these decisions. In a total of 2,757 of the decisions, advice was good and the rational action was for the participant to accept the advice (accepting would have strictly increased expected payoff). In 176 cases, the participant's score tied with the adviser's making either decision rational. In the remaining 73 decisions, the participant achieved a strictly higher score than the adviser so the rational decision was to ignore the advice.

Table 2: The rational decisions

The rational decision was to:	n
Accept	2,757
Ignore	73
Indifferent	176
Total	3,006

Our first result is that good advice is frequently ignored. Moreover, we show that the rate at which good advice is accepted is significantly lower than the frequency implied instead by some random error rate. To do so, we compare the proportion of participants making the rational decision to ignore the advice against the proportion making the rational decision to accept the advice. The results are presented in Table 3. Of the 73 decisions where it was rational to ignore the advice, 71 took the rational action (97.3%). However, of the 2,757 decisions where it was rational to accept the advice, 2,078 took the rational action (75.4%). The difference in the proportion of rational decisions between the two groups is significantly different from zero ($P < 0.001$): good advice is frequently ignored.

To offer some context, we note that in experiments of choices over lotteries, the rate at which participants violate stochastic dominance varies depending on the complexity of the lotteries offered. However, when they are in their simplest forms, violation rates have been documented to be low. For example, Birnbaum (1999) studied violations of stochastic dominance in online samples where in perhaps the simplest binary choice offered, 6% of participants chose the dominated option.[19] We expected this level of error

[18] Demographic information on these participants is shown in Table A1 in the Appendix. Wave 1 collected data from 75 participants, from which the advisers were chosen, results from this wave are available upon request.

[19] This choice, offered to participants was between [$4 with 0.5, $96 with 0.3, $100 with 0.2] and [$4 with 0.5, $12 with 0.3, $100 with 0.2] (choice 3 of his Table 3) where 6% chose the latter.

to constitute an upper bound on the error rate in our experiment for three reasons: our participants face an even simpler, binary choice, where each lottery has two outcomes rather than three; the outcomes we offered (dollar amounts) were the same in the two lotteries where one was zero (the other was $0.50); and we offered a verbal explanation of the choices. Indeed, we found that only $< 3\%$ irrationally accepted bad advice by choosing the stochastically-dominated option. Against such a marker, the rate of 25% at which participants ignored good advice stands out starkly; see Table 3 below.

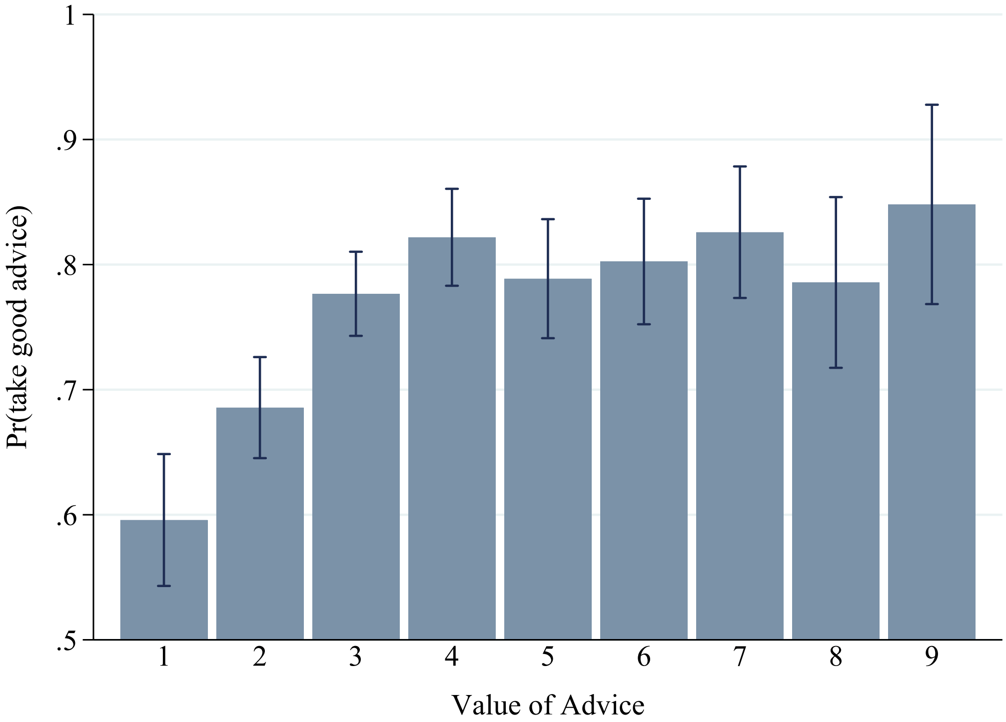
Table 3: Good advice is ignored

	Advice			Pr(rational)
	accepted	ignored	n	
Rational to accept	2,078	679	2,757	0.754
Rational to ignore	2	71	73	0.973
Difference in proportions				0.219
P-value from test of difference in proportions				<0.001

Our experimental design allows for variation in the value of advice (as defined in Section 3). Figure 1 below shows how the proportion of participants taking good advice varied with value. The correlation between the proportion taking good advice and value is 0.135 ($P < 0.001$), although Figure 1 shows the relationship over the whole domain to be non-linear. Value was at its lowest of 1 when a participant scored one less than the adviser. There, we find that only 59.6% of participants accepted the good advice. When value was 2, this rose to 68.6% and for a value of 3, it was 77.7%. For higher levels of value, the proportion plateaued at about 80%. Together, this suggests that our participants were trading off the value of advice against other, non-rational forces.

We have established that participants did not always take the rational decision to accept good advice. We now investigate the effect of our treatments. Firstly, we use our between-subject treatment to investigate the effect of adviser remuneration on the propensity to take good advice. In the control condition, advisers received $0.50 as a reward for being made advisers. In our first remuneration treatment, advisers instead received a high lump-sum amount of $100.00 as a reward. In our second remuneration treatment, advisers received $0.50 plus an additional $0.50 "per-follower" i.e., for every advisee that accepted their advice. In all conditions, the remuneration received by advisers was told to advisees. The results for the two treatments are reported in Tables 4 and 5 respectively. We find that those assigned to the treatment where the adviser received a higher level of remuneration exhibited a lower propensity to follow good advice. This treatment effect was 5.4 percentage points ($P = 0.008$). In contrast, we did not find a

Figure 1: Trading-off rationality



The value of advice is defined as the difference between the adviser's score and the participant's score (see Section 3). This chart only shows data from participants who should (rationally) have accepted the advice, hence value is positive. There are at least 50 decisions for each level of value.

significant difference in the propensity to take good advice across the control and the "per-follower" treatment, as reported in Table 5.

Table 4: The effect of high lump-sum adviser remuneration

Adviser remuneration	Good advice			Pr(good advice
	accepted	ignored	n	accepted)
\$0.50 lump sum	703	212	915	0.768
\$100.00 lump sum	647	259	906	0.714
Difference in proportions				0.054
P-value from test of difference in proportions				0.008

Table 5: The effect of per-follower adviser remuneration

Adviser remuneration	Good advice			Pr(good advice
	accepted	ignored	n	accepted)
\$0.50 lump sum	703	212	915	0.768
\$0.50 lump sum + \$0.50 per follower	728	208	936	0.778
Difference in proportions				0.009
P-value from test of difference in proportions				0.627

Next, we use our within-subject treatment to test whether good advice is ignored more or less frequently when the adviser has achieved their status through skill rather than luck. Each participant completed both a luck and a skill task as well as deciding whether to take advice in each. In order to look at the effect of adviser skill per se, we restrict attention to participants for whom: i) advice was good and ii) the value of advice was the same in each task. There were 143 such participants, corresponding to 286 decisions. Data from these participants is displayed in Table 6. Of these 143 participants, there were 12 (8.4%) who ignored the skilled adviser's advice but accepted the lucky adviser's advice. Conversely, there were no participants (0.0%) who accepted the skilled adviser's advice and ignored the lucky adviser's advice. An exact binomial test strongly rejects the null hypothesis that the probability of participants belonging to one of these two groups is equal ($P < 0.001$).

Table 6: Good advice is ignored more often when the adviser is skilled

	action given value equal across tasks			
	accepted	ignored	n	Pr(accepted)
Luck	127	16	143	0.888
Skill	115	28	143	0.804
Difference in proportions				0.084
P-value from binomial test				<0.001

The data here are from the 143 participants for whom i) advice was good and ii) the value of advice was the same in each task.

We now provide evidence that this behavior is driven by underlying psychological variables. We consider two particularly relevant behavioral forces behind the decision not to take advice: envy and stubbornness.

It was not clear ex-ante whether or when envy would have a positive or negative effect on the propensity to take good advice. In general, and as discussed earlier, it has been said that there are two sides to envy (Russell, 1930; Van de Ven et al., 2009). It could be that through relative comparisons to the adviser, envy could encourage self-betterment and directly increase the proportion of individuals taking good advice. Alternatively, it may be that those more disposed to envy incur a negative emotional response to exposure to someone superior to themselves in some dimension, which here could result in good advice being ignored. Through the use the dispositional envy scale of Smith et al. (1999) and our design, we investigate when the effect of envy is to drive a positive or negative response in the propensity to take good advice.

Stubbornness reflects a general unwillingness to budge from one's current position or opinion. We consider both a general measure of stubbornness (Wilkins, 2015) and a measure of susceptibility to the sunk-cost fallacy for which we constructed a novel scale. Susceptibility to the sunk-cost fallacy is especially relevant to advice-taking contexts because individuals have often invested resources (time, effort, money etc.) in forming their position. In our experiment, this was reflected by the time and effort participants put into the tasks. By pitting these two measures against each other in regressions, we are able to retrieve the relative importance of the resources sunk in forming a position over and above a general reluctance to leave one's position.[20]

We now examine the effect of our psychometric measures on the probability of accepting good advice. Table 7 provides the average marginal effects (AMEs) from probit regressions as per the specification derived in Section 3. The AME of envy on the

[20]Histograms of our participants' scores on all three psychometric scales are provided in Figures A1-A3 in the Appendix.

probability of accepting good advice is positive and significant in the domain of skill ($P = 0.030, 0.029, 0.022$ in specifications (4)-(6) respectively) but not in the domain of luck ($P = 0.679, 0.641, 0.540$ in (1)-(3) respectively). Interpreting the estimated AMEs in the domain of skill, a one-standard-deviation increase in dispositional envy increases the probability of taking good advice by roughly 2.5 percentage points on average.

Table 7: Determinants of following good advice

Average Marginal Effects	Luck			Skill		
$y = \mathbb{I}\,$(took good advice)	(1)	(2)	(3)	(4)	(5)	(6)
Envy	0.005	0.006	0.008	0.025	0.025	0.026
	(0.013)	(0.013)	(0.013)	(0.012)	(0.012)	(0.011)
Sunk-Cost Fallacy	-0.038	-0.035	-0.034	-0.027	-0.025	-0.026
	(0.012)	(0.013)	(0.013)	(0.011)	(0.011)	(0.011)
Stubbornness	-0.005	-0.007	-0.007	-0.013	-0.013	-0.014
	(0.013)	(0.013)	(0.013)	(0.012)	(0.012)	(0.011)
Remuneration Treatment 1			-0.060			-0.055
			(0.030)			(0.026)
Remuneration Treatment 2			-0.013			-0.001
			(0.029)			(0.025)
Remuneration × envy			X			X
Psych. interactions		X	X		X	X
Value dummies	X	X	X	X	X	X
Participant demographics	X	X	X	X	X	X
Observations	1,317	1,317	1,317	1,427	1,427	1,427

Average marginal effects are shown with standard errors in parentheses following probit regressions. The estimates from the probit regression are included in Table A2 in the Appendix. All specifications pass various mis-specification tests of functional form and heteroskedasticity. The binary dependent variable is 1 (0) if the decision made was to accept (ignore) the good advice. "Value dummies" refers to the inclusion of a dummy for every level of value in the underlying probit regression (bar one, to serve as the reference category). "Psych. interactions" refers to the inclusion in the underlying probit regression of all possible interaction terms between the three psychometric variables (envy, susceptibility to the sunk-cost fallacy and stubbornness). Each participant decided whether to take the advice in each task, hence there were 1,503 decisions for each task. Of those in the skill (luck) task, 68 (181) were excluded because they out-performed the adviser i.e., advice was bad. Of the remaining participants, 8 (4) were excluded because they did not disclose their sex. Finally, there was only one respondent with a value of 7 in the luck task, who was dropped for the analysis.

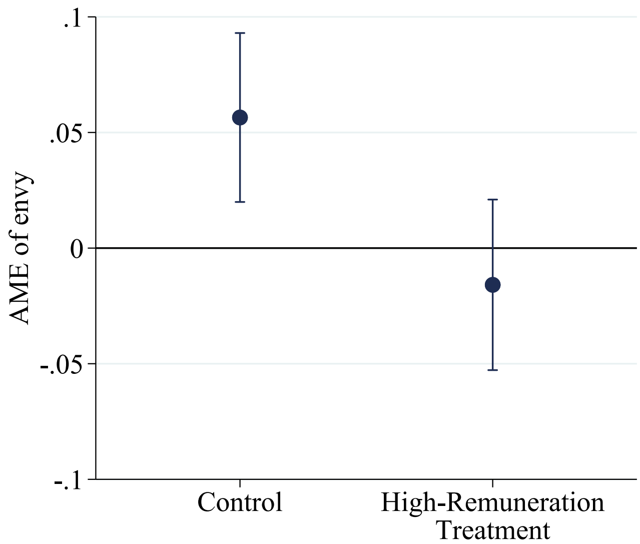
However, this positive effect of envy is averaged across the different remuneration treatments. We now show that the effect of envy on the propensity to take good advice changes significantly depending on whether or not the adviser was highly remunerated. In other words, we find a significant interaction effect between envy and whether a subject was in the high-remuneration treatment condition. Figure 2 shows the AMEs of a one standard deviation increase in envy on the propensity to take good advice in the hypothetical scenario where all participants are allocated to the control versus where all are allocated to the high-remuneration treatment. The estimates reveal that the positive overall effect of envy found in Table 7 can be meaningfully decomposed by whether the adviser was highly remunerated or not. When all are assumed to be in the control, where the adviser was remunerated with the same low amount that the advisee could achieve, envy had a large positive effect: 5.6 percentage points for a one standard deviation increase in envy ($P = 0.002$). However, when all participants are assumed to be in the treatment where the adviser had been highly remunerated, the effect of envy was significantly different ($P = 0.004$ from a contrast of AMEs test) with an estimate of -1.6 percentage points for a one standard deviation increase in envy.[21] To describe the effect of envy more precisely, Figure 3 provides average adjusted predictions of the probability of accepting good advice in the two scenarios, over the range of participants' dispositional level of envy. Where participants are assumed to be in the control condition (left panel), there is a monotonic increasing relationship between envy and the propensity to take good advice (we knew the relationship was increasing on average from the AME of 5.6 in Figure 2). The effect of envy can be seen to be large here. For the least envious participants, we predict on average that they will follow good advice about 71% of the time. In contrast, we predict on average that the most envious participants will follow good advice about 93% of the time, an increase of more than 20 percentage points. On the other hand, where participants are assumed to be in the high-remuneration treatment (right panel), the positive relationship is lost. For less-envious participants, the differences in the predicted probabilities (across panels) are not significant. However, as we consider more envious participants, the predictions across conditions diverge, culminating in the most envious participants following advice 24% less often when the adviser was highly-remunerated.

Regarding measures of stubbornness, the AME of our measure of susceptibility to the sunk-cost fallacy is significant and negative across all the specifications of Table 7. The AMEs reported suggest that a one-standard-deviation increase in susceptibility to the sunk-cost fallacy increases the probability of taking good advice by between 2.5-3.8 percentage points on average. Figure 4 provides the average adjusted predictions of the probability of accepting good advice in the domains of luck and skill, over the range of participants' susceptibility to the sunk cost fallacy. The difference between the least and most susceptible participants' in the propensity to follow good advice was 21 percentage points (79% for the least, 58% for the most) and in the skill task, 16 percentage points (84% for the least, 68% for the most).

In contrast, we did not find any predictive power of the general measure of stubbornness. This provides support for the notion that people may be unwilling to take good advice, not because they are stubborn per se, but because they are sensitive to the fact they used resources in order to form their position on a matter.

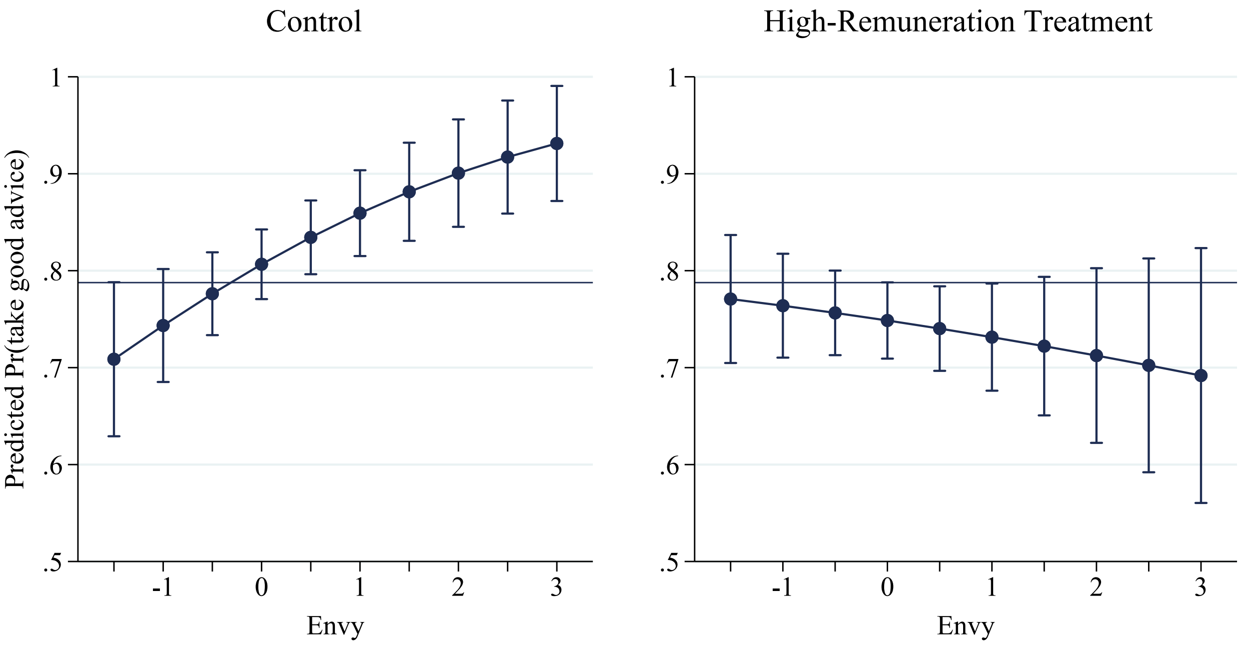
[21]Congruent with earlier results, the corresponding AME of envy for remuneration treatment 2 was not significantly different from the AME where all participants are assumed to be in the control condition.

Figure 2: The two sides of envy: average marginal effects



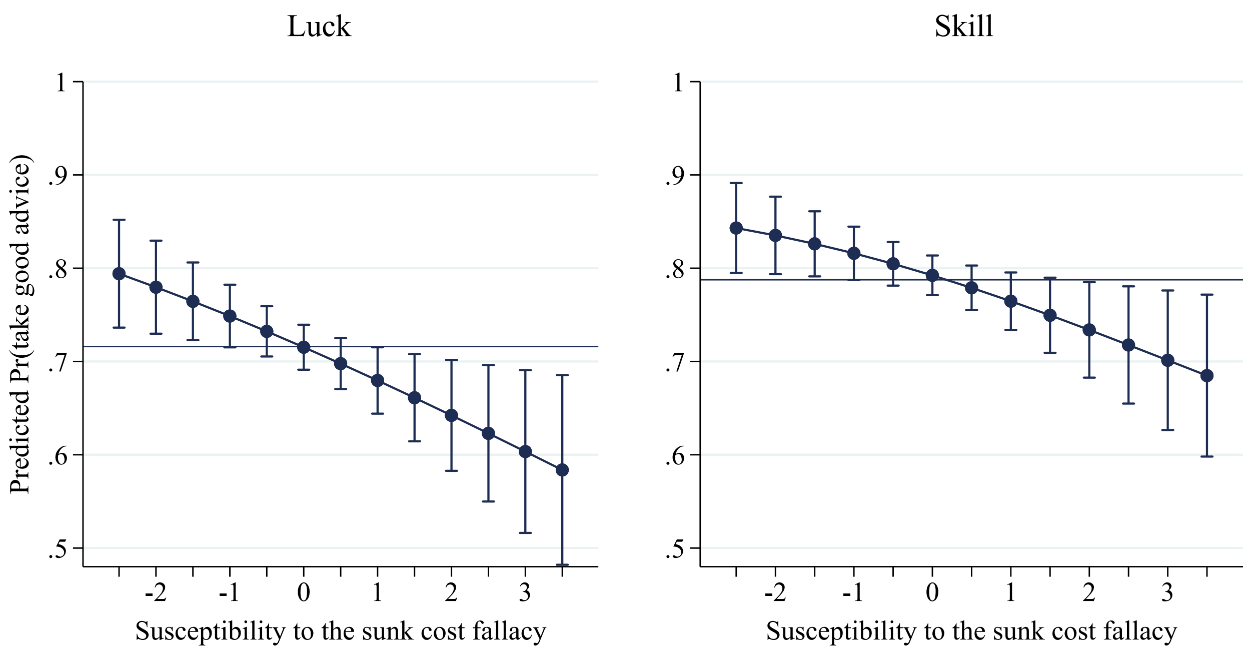
The average marginal effects shown were computed using the estimates from specification (6) of Table 7. 95% confidence intervals are shown. A contrast of AMEs test provides evidence ($P = 0.004$) that the AMEs shown are significantly different.

Figure 3: The two sides of envy: average adjusted predictions



The average adjusted predictions shown were computed using the estimates from specification (6) of Table 7 under the counter-factual assumptions that all participants have the dispositional envy level as shown on the x-axis, and were in the: control (left panel); high-remuneration treatment (right panel). The horizontal reference line is the average predicted value of the probability of accepting good advice where the prediction for each participant is generated using their actual (observed). The lowest (highest) observed standardized level of envy in our sample was -1.60 (3.27). To provide a smoother illustration we allow envy to take 0.5 standard-deviation increments between -1.5 and 3.0. 95% confidence intervals are shown.

Figure 4: Susceptibility to the sunk cost fallacy leads to a lower propensity to take good advice: average adjusted predictions



The average adjusted predictions shown in the left and right panels were computed using the estimates from specifications (3) and (6) of Table 7 respectively under the counter-factual assumptions that all participants have the susceptibility level as shown on the x-axis. The horizontal reference lines are the average predicted value of the probability of accepting good advice where each participant's prediction is generated using their actual (observed) data in the domain of luck and skill respectively. The lowest (highest) observed standardized level of susceptibility in our sample was -2.57 (3.57). To provide a smoother illustration we allow envy to take 0.5 standard-deviation increments between -2.5 and 3.5. 95% confidence intervals are shown.

6 Concluding Remarks

We have offered an analysis of whether, when and why people ignore good advice. We provided an abstracted experimental setting in which it was clear that the rational action was to accept good advice. However, we found good advice was frequently ignored; in our experiment, about 25% of the time. When good advice was more valuable, participants were more likely to accept it which suggests participants traded off rationality with other forces.

The literature documents that many practical and context-specific factors can play a role in determining the propensity to follow advice. In this paper, we provide evidence that two novel treatments have effects in a relatively context-free setting which allows us to interpret our results in a more general sense. In contrast to existing studies, we also relate whether people follow good advice to underlying psychological traits. We found that the fundamental human trait of envy played a major and varied role in determining whether good advice was followed. Our within-subject treatment revealed that good advice was followed less often when the adviser was more skillful than advisees, rather than luckier. Moreover, when the adviser was superior in skill, envy played a "positive" role: those with a higher dispositional level of envy were on average more likely to take good advice. In a between-subject treatment, we showed that good advice was followed less often when the adviser was highly remunerated, and there, a "negative" side of envy emerged: the positive association between envy and the propensity to take good advice was significantly lower, becoming insignificantly different from zero. We also showed that susceptibility to the sunk-cost fallacy was a robustly negatively associated to whether good advice was accepted suggesting an unwillingness to leave one's position, because it was costly to form. We measured this susceptibility through a novel scale we constructed, based on works such as Thaler (1999) and Arkes and Blumer (1985). In contrast, a more general scale of stubbornness was not found to have predictive power.

Envy has long been viewed as complex human trait. It is a latent characteristic; sitting dormant waiting to be activated. The Oxford English Dictionary defines envy as "the feeling of mortification and ill-will occasioned by the contemplation of superior advantages possessed by another". Important questions for social science are: what superior advantages possessed by another provoke envy and which decisions does envy then affect? To this end, our work suggests that comparisons of one's skill relative to another's, and one's remuneration relative to another's can lead envy to be activated and to affect the decision of whether to accept good advice from that other individual. In our experiment we also demonstrated both positive and negative sides of envy. When the adviser was superior in skill, more envious types were more likely to accept good advice, increasing their expected payoffs and reducing inequality between themselves and the adviser, which seems supportive of notions such as inequality aversion (e.g., Fehr and Schmidt, 1999).

However, we showed that other payoff-irrelevant factors, e.g., adviser remuneration, can substantially reduce such an effect. Our treatments were instrumental in showing envy's multifaceted role, but naturally leave many interesting questions open. For example, our study focused on the behavioral determinants of following good advice conditional on being an advisee. Therefore, we study the intensive margin of advice-taking, rather than the extensive margin i.e., the propensity to search for advice.

In summary, our results suggest that individuals may ignore advice even when the rational action is to take it, seemingly trading off rationality with other forces. More generally, the findings also suggest that whether advisers or experts are perceived to have reached their status by luck or skill, adviser attributes such as remuneration, and underlying fundamental psychological traits of advisees such as envy, could be key to understanding when and why good advice is ignored.

References

Alesina, A F, E L Glaeser, and B Sacerdote (2001), "Why doesn't the US have a European-style welfare state?" Harvard IER Discussion Paper 1933.

Arkes, H R and C Blumer (1985), "The psychology of sunk costs." Organizational Behavior and Human Decision Processes, 35, 124–140.

Banerjee, Dipak (1990), "Envy." In Economic Theory and Policy: Essays in Honor of Dipak Banerjee (Debraj Ray, Bhaskar Dutta, S Gangopadhyay, and Dilip Mookherjee, eds.), Oxford University Press.

Baumol, William J (1986), Superfairness: Applications and Theory. MIT Press.

Birnbaum, M H (1999), "Testing critical properties of decision making on the internet." Psychological Science, 10, 390–407.

Blanco, Mariana, Dirk Engelmann, and Hans Theo Normann (2011), "A within-subject analysis of other-regarding preferences." Games and Economic Behavior, 72, 321–338.

Brennan, Geoffrey (1973), "Pareto desirable redistribution: the case of malice and envy." Journal of Public Economics, 2, 173–183.

Brenner, Reuven (1987), Rivalry: In Business, Science, among Nations. Cambridge University Press.

Cain, Daylian M, George Loewenstein, and Don A Moore (2005), "The dirt on coming clean: Perverse effects of disclosing conflicts of interest." Journal of Legal Studies, 35, 1–25.

Calvert, Randall L (1985), "The value of biased information: A rational choice model of political advice." Journal of Politics, 47, 530–555.

Cook, John and Stephan Lewandowsky (2016), "Rational irrationality: Modeling climate change belief polarization using bayesian networks." Topics in Cognitive Science, 8, 160–179.

Davis, Milton S (1968), "Variations in patients' compliance with doctors' advice: an empirical analysis of patterns of communication." American Journal of Public Health and the Nations Health, 58, 274–288.

DellaVigna, Stefano and Devin Pope (2017), "What motivates effort? Evidence and expert forecasts." forthcoming: Review of Economic Studies.

Doberstein, Carey (2016), "Whom do bureaucrats believe? A randomized controlled experiment testing perceptions of credibility of policy research." Policy Studies Journal, 45, 384–405.

Doberstein, Carey (2017), "The credibility chasm in policy research from academics, think tanks, and advocacy organizations." Canadian Public Policy, 1–13.

Elster, Jon (1987), "Envy in economic life." In Strategy and Choice (Richard Zeckhauser, ed.), MIT Press.

Fehr, Ernst and Klaus M Schmidt (1999), "A theory of fairness, competition, and cooperation." Quarterly Journal of Economics, 114, 817–868.

Foley, Duncan K (1967), "Resource allocation and the public sector." Yale Economic Essays, 7, 45–98.

Gino, Francesca and Don A Moore (2007), "Effects of task difficulty on use of advice." Journal of Behavioral Decision Making, 20, 21–35.

Hilger, Nathaniel G (2016), "Why dont people trust experts?" Journal of Law and Economics, 59, 293–311.

John, Oliver P and Sanjay Srivastava (1999), "The big five trait taxonomy: History, measurement, and theoretical perspectives." Handbook of Personality: Theory and Research, 2, 102–138.

Kardas, Przemyslaw, Pawel Lewek, and Michal Matyjaszczyk (2013), "Determinants of patient adherence: a review of systematic reviews." Frontiers in Pharmacology, 4.

Kirchsteiger, Georg (1994), "The role of envy in ultimatum games." Journal of Economic Behavior & Organization, 25, 373–389.

Kuziemko, Ilyana, Michael I Norton, Emmanuel Saez, and Stefanie Stantcheva (2015), "How elastic are preferences for redistribution? Evidence from randomized survey experiments." American Economic Review, 105, 1478–1508.

Lachapelle, Erick, Éric Montpetit, and Jean-Philippe Gauvin (2014), "Public perceptions of expert credibility on policy issues: The role of expert framing and political worldviews." Policy Studies Journal, 42, 674–697.

Leibbrandt, Andreas and Raúl López-Pérez (2012), "An exploration of third and second party punishment in ten simple games." Journal of Economic Behavior & Organization, 84, 753–766.

Morris, Peter A (1974), "Decision analysis expert use." Management Science, 20, 1233–1241.

Mui, Vai-Lam (1995), "The economics of envy." Journal of Economic Behavior & Organization, 26, 311–336.

Mujcic, Redzo and Andrew J Oswald (2017), "Is envy harmful to a societys psychological health? A longitudinal study of 18,000 adults."

Mullainathan, Sendhil, Markus Noeth, and Antoinette Schoar (2012), "The market for financial advice: An audit study." National Bureau of Economic Research.

Önkal, Dilek, Paul Goodwin, Mary Thomson, Sinan Gönül, and Andrew Pollock (2009), "The relative influence of advice from human experts and statistical methods on forecast adjustments." Journal of Behavioral Decision Making, 22, 390–409.

Paolacci, Gabriele and Jesse Chandler (2014), "Inside the Turk: Understanding Mechanical Turk as a participant pool." Current Directions in Psychological Science, 23, 184–188.

Paolacci, Gabriele, Jesse Chandler, and Panagiotis G Ipeirotis (2010), "Running experiments on Amazon Mechanical Turk." Judgment and Decision Making, 5, 411–419.

Parrott, W G and R H Smith (1993), "Distinguishing the experiences of envy and jealousy." Journal of Personality and Social Psychology, 64, 906–920.

Ronayne, David and Daniel Sgroi (2017), "When good advice is ignored: The role of adviser remuneration." AEA RCT Registry, June 28, URL `https://www.socialscienceregistry.org/trials/2022/`.

Russell, Bertrand (1930), The Conquest of Happiness. H. Liverwright.

Smith, R H, W G Parrott, E F Diener, R H Hoyle, and S H Kim (1999), "Dispositional envy." Personality and Social Psychology Bulletin, 25(8), 1007–1020.

Swol, Lyn M and Janet A Sniezek (2005), "Factors affecting the acceptance of expert advice." British Journal of Social Psychology, 44, 443–461.

Thaler, R H (1999), "Mental accounting matters." Journal of Behavioral Decision Making, 12, 183–206.

Van de Ven, N, M Zeelenberg, and R Pieters (2009), "Leveling up and down: The experiences of benign and malicious envy." Emotion, 9, 419–29.

Varian, Hal R (1974), "Equity, envy, and efficiency." Journal of Economic Theory, 9, 63–91.

Weizsäcker, Georg (2010), "Do we follow others when we should? A simple test of rational expectations." American Economic Review, 100, 2340–2360.

Wilkins, M M (2015), "Signs that you're being too stubborn." Harvard Business Review, May 21, https://hbr.org/2015/05/signs–that–youre–being–too–stubborn.

Winkelmann, Rainer (2012), "Conspicuous consumption and satisfaction." Journal of Economic Psychology, 33, 183–191.

Yaniv, Ilan and Eli Kleinberger (2000), "Advice taking in decision making: Egocentric discounting and reputation formation." Organizational Behavior and Human Decision Processes, 83, 260–281.

Zizzo, Daniel John (2010), "Experimenter demand effects in economic experiments." Experimental Economics, 13, 75–98.

Zizzo, Daniel John and Andrew J Oswald (2001), "Are people willing to pay to reduce others' incomes?" Annales d'Économie et de Statistique, 63/64, 39–65.

Appendix

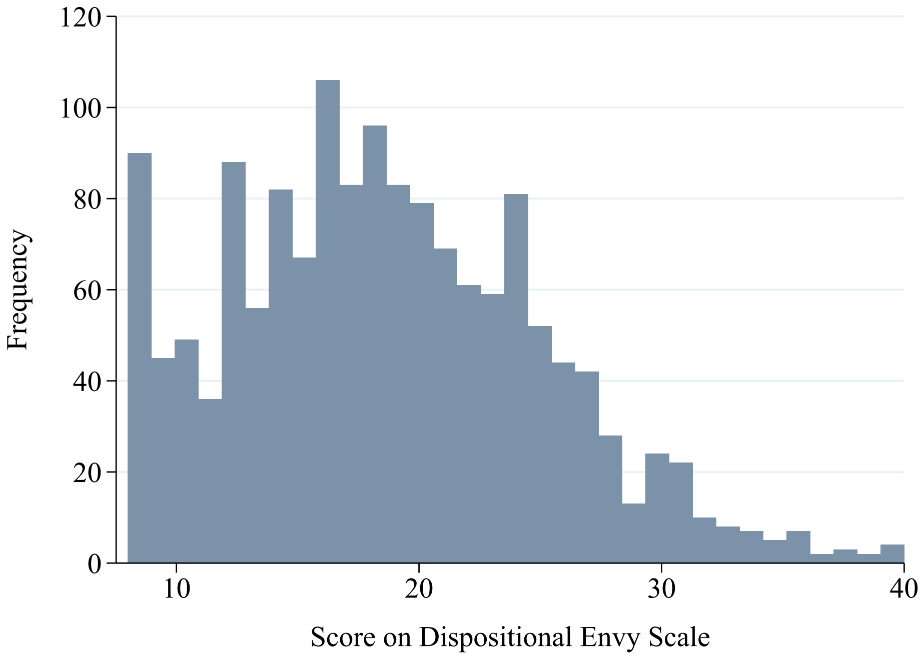
The Appendix provides demographic information on the participants of the main wave (Table A1); the distribution of scores obtained from participants' responses to three psychometric scales we used (Figures A1-A3); coefficient estimates of the probit coefficients underlying the regressions of Table 7 (Table A2); and a copy of the experiment's transcript (modified for readability).

Table A1: Participant Demographics from Main Wave

Characteristic	
Gender	
Male	701 (47)
Female	794 (53)
Other	4 (0)
Prefer not to say	4 (0)
Age, mean years [sd]	35.7 [11.7]
18-25	284 (19)
26-30	326 (22)
31-40	484 (32)
41-50	210 (14)
51+	199 (13)
Race	
White	1,105 (74)
Black or African American	125 (8)
Hispanic or Latino	96 (6)
American Indian or Alaska Native	11 (1)
Asian American	130 (9)
Native Hawaiian or Pacific Islander	6 (0)
Other	30 (2)
Income[a]	
0 – 9.999	77 (5)
10 – 19.999	158 (11)
20 – 29.999	197 (13)
30 – 39.999	215 (14)
40 – 49.999	175 (12)
50 – 59.999	167 (11)
60 – 69.999	127 (8)
70 – 79.999	113 (8)
80 – 89.999	54 (4)
90 – 99.999	53 (4)
100 – 124.999	85 (6)
125 – 149.999	38 (3)
150+	44 (3)
Education	
No schooling	1 (0)
Nursery, kindergarten, and elementary (grades 1-8)	1 (0)
High school (grades 9-12, no degree)	23 (2)
High school graduate (or equivalent)	160 (11)
Some college (1-4 years, no degree)	546 (36)
Bachelors degree (BA, BS, AB, etc)	591 (39)
Masters degree (MA, MS, MENG, MSW, etc)	139 (9)
Professional school degree (MD, DDC, JD, etc)	26 (2)
Doctorate degree (PhD, EdD, etc)	16 (1)
Political Affiliation[b] [sd]	38.7 [29.6]
N	1,503

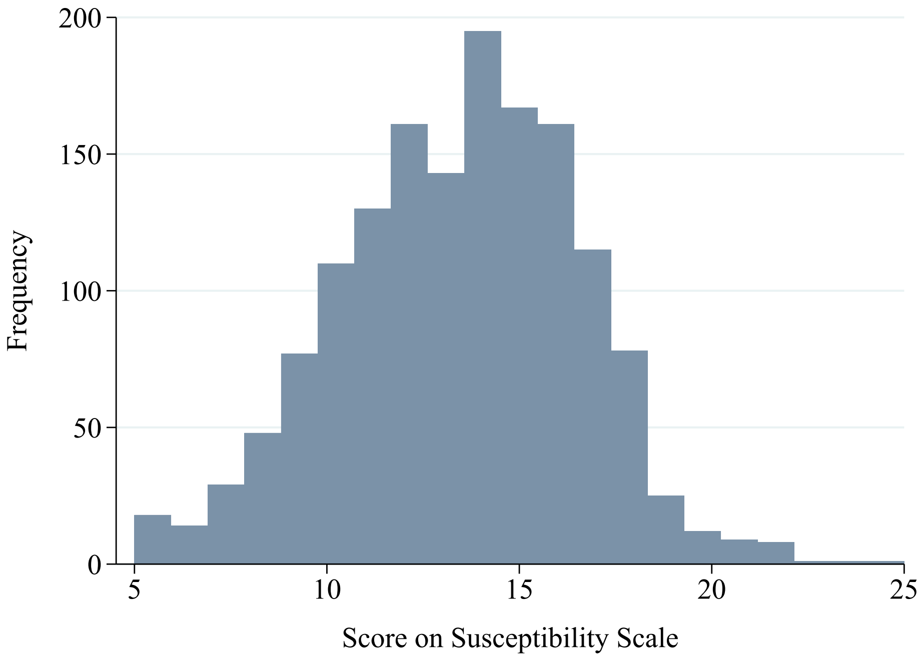
Frequencies; (% within characteristic); [standard deviation]
[a] Household annual pre-tax income in '000 USD
[b] 0 = "Entirely Liberal"; 100 = "Entirely Conservative"

Figure A1: Envy



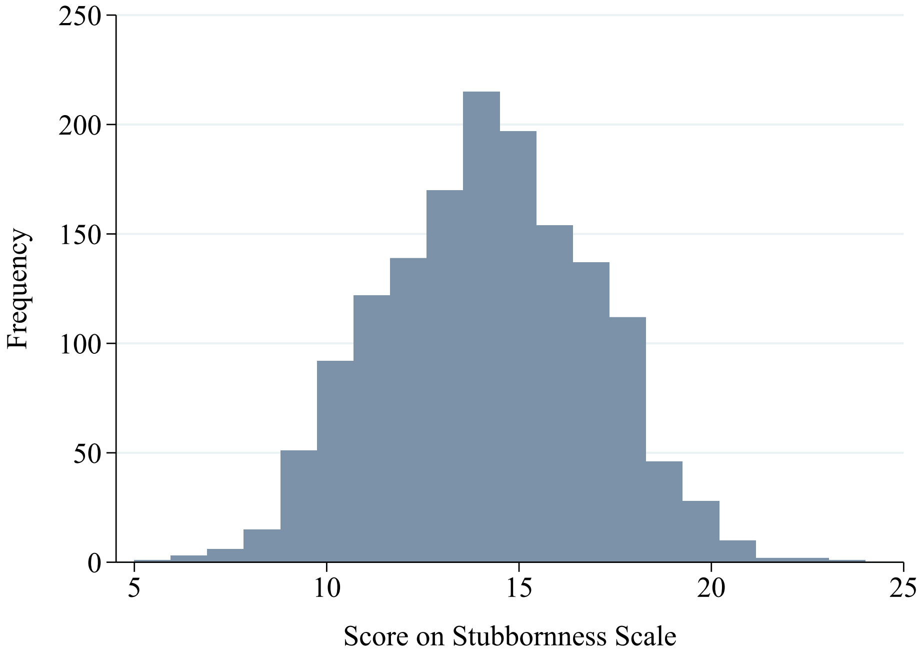
Participants' scores from the dispositional envy scale of Smith et al. (1999) that comprises of 8 statements to which the subject must select how much they agree with them on a 5-point Likert scale from "Strongly disagree" to "Strongly agree". Before summing, the responses to each question are coded from 1-5 such that the higher the score, the higher the level of dispositional envy, $N = 1,503$.

Figure A2: Susceptibility to the Sunk Cost Fallacy



Participants' scores from our susceptibility scale that comprises of 5 statements to which the subject must select how much they agree with them on a 5-point Likert scale from "Strongly disagree" to "Strongly agree". Before summing, the responses to each question are coded from 1-5 such that the higher the score, the higher the level of susceptibility to the sunk cost fallacy, $N = 1,503$.

Figure A3: Stubbornness



Participants' scores from the responses to the stubbornness criteria listed inWilkins (2015) that comprises of 5 statements to which the subject must select how much they agree with them on a 5-point Likert scale from "Strongly disagree" to "Strongly agree". Before summing, the responses to each question are coded from 1-5 such that the higher the score, the higher the level of stubbornness, $N = 1,503$.

Table A2: Probit Coefficients Behind AMEs of Table 7

Coefficients	Luck			Skill		
$y = \mathbb{I}$ (took good advice)	(1)	(2)	(3)	(4)	(5)	(6)
Envy	0.017	0.019	0.047	0.089	0.078	0.269
	(0.041)	(0.042)	(0.071)	(0.041)	(0.042)	(0.073)
Sunk-Cost Fallacy (SCF)	-0.119	-0.108	-0.105	-0.096	-0.091	-0.096
	(0.039)	(0.041)	(0.042)	(0.038)	(0.040)	(0.041)
Stubbornness	-0.016	-0.021	-0.022	-0.045	-0.044	-0.048
	(0.042)	(0.042)	(0.042)	(0.041)	(0.041)	(0.041)
Envy × SCF		0.040	0.037		0.085	0.075
		(0.038)	(0.038)		(0.037)	(0.037)
Envy × Stubbornness		-0.029	-0.029		-0.001	0.003
		(0.036)	(0.036)		(0.035)	(0.034)
SCF × Stubbornness		-0.003	0.001		-0.105	-0.103
		(0.040)	(0.040)		(0.037)	(0.037)
Envy × SCF × Stubbornness		-0.010	-0.008		-0.024	-0.025
		(0.026)	(0.026)		(0.025)	(0.025)
Remuneration T1			-0.184			-0.202
			(0.092)			(0.093)
Remuneration T1 × Envy			-0.072			-0.269
			(0.091)			(0.092)
Remuneration T2			-0.044			-0.012
			(0.094)			(0.096)
Remuneration T2 × Envy			0.010			-0.072
			(0.096)			(0.097)
Value dummies	X	X	X	X	X	X
Participant demographics	X	X	X	X	X	X
Observations	1,317	1,317	1,317	1,427	1,427	1,427

Coefficient estimates from probit regressions are shown with robust standard errors in parentheses. All specifications pass various mis-specification tests of functional form and heteroskedasticity. The binary dependent variable is 1 (0) if the decision made was to accept (ignore) the good advice. "Value dummies" refers to the inclusion of a dummy for every level of value (bar one, to serve as the reference category). Each participant decided whether to take the advice in each task, hence there were 1,503 decisions for each task. Of those in the skill (luck) task, 68 (181) were excluded because they out-performed the adviser i.e., advice was bad. Of the remaining participants, 8 (4) were excluded because they did not disclose their sex. Finally, there was only one respondent with a value of 7 in the luck task, who was dropped for the analysis.

Transcript (modified for inclusion in the Appendix)

Participation Agreement

You have been invited to take part in a research study run by researchers at the University of Oxford. Please read the following statements carefully.

<u>Our Commitments and Privacy Policy</u>

- We never deceive participants. For example, if we inform you that another participant is making a choice on which you can then react, this is indeed the case.
- We keep our promises made to participants. For example, if we promise a certain payment, participants will indeed receive it.
- In the event that we are responsible for a mistake that is to the disadvantage of participants, we will inform and compensate the respective participants.
- We design, conduct and report our research in accordance with recognised scientific standards and ethical principles.

We adhere to the terms of our privacy policy as stated below.

- The data in the participants' database will only be used for the purpose of the study.
- There is no link between the personal data in the participants' database and the data collected during a study.
- The generated anonymous data will be used for analysis. The end product will be publicly available.
- Your participation in this study is purely voluntary, and you may withdraw your participation or your data at any time without any penalty to you.
- Please be aware that Amazon user information, connected to Mturk worker IDs, can be visible to the public, depending on the privacy settings of your Amazon.com account. See also https://www.mturk.com/mturk/privacynotice for further information on Amazon.com's privacy policies.

There are no known risks associated with your participation in this research beyond those of everyday life. If there is anything about the study or your participation that is unclear or that you do not understand, if you have questions or wish to report a research-related problem, you may contact the Requester via MTurk. For questions about your rights as a research participant, you may contact The Center for Experimental Social Science, Nuffield College, University of Oxford, 3 George Street Mews, Oxford OX1 2AA, at cess@nuffield.ox.ac.uk.

○ I agree

Coins: Task & Bonus Instructions

You will face ten timed questions where you are simply asked to guess whether a fair two-sided coin landed with "Heads" or "Tails" facing up.

You have ten seconds to answer each question. If you are happy with your answer and wish to move on to the next question before the ten seconds are up, simply hit the blue ">>" button at the bottom of the screen. If the time runs out and you have selected an answer, that answer will be submitted and you will move on automatically to the next question. If the time runs out and you have not selected any answer, that question will be marked as incorrect and you will move on automatically to the next question. Once you have moved on from a question, you cannot go back to it.

If you do especially well in this task you will be awarded a bonus payment of at least $0.50. Bonuses will be paid after the required number of workers have completed the HIT. Those who do not win a bonus will not be notified.

Before you take the quiz of ten questions, you first have two practice questions. These will allow you to get a feel for the format and time limit. They do not count for the bonus payments. The first practice question will begin immediately on the next page.

- ○ I understand these instructions

Coin Toss (Practice x2)

One coin is tossed. Guess which side landed face up:

- ○ Heads
- ○ Tails

Coin Tosses Practice Feedback

You scored <<their score>> out of 2 in the practice.

First coin-toss
Your guess: <<their guess>>

The coin showed: Tails

Second coin-toss
Your guess: <<their guess>>
The coin showed: Heads

The real questions will **begin immediately** on the next page. Make sure you are ready.

Coin Tosses (x10)

One coin is tossed. Guess which side landed face up:

- ○ Heads
- ○ Tails

Logic Puzzles: Task & Bonus Instructions

You will face ten timed multiple choice questions about logic.

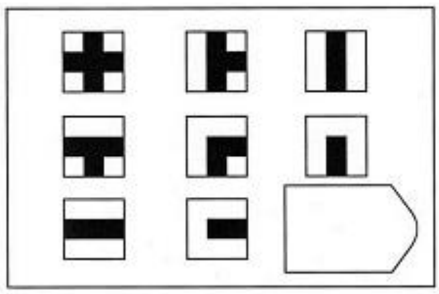
Each question shows a sequence of nine patterns with one missing. Your task is to select the missing pattern from the drop-down list. There is only one correct answer for each question. You have 30 seconds to answer each question. If you are happy with your answer and wish to move on to the next question before the 30 seconds are up, simply hit the blue ">>" button at the bottom of the screen. If the time runs out and you have selected an answer, that answer will be submitted and you will move on automatically to the next question. If the time runs out and you have not selected any answer, that question will be marked as incorrect and you will move on automatically to the next question. Once you have moved on from a question, you cannot go back to it.

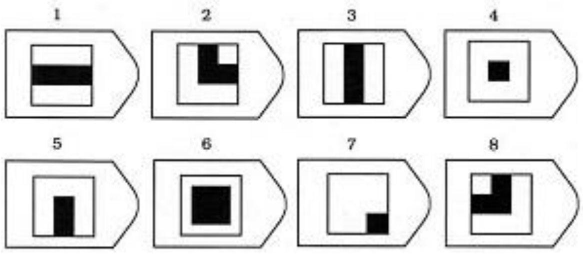
If you do especially well in this task you will be awarded a bonus payment of at least $0.50. Bonuses will be paid after the required number of workers have completed the HIT. Those who do not win a bonus will not be notified.

Before you take the task of ten questions, you first have a practice question. This will allow you to get a feel for the format and time limit. It does not count for the bonus payments. The practice question will begin immediately on the next page.

○ I understand these instructions

Logic Puzzle Practice





(Dropdown menu of 1-8.)

Logic Puzzles Practice Feedback

You got the practice question correct!
[You got the practice question incorrect. Hopefully you will have better luck with the next questions.]

The real task will **begin immediately** on the next page. Make sure you are ready.

Logic Puzzles (x10)

<<Ten puzzles similar in nature to the practice were presented, each with 8 possible answers>>

Description of Other Worker

Before we reveal your scores...

... we here describe the experiences of two other workers who, some time ago, completed the exact same tasks you have just tried for the same $2.00 HIT reward. Please pay attention as we will be asking you some **comprehension questions** about them on the next screen.

These two workers both saw the same instructions you did. This means they were both told that they could receive at least $0.50 for doing especially well in a task. In addition, they each did especially well in one of the tasks: one scored 7/10 in the coin task, and got [If not in remuneration treatment 1: “a bonus of $0.50”] [If in remuneration treatment 1: “a large bonus of $100”]; the other scored 9/10 in the logic puzzles task, and [If not in remuneration treatment 1: “got a bonus of $0.50”] [If in remuneration treatment 1: “also got a large bonus of $100”].

○ I am ready for comprehension questions on the text above

Comprehension Questions

What did the two workers who did the tasks some time ago know about the potential bonus payment for doing "especially well" in a task **before** they started it?

○ It would be at least $0.50

○ It would be exactly $0.50

What bonus payment did both the workers actually get for doing especially well?

○ $0.50

○ $100.00

A Chance for a Bonus

We are not going to ask you to repeat the tasks. On this page we explain how the answers you gave in each task translate into your chances of winning a bonus.

We will choose ten workers at random. If you are chosen, we will pick one question at random, and if you got it right, you will win a bonus. Let's look at your scores:

You got <<their coins score>>/10 in the coin task
You got <<their logic score>>/10 in the logic puzzles task

That means if you are chosen and we pick one of the coin questions, there is a <<their coins score>> in 10 chance of you winning the bonus. Similarly, if you are chosen and we pick one of the logic puzzle questions, there is a <<their logic score>> in 10 chance of you winning the bonus.

But before we go ahead, we would like to give you an opportunity to perhaps boost your odds of getting the bonus. Remember those two other workers we described earlier who did especially well? If you like, instead of us using your answers when we check if you have won the bonus we will look at their answers: that means your chances of getting the bonus would be 7 in 10 if we pick from the coin task or 9 in 10 from the logic puzzles task.

Regarding payment: The other worker got a bonus of $0.50. In your case, the bonus you might win is also $0.50. [If in remuneration treatment 1: “The other worker got a large bonus of $100.00. In your case however, the bonus you might win is $0.50.”] [If in remuneration treatment 2: “The other worker got a bonus of $0.50. In your case, the bonus you might win is also $0.50. Additionally, if you decide to use another worker's answers, they will get a further bonus of $0.25.”]

So, would you like us to use the answers you already gave, or the answers the other worker gave when we check whether you have won a bonus?

For the coin task:

- ◯ Keep my answers
- ◯ Use the other worker's answers

For the logic puzzles task:

- ◯ Keep my answers
- ◯ Use the other worker's answers

Comprehension Questions

What did the two workers who did the tasks some time ago know about the potential bonus payment for doing "especially well" in a task before they started it?

- ◯ It would be at least $0.50
- ◯ It would be exactly $100.00

What bonus payment did both the workers actually get for doing especially well?

- ◯ $0.50
- ◯ $100.00

Final Questions (page 1 of 2): Demography

What is your sex?

- ◯ Male
- ◯ Female
- ◯ Other
- ◯ Prefer not to say

What is your age?

What is your race?

- ◯ White
- ◯ Black or African American
- ◯ Hispanic or Latino
- ◯ American Indian or Alaska Native
- ◯ Asian American
- ◯ Native Hawaiian or Pacific Islander
- ◯ Other

What is your household's annual income? (US dollars, before tax)

- 0-9,999
- 10,000 - 19,999
- 20,000 - 29,999
- 30,000 - 39,999
- 40,000 - 49,999
- 50,000 - 59,999
- 60,000 - 69,999
- 70,000 - 79,999
- 80,000 - 89,999
- 90,000 - 99,999
- 100,000 - 124,999
- 125,000 - 149,999
- 150,000 +

What is the highest grade of school you have completed, or the highest degree you have received?

- No schooling (or less than 1 year)
- Nursery, kindergarten, and elementary (grades 1-8)
- High school (grades 9-12, no degree)
- High school graduate (or equivalent)
- Some college (1-4 years, no degree)
- Bachelor’s degree (BA, BS, AB, etc)
- Master’s degree (MA, MS, MENG, MSW, etc)
- Professional school degree (MD, DDC, JD, etc)
- Doctorate degree (PhD, EdD, etc)

Generally speaking, which point on this scale best describes your political affiliation?
(A slider was presented with range [0,100] with “Entirely Liberal” over 0 and “Entirely Conservative” over 100.)

What is your Mturk ID? (please copy and paste it to avoid typos)

Final Questions (page 2 of 2): Personality

Please respond to the statements below using the scales provided: (each scale was a 5-point Likert scale with “Strongly disagree”, “Moderately disagree”, “Neither agree nor disagree”, “Moderately agree” and “Strongly agree”.)

(Dispositional Envy Scale)
I feel envy every day.
The bitter truth is that I generally feel inferior to others.
It doesn't frustrate me to see some people succeed easily.
Feelings of envy rarely torment me.
No matter what I do, envy always plagues me.
I am rarely troubled by feelings of inadequacy.
It somehow doesn’t seem fair that some people seem to have all the talent.
The success of my neighbors doesn't make me resent them.

(Susceptibility to the Sunk-Cost Fallacy Scale)
You have invested a good deal of your time into a project and it is failing. You have the option to start on something different that you now know is more likely to be successful but you know you cannot get the time back that you spent on the project so you decide to keep going with it.

You have an investment strategy that you have developed over several months. It is not working and you are losing money. There is no way for you to recover the lost effort put in to developing the strategy but you decide that it is better to start afresh anyway.

Imagine that you have spent $20 on a ticket to a concert. The day of the concert comes and unfortunately it is snowing heavily, and you feel tired after a tough day. You know you would not have decided to go to the concert if you hadn’t already bought the ticket, but you also know that you cannot get a refund. On balance you decide not to go to the concert.

You are staying in a hotel room, and you have just paid $6.95 to see a movie on pay TV. You find that you are bored 5 minutes into the movie and that the movie seems pretty bad. You decide that since you cannot get a refund you might as well continue watching the movie.

Your relationship with your partner is not going well. You have reasoned it out and you have realized that if you knew how it would go when you started the relationship you would not have gone through with it. You have the opportunity to break up but since you have been together for many months you decide to keep going.

(Stubbornness Scale)

I do something I want to do even if no one else wants to do it.

I never keep at an idea (or plan) when I know I am wrong.

When others present an idea, I tend to point out all the reasons it won’t work.

I agree to or commit half-heartedly to others’ requests, when I know all along that I’m going to do something entirely different.

I visibly feel anger, frustration, or impatience when others try to persuade me of something I don’t agree with.