

Ambuehl, Sandro

Working Paper

An Offer You Can't Refuse? Incentives Change What We Believe

CESifo Working Paper, No. 6296

Provided in Cooperation with:

Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Ambuehl, Sandro (2017) : An Offer You Can't Refuse? Incentives Change What We Believe, CESifo Working Paper, No. 6296, Center for Economic Studies and ifo Institute (CESifo), Munich

This Version is available at:

<http://hdl.handle.net/10419/155538>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



Working Papers

www.cesifo.org/wp

An Offer You Can't Refuse? Incentives Change What We Believe

Sandro Ambuehl

CESIFO WORKING PAPER NO. 6296

CATEGORY 13: BEHAVIOURAL ECONOMICS

JANUARY 2017

An electronic version of the paper may be downloaded

- from the SSRN website: www.SSRN.com
- from the RePEc website: www.RePEc.org
- from the CESifo website: www.CESifo-group.org/wp

ISSN 2364-1428

CESifo

Center for Economic Studies & Ifo Institute

An Offer You Can't Refuse? Incentives Change What We Believe

Abstract

Much of economics assumes that higher incentives increase participation in a transaction only because they exceed more people's reservation price. This paper shows theoretically and experimentally that when information about the consequences is costly, higher incentives also change reservation prices to further increase participation. A higher incentive makes people gather information in a way that is more favorable to participation—as if they were persuading themselves to participate. Hence, incentives change not only what people choose, but also what they believe their choices entail. This result informs the debate about laws around the world that severely restrict incentives for transactions such as organ donation, surrogate motherhood, human egg donation, and medical trial participation. It helps bridge a gap between economists on the one hand and the policy makers and ethicists on the other.

JEL-Codes: D030, D040, D840.

Keywords: incentives, repugnant transactions, information acquisition, inattention, experiment.

Sandro Ambuehl
UTSC Department of Management &
Rotman School of Management
University of Toronto
105 St George Street
Canada – Toronto, ON M5S 3E6
sandro.ambuehl@utoronto.ca

October 14, 2016

This is a thoroughly revised version of my job market paper I wrote at Stanford University. I am deeply indebted to my advisors Muriel Niederle, B. Douglas Bernheim, and Alvin E. Roth. I am very grateful to Debra Satz for many discussions about philosophers' work on the ethics of incentives. I am extremely grateful to P.J. Healy and to Yan Chen for letting me use their experimental economics laboratories. This work benefitted from countless helpful comments by seminar audiences at CERGE-EI, the University of Melbourne, the University of Toronto, the Wharton School at the University of Pennsylvania, Carnegie Mellon University, the University of California in Los Angeles, the University of Cologne, the University of Munich, Stanford University, as well as at the Behavioral Economics Annual Meetings, the Congress of the European Economic Association, and the Bay Area Behavioral and Experimental Economics Workshop. I thank Constantine Yannelis for narrating the videos. This work has been approved by the Stanford IRB in protocols 29615 (online experiments) and 34001 (laboratory experiments). Funding from the Stanford Department of Economics is gratefully acknowledged.

Offers of large sums of money ... could lead some prospective participants to enroll ... when it might be against their better judgment and when otherwise they would not do so.

- [National Bioethics Advisory Commission \(2001\)](#)

Payments to women providing oocytes should be ... not so substantial that they ... will lead donors to discount risks. ... The higher the payment, the greater the possibility that women will discount risks.

- [American Society for Reproductive Medicine \(2007\)](#)

1 Introduction

Much of economics assumes that higher incentives increase participation in a transaction *only* because they exceed the reservation price of a larger number of people. This paper shows both experimentally and theoretically that incentives may also *change* reservation prices in a way that further *increases* participation. As incentives rise, people search and interpret information about the transaction differently. They do so in a way that is more favorable to participation, as if they were trying to persuade themselves to participate. Thus, incentives do not only change what people choose, but also what they believe their choices entail—even before they make a decision.

Policy makers and ethicists have long intuited such effects, and have viewed them as an indication of deficient decision making ([Kanbur, 2004](#); [Satz, 2010](#); [Grant, 2011](#); [Sandel, 2012](#)). This intuition is one major reason behind laws around the world that tightly restrict whether and how transactions such as organ donation, surrogate motherhood, egg donation, medical trial participation and others can be incentivized. Indeed, worries concern incentives, not the activities *per se*—altruistic participation is often applauded.¹ The welfare costs of restricting incentives are potentially enormous. For instance, [Held et al. \(2016\)](#) estimate that the ban on incentives for kidney donation is responsible for the premature death of up to ten thousand Americans on the waiting list each year.² Hence, whether and how incentives change the quality of decision making is a critical question.

The debate has suffered from the lack of two crucial pieces of information. First, is it true *empirically* that incentives change what people believe their choice options entail, and do they do so in the way policy makers and ethicists are concerned about? Second, is such behavior really an indication of unsound decision making, or is it merely an expression of Bayesian rationality? These questions are of independent interest. If incentives indeed alter people’s beliefs about their choice options, incentives do not merely change choices, but genuinely persuade. Implications then extend to all kinds of situations in which information acquisition and incentives interact, in fields as diverse as finance, marketing, and political economy.

¹[Satz \(2010\)](#) explicitly mentions “a paternalistic concern that sellers would actually be harmed by the sale of their organs, but that ... they would sell their organs if it were legal.”

²The same authors estimate that restrictions on incentives for kidney donation cause taxpayers annual costs of \$12 billion because current treatments for kidney disease are more expensive than the presumed costs of a donor kidney.

In this paper I answer both questions, in three parts. First, an experiment demonstrates that higher incentives cause people to acquire and interpret information about a transaction in a way that is more favorable to participation. The experiment emphasizes ecological validity, and thus uses an unfamiliar, visceral transaction that is, in select aspects, similar to those for which incentives are restricted by law. It features two conditions. In the control condition, subjects cannot acquire information about the transaction. Hence, incentives increase participation only because they exceed the reservation price of a larger number of people. In the treatment condition, subjects can gather information, perhaps selectively, after learning the incentive they will be offered for participating. Hence, incentives may *change* reservation prices. Indeed, in that condition, as incentives rise, subjects become more interested in encouraging information about the transaction, and less keen to learn about the risks and downsides. Importantly, the supply of people willing to participate in the transaction responds more strongly to an increase in the incentive amount if endogenous information acquisition is possible than if it is not. Hence, the change in information demand affects participation decisions; it does not merely serve to justify a choice subjects would have made anyway. A direct measurement of reservation prices confirms that higher incentives deflate them, *ceteris paribus*, reflecting changed expectations about what the transaction entails. Higher incentives cause subjects to successfully persuade themselves to participate.

Second, a model of costly information acquisition shows that this perhaps worrisome-looking behavior is, in principle, entirely consistent with Bayesian rationality. When the consequences of a transaction are uncertain, and information about them is costly, agents will typically decide whether to participate based on incomplete information. Their choice may be different from the one they would have made under complete information, akin to type I and type II errors in classical statistics. Optimal information demand reflects the costs of these errors. Because an increase in the incentive for participation changes these costs, it changes the *kind* of information the Bayesian demands, as follows. With a higher incentive for participation, type I errors (mistaken participation) are less costly, because the additional payment partially insurances against *ex post* undesirable outcomes. Simultaneously, type II errors (mistaken abstention) are more costly because the opportunity cost of abstention is higher. In order to better eschew type II errors, and potentially make more type I errors, the information the Bayesian chooses to consult will be more favorable to participation, both in states of the world in which this is warranted (fewer type II errors) and in those states in which it is not (more type I errors). Consequently, he will have different beliefs about what the transaction entails. Hence, even a Bayesian may look as if he were trying to talk himself into participation. This mechanism applies whenever the stakes of the decision and the costs of information acquisition are of comparable magnitude, so it should be expected both in the limited-stakes experiments in this paper, and in more consequential decisions in which acquiring encompassing information is comparatively more costly, such as organ donation.³

³To illustrate, most people do not acquire a degree in nephrology before deciding whether to donate a kidney.

Third, an experiment in a stylized setting conceptually replicates the first experiment, but allows for an explicit test of rationality. After all, the mere consistency of a behavioral pattern with Bayesian rationality does not, by itself, imply that humans avoid (non-Bayesian) decision errors. But whether they do is a crucial determinant of the welfare effects of increased incentives.

I find that empirical behavior follows the theoretical predictions. Higher incentives make subjects gather information that is more encouraging, and the same increase in the incentive amount induces a larger increase in participation when endogenous information acquisition is possible than when it is not. Importantly, elicited and objective posteriors align closely, even though information is presented in a fashion that makes it impossible to calculate individual-level posterior beliefs explicitly. This indicates a large degree of rationality. But to the extent that people do deviate from rationality, their behavior may be concerning to ethicists and policy makers—higher incentives make subjects overly optimistic about the consequences of participation.

Overall, these results show that incentives do not only change what people choose, but also how they evaluate their choice options. Because they make people gather and interpret information in a way that is more favorable to the incentivized choice, this behavior may appear worrisome. It is, however, largely consistent with Bayesian rationality, both in theory, and in controlled experiments.

The paper contributes to several literatures. It bridges a gap between disciplines by using standard economic methodology to inform a concern about the effects of incentives that is both widely-held in the applied ethics literature and highly influential in policy.⁴ It thus contributes to a burgeoning literature on *repugnant transactions* that characterizes people’s motives for preventing others from engaging in voluntary transactions (Kahneman, Knetsch and Thaler, 1986; Basu, 2003, 2007; Roth, 2007; Leider and Roth, 2010; Niederle and Roth, 2014; Ambuehl, Niederle and Roth, 2015; Elias, Lacetera and Macis, 2015a,b). In contrast to that literature, the present paper focuses on a specific such motive and examines whether its underpinnings stand empirical and theoretical scrutiny.

It forges a new connection between a vast literature on motivated reasoning (see Kunda (1990), Klayman (1995), and Epley and Gilovich (2016) for reviews) and formal economic theory. By capturing the phenomenon in a rational model with standard preferences, it also contributes to work in behavioral economics on this topic (see Gino, Norton and Weber (2016); Bénabou (2015); Bénabou and Tirole (2016) for reviews).

The paper also contributes to multiple literatures within economics. The mechanism it identifies is an important comparative static implication of the theory of rational inattention (Sims, 2003, 2006; Woodford, 2012, 2014; Martin, 2014; Caplin and Martin, 2014; Yang, 2014; Matějka and McKay, 2015; Caplin and Dean, 2015; see Caplin (forthcoming) for a review), and the experiment in section 5 is an explicit test thereof. The paper more generally relates to a theoretical literature on endogenous information acquisition by Bayesians (Caplin and Leahy, 2001; Bénabou and Tirole, 2002, 2011; Suen,

⁴By focusing on a single mechanism, this paper cannot take a stance on whether incentives for any particular transaction such as living organ donation should be limited. The policy implications are more nuanced; they are discussed in section 6.

2004; Koszegi, 2006; Eliaz and Spiegler, 2006; Gentzkow and Kamenica, 2011). Like that literature, it relies on the fact that a Bayesian decision maker whose utility function is non-linear in posterior beliefs will have non-trivial preferences over information structures, and characterizes these preferences.

The paper shows that incentives change how subjects acquire and interpret *external* information about a transaction in a way that *amplifies* their effect.⁵ It thus complements a literature that studies the inferences subjects draw from the incentive *per se* (either rationally, as in Kamenica, 2008; Cryder et al., 2010; Bénabou and Tirole, 2006, or irrationally, as in the literature on anchoring, Ariely, Loewenstein and Prelec (2003), see Maniadis, Tufano and List (2014) for a review). There, reservation prices typically change in a way that *diminishes* the effect of incentives; for instance because high incentives may signal that the transaction is unattractive. The first experiment in this paper shows that people will behave as if trying to persuade themselves when they mostly rely on *external* information about the transaction, whereas they will act in the opposite manner if the main source of information about the transaction is the incentive itself.

Finally, the paper has applications in fields like finance, political economics, and health economics, in which experts charged with acquiring and interpreting are influenced by incentives, as well as in information economics (informational moral hazard), and in marketing (bait and switch). These are detailed in Section 6.

The remainder of this paper proceeds as follows. Section 2 reviews some of the laws and policies that restrict incentives for transactions. Section 3 demonstrates empirically that incentives do not only change what people choose but also what they think their choices entail, in the context of an experiment that emphasizes ecological validity. Section 4 presents a theoretical analysis of the effects of incentives when information acquisition is costly, and shows that, in principle, the results of the experiment in the preceding section are consistent with Bayesian rationality. Section 5 conceptually replicates the first experiment and tests the extent to which behavior in an experiment in a stylized setting conforms to Bayesian rationality. Section 6 discusses policy implications of these findings and applications to other subfields of economics. Section 7 concludes.

2 Policies that restrict incentives

While this paper identifies a general mechanism, its focal application are the laws and guidelines that restrict incentive payments. Here, I review a selection of such laws. They have two things in common. First, protecting the person whom incentives would target is an important motivation (amongst others). Second, they are not intended to discourage these activities *per se*. On the contrary, altruistic participation is often applauded (e.g. Macklin, 1981).

⁵By studying the effect of increasing incentives on the quality of decision making, this paper also relates to Ariely et al. (2009). While the current paper focuses on the effects on the acquisition and interpretation of information, the latter work investigates choking under pressure, which is an altogether different mechanism.

- *Research participation.* The basis of laws on research ethics around the world is the [Belmont Report \(1978\)](#). It states that “undue influence ... occurs through an offer of an excessive ... reward or other overture in order to obtain compliance.” In similarly vague terms, the [National Bioethics Advisory Commission \(2001\)](#) fears that “offers of large sums of money ... could lead some prospective participants to enroll in a research study when it might be against their better judgment and when otherwise they would not do so.” The concern is that an offer may be “so excessively desirable that it compromises judgment” ([Emanuel, 2004](#)). These regulations apply to both medical and non-medical (e.g. psychological or experimental economic) research.
- *Reproduction.* Incentives in the context of reproductive technologies are frequently limited. A first example is human egg donation. The majority of countries surveyed by the Council of Europe in 1998 prevented human egg donation for commercial gain ([Council of Europe, 1998](#)).⁶ Protecting donors, who undergo substantial hormonal treatment, is a critical aim. The U.S. permits commercial human egg donation, but the [American Society for Reproductive Medicine \(2007\)](#) recommends that “payments to women providing oocytes should be fair and not so substantial that they ... lead donors to discount risks” concluding that “sums of \$5,000 or more require justification and sums above \$10,000 are not appropriate”.⁷ A second example is surrogate motherhood. Many U.S. states strictly limit material benefits for surrogate mothers, and partially out of concern for the surrogate mother. Nevada, New Hampshire and Washington, for instance, prohibit payments to surrogate mothers except for particular expenses that are explicitly listed in the states’ statutes.⁸
- *Tissue donation.* A prominent current policy debate concerns incentives for kidney donation ([Open Letter To President Obama, 2014](#); [Vatican Radio, 2014](#)). Paid living kidney donation is outlawed in every country of the world, except for the Islamic Republic of Iran ([Rosenberg, 2015b](#)). A frequent argument is that incentives would be coercive ([Choi, Gulati and Posner, 2014](#)) and that they would distort prospective participants’ assessment of the costs and benefits of the transaction, possibly to their detriment ([Satz, 2010](#); [Grant, 2011](#); [Kanbur, 2004](#)). A related example is bone marrow donation. The [World Marrow Donor Association \(2011\)](#) disapproves of incentives for bone marrow donation out of concern that “remuneration may ... cause the prospective donor to withhold personal health information for fear of being disqualified from donation, preventing an accurate risk assessment and disclosure of risks specific to that donor.”

⁶Ten of these had regulations allowing some reimbursement of expenses. Others outlawed human egg donations entirely.

⁷The class action lawsuit *Kamakahi v. ASRM* challenged these recommendation on antitrust grounds. A recent proposed settlement would require the ASRM to eliminate the explicit price recommendations from these guidelines ([Knaub, 2016](#)).

⁸Legislation varies widely within the U.S. On one extreme, California and Illinois support commercial surrogacy. On the other extreme, Michigan declares any participation in a surrogacy agreement a gross misdemeanor punishable with jail. This demonstrates how little consensus there is about the ethics of incentives, possibly because of a lack of information about how prospective participants would be affected.

- *Prostitution.* There are many views about the effects of selling sex on those who supply it. A prominent one maintains that overall, it harms prostitutes, even if they chose to engage in it voluntarily. This view, amongst others, has been instrumental in laws around the world that limit the incentives people can be paid for engaging in sexual activities (Farley, 2013; Danna, 2014).

Concerns about incentives are particularly prevalent for transactions involving bodily products and parts. But they are neither limited to that domain, nor do they fully encompass it. On the one hand, several laws limit incentives in domains that do not concern bodily products or parts. For instance, the U.S. outlaws selling oneself into voluntary slavery (42 U.S. Code §1994), and “excessive payments” are prohibited also for participants in non-medical experiments. Inducements are outlawed in student athlete recruiting on account that they would consist “undue influence” (National Collegiate Athletic Association, 2015). And the World Council of Churches compels its members not to use material incentives to induce individuals to change their confession, arguing that incentives would be coercive and impair religious freedom (Clark, 1996). On the other hand, there are transactions with bodily products for which no comparable laws apply. Many U.S. states, for instance, explicitly exclude the trade with human hair from statutory regulations on trade with bodily products and parts.⁹ Relatedly, no concerns about paying donors of human feces have surfaced, although donors can earn up to \$13,000 a year (Feltman, 2015).

3 Experiment: Do incentives cause self-persuasion?

Much of economics assumes that higher incentives increase participation in a transaction *only* because they exceed the reservation price of a larger number of people. However, incentives might also change how people acquire and interpret information about the transaction. This hypothesis lies behind the concerns that have led policy makers to restrict incentives for the transactions reviewed in Section 2.

In this section, I conduct a laboratory experiment to test this hypothesis. I say that incentives cause people to *self-persuade* if they change the distribution of reservation prices in a way that increases the likelihood of participation.¹⁰ The smoking gun is a comparative static on the supply curve. If people self-persuade, then the same increase in incentives will have a larger effect on the supply of people willing to participate in a transaction when they can access information about the transaction after learning the incentive than when they cannot.

This experiment captures two features of transactions for which incentives are restricted by law, such as living organ donation. First, they are unfamiliar to most prospective participants, and they

⁹For instance Connecticut, the District of Columbia, Illinois, Michigan, and Texas. Trade in human hair is a multi-million dollar industry (Khaleeli, 2012).

¹⁰The term *self-persuasion* is merely a label for a particular pattern of choices. It might arise both due to entirely rational reasons (as modeled in section 4, and similar to Bayesian Persuasion in Gentzkow and Kamenica (2011)), or due to irrational ones (as the term *self-deception*, that features in the psychology literature, suggests). The extent to which it reflects irrationality is a crucial determinant of welfare conclusions; see section 6.

are highly visceral. It is therefore difficult to assess the disutility with participation in monetary terms. Second, they are complex, and thus give prospective participants the opportunity to consult rich and multifaceted information, perhaps selectively. The transaction in this experiment must, in addition, be legal and reasonably feasible. Hence, in this experiment, subjects decide whether to ingest whole insects in exchange for cash payment. Like the motivating transactions, the experimental transaction is also unfamiliar and visceral, and thus hard to assess in monetary terms; and there is multifaceted, rich information about insects as food that laboratory subjects can be presented with. Crucially, insects are produced for human consumption in certified facilities, so that eating them is safe, and inducing people to eat them is legal.¹¹

Clearly, insect-eating is different from transactions such as organ donation on countless dimensions (examples include altruistic concerns and irreversibility). This experiment does *not* attempt to replicate these other dimensions, and does not need to. It demonstrates that the attributes it replicates are *sufficient* to cause self-persuasion. Consequently, self-persuasion should be expected whenever these attributes are present.

By examining a setting with genuine preference uncertainty about a real transaction, this experiment emphasizes ecological validity. The cost of this design choice is that distinguishing mechanisms becomes more challenging. This prevents a definitive assessment of the extent to which the observed behavior is consistent with rationality. Section 5 therefore presents a complementary experiment that examines behavior in a more stylized setting that allows for a precise delineation of mechanisms, and for a rigorous test of rationality.

3.1 Design

Structure. The experiment follows a 2×2 across-subjects design. The first dimension varies the incentive amount. A subject is offered either a \$3 or a \$30 incentive for eating an insect. The second dimension varies whether a subject can select and watch one of two videos about insects as food (the *video* and *no video* conditions, respectively). The *no video* condition serves a dual purpose. On the one hand, it allows me to study whether endogenous information acquisition changes supply curves. On the other hand, it serves as a control condition that allows me to difference out any effects that pertain to the payment amounts *per se*.¹²

¹¹Nonetheless, the transaction is rather intense to most subjects. Some reported that the experiment was “stressful” or that the “insects were scary”, and others refused to even just touch the containers they were packaged in. Even in countries such as China, Thailand, and Mexico, insect-eating is not practiced by a wide majority. Rather, it is associated with particular regions and / or communities, and often limited to a small number of insect species. In my data, Asians and Hispanics are neither more nor less willing to eat insects than Caucasians.

¹²Payment amounts *per se* might affect subjects’ decisions through at least four channels. First, because subjects are informed of both payment amounts, they might be unhappy to learn that they have been assigned to the \$3 condition. Second, eating insects for a lower amount of money may give the subject more bragging rights. Third, a high incentive amount might crowd out intrinsic motivation to eat insects (Frey and Jegen, 2001). Fourth, incentive amounts may serve as an anchor for subsequent decisions (Maniadi et al., 2014).

Briefly, the experiment follows four steps which are described in detail below. First, subjects learn the incentive amount they will be offered. Second, they select and watch one of the two videos (in the *video* treatment only). Third, they reveal their reservation price for eating insects. Fourth, they decide whether to eat the insect for the incentive amount offered in the beginning. Subjects then participate in ancillary treatments. The rules for payment and consumption of insects ensure that subjects find it in their interest to reveal their genuine preferences in each decision.

Main stages. In the first stage, subjects learn the incentive amount they are assigned to, and that they will decide, for each of five food items, whether to eat the item in exchange for that amount. They then learn that all of the food items are whole insects that are either baked, or cooked and dehydrated, and produced for human consumption. Subjects do not make these decisions yet, but their behavior in the intervening stages may be affected by the incentive. Subjects know of both incentive amounts and that they are randomly assigned to one of them, so that they cannot *rationally* draw inferences about the experience of eating insects (although effects such as anchoring may still apply).¹³

Only subjects in the *video* condition participate in the second stage; subjects in the *no video* condition proceed directly to the third stage.¹⁴ They decide between watching a 6-minute video entitled “Why you may want to eat insects” (the encouraging video) and one called “Why you may not want to eat insects” (the discouraging video). Directly after, they watch the selected video. The titles, and the approximate length of 6 minutes is all the information subjects have about the videos.¹⁵ Each subject in the video condition must view one of the videos, and nobody can watch both.¹⁶ Because the videos are relatively long and contain significant detail, incentives may not only change which video subjects choose, but also which parts of the video they pay attention to.¹⁷

In the third stage, subjects reveal reservation prices by filling in five multiple decision lists, one for each of the following five food items: 2 house crickets, 5 large mealworms, 3 silkworm pupae, 2 mole crickets, 2 field crickets. On each line in each list they decide between the options “Get \$ p . In exchange, eat the food item” and “Do not participate in this transaction,” for multiple values of p

¹³In contrast, subjects are only aware of the video condition they are in themselves.

¹⁴Hence, subjects in the *video* condition have more time to contemplate their choice, and thus additional opportunity to self-persuade. The objective is *not* to identify the effect of the specific menu of videos I offer, but rather to identify the effect of opportunities for self-persuasion in general. Hence, the additional contemplation time in the *video* treatment is desired.

¹⁵Each video lists various reasons for or against insect-eating. Transcriptions and links to the videos are in Appendix D.

¹⁶Because subjects can only watch either one or the other video, but not both, they can choose the kind, but not the amount of information to acquire. This design choice helps prevent diluting the treatment effect. The experiment in section 5 imposes no such restriction.

¹⁷Subjects in the *video* condition also select at least four out of a selection of 9 video clips, grouped in bins of three named “Reasons for eating insects”, “Reasons against eating insects”, “Other information about eating insects”. Subjects faced a 3% chance of watching the selected clips, and a 97% chance of watching the selected 6-minute video. See Appendix A.6 for details and analysis.

ranging from \$0 to \$60 in 21 increasingly larger steps.¹⁸ They select the least amount for which they are willing to eat the food item by clicking on the respective line; the remaining choices are filled in automatically.

In the fourth stage, subjects decide, for each of the insects, whether to eat it in exchange for the incentive they were assigned to in the first stage. Up to, and including this stage, subjects have no information about the food items, except for a verbal description and, potentially, the information contained in the video they had chosen. This motivates subjects in the video treatment to carefully decide which video to watch, and to pay attention.

		Implementation probability
<i>Main stages</i>		
1.	Learn incentive that will be offered in step 4: \$3 or \$30.	
2.	Select encouraging or discouraging video and watch it (<i>video</i> condition only).	
3.	Reveal reservation price for each species.	7%
4.	Make the decision announced in step 1 for each species.	80%
<i>Ancillary stages</i>		
5.	Insects handed out.	
6.	Reveal reservation price for each species.	7%
7.	Predict others' reservation prices.	6%

Table 1: Experiment timeline. Instructions for stages 1 through 6 are read out aloud in the beginning of the experiment. Stage 7 is a surprise, instructions are displayed on subjects' screens immediately before that stage.

Payment and execution of consumption decisions. Throughout the experiment, subjects make a large number of decisions. To incentivize subjects to reveal their genuine preferences in each of them, exactly one of all decision is randomly chosen for implementation at the end of the experiment. That decision entirely determines their payment and consumption of insects. The implementation probability is *not* uniform across decisions. The reason is that the decisions subjects make in the fourth stage (whether to eat insects for the promised cash amount) are not only an outcome measure, they are also intended to influence how they acquire and interpret information about insects as food. Therefore, with an 80% chance the decision selected for implementation is from stage 4. Table 1 details the implementation probability of the remaining stages.

All subjects know that the insects will be eaten in a visually secluded space in the presence of only the experimenter who ensures that the participant completely consumes the animals. This minimizes social motives, such as trying to impress others.

¹⁸ The amounts are 0, 1, 2, 3, 4, 6, 8, 10, 12.5, 15, 17.5, 20, 22.5, 25, 27.5, 30, 33, 36, 39, 44, 50, 60. Resolution is finer at lower levels since the distribution of reservation prices is positively skewed. The amount \$3 was not included in the decision lists for the first 79 Stanford subjects.

Subjects make many consumption decisions before having seen the actual insects, and thus may be unpleasantly surprised (see Appendix D for pictures). Participants cannot be forced to ingest insects. They nonetheless have an incentive to reveal their genuine preferences during the experiment, since reneging on a decision to eat an insect costs \$20. (Subjects who rejected the offer selected for implementation cannot renege.) Specifically, each subject receives a completion payment of \$35 which is reduced to \$15 if she reneges. In addition, the subject forfeits whatever she would have received for eating the insect. Subjects know all of this from the outset of the experiment.

Ancillary stages In the fifth stage, all subjects receive five containers. Each is filled with insects and a folded piece of paper with a code. They must enter all codes into the computer, which forces them to open each container, remove the label from within and thus view and (inadvertently) smell each of the insects.¹⁹ In the sixth stage, they again reveal their reservation prices.

In order to measure whether subjects are aware of the effects of incentives on others, the seventh stage asks them to predict the reservation prices of previous participants. Subjects make separate predictions for others in the \$3 and in the \$30 incentive conditions, but only for previous participants in the same video condition as themselves.²⁰ This stage comes as a surprise, so subjects' own decisions are not influenced by considerations of how others would decide. There is a 6% probability that a random such prediction will determine a subject's payment. If so, she will not consume any insects, and her completion payment is reduced by \$0.50 for each \$1 from which her prediction differs from the true mean.²¹

3.2 Implementation and preliminary analysis

Implementation A total of 671 participants participated in one of 39 computerized sessions in May, June, and July 2015 at the Ohio State University (499 subjects), Stanford University (110 subjects), and the University of Michigan (62 subjects). 271 subjects participated in the *no video* treatment (136 and 135 with \$3 and \$30 incentives, respectively), and 400 participated in the *video* treatment (197 and 203 with \$3 and \$30 incentives, respectively). A large number of subjects is required since individuals' willingness to eat insects is highly heterogeneous. Each session lasted about 2.5 hours, contained both payment conditions, and either all or none of the subjects in a session were in the *video* condition. At the beginning of the each session I read the instructions aloud.²²

¹⁹As a filler task during the handing out of the insects, subjects complete an extended version of the Cognitive Response Scale (Toplak, West and Stanovich, 2014), and sets D and E of Raven's (1960) standard progressive matrices.

²⁰Subjects first make a prediction for an average participant. They then separately predict the mean reservation price of those who were offered \$3 and \$30, respectively, in randomized order.

²¹A subject thus maximizes her expected payoff by stating the median of her belief distribution.

²²See Appendix D for the instructions. The experiment was coded in Qualtrics. I had expected a much larger number of subjects at the University of Michigan, but only 62 were available. I oversampled the video condition since only that condition reveals information choice. The 79 Stanford students who first participated in this experiment were not given any decisions to eat field crickets. The 48 first Stanford students also did not make any predictions about other participants. In addition, 68 Stanford students participated in an exploratory treatment. Data from the

I recruited subjects using the universities’ experimental economics participant databases. The invitation emails mentioned that the experiment would involve the consumption of food items on the spot, but did not mention insects.²³ It asked recipients not to participate if they have food allergies, are vegetarian or vegan, or eat kosher or halal.

Randomization check. Randomization into treatments was successful. Of 24 F -tests for differences in subjects’ predetermined characteristics across treatments, only one is significant at the 5% level, and four more are significant at the 10% level. This falls within the expected range. Details are in Appendix A.1.

Summary statistics. Eating insects is aversive to most participants. For each of the five species, Column 1 of Table 2 lists the fraction of subjects who have a positive reservation price in stage 3. For each item, at most 5% of subjects would eat it for free. The median reservation price is substantial, ranging from \$9 to \$18.75 (Column 2). There is also a substantial percentage of subjects who would not eat the insect even for the highest incentive amount offered in the multiple price lists (\$60), ranging from 18% to 30% (Column 3). A total of 9% of subjects are not willing to eat any insect for any price in any multiple price list.²⁴

Subjects could renege on the decision selected for implementation in exchange for \$20 if they had agreed to eat an insect. Five participants (0.8%) chose to do so. These participants would have been better off never having been offered the voluntary choice to eat an insect in exchange for money. All of them were in the \$30-condition.²⁵

	(1)	(2)	(3)
	Reservation price		
	Fraction \leq \$0	Median	Fraction \geq \$60
2 house crickets	0.96	9.00	0.18
5 large mealworms	0.96	18.75	0.30
3 silkworm pupae	0.95	13.75	0.23
2 mole crickets	0.96	13.75	0.24
2 field crickets	0.95	13.75	0.22

Table 2: Summary statistics of reservation prices elicited in stage 3. \$60 is the highest price offered in the multiple decision lists. 9% of subjects reveal a reservation price \geq \$60 for all five species. Data pooled over treatment conditions. Interval midpoints are used for analysis. $n = 671$.

exploratory treatment are not included in any analysis. In those treatments, the overwhelming majority of participants were presented with highly visceral images of insects, which muted self-persuasion.

²³An exception are the invitation emails in Michigan, and those for the last 31 Stanford subjects. Those mentioned that the experiment involves the voluntary consumption of food items, including edible insects. This information had no statistically measurable effect on the fraction of participants who refused to eat an insect for any price offered.

²⁴Each decision made in stage 4 of the experiment is also made as a part of a multiple price list in step 3. These decisions are sometimes inconsistent. See Appendix A.2 for details.

²⁵Four of them were in the video condition, and three had opted for the encouraging video.

Analysis. In all analyses, I control for subjects’ gender, age, age² and ethnicity, and include university and species fixed effects. All constants are estimates for the mean participant. Data from the multiple price lists are interval-coded. I use the interval-midpoints for analysis. The results are robust to alternative specifications (see Appendix A.7).

3.3 Main analysis

Result 1: Incentives raise the demand for encouraging information. With higher incentives subjects become more likely to demand information that encourages rather than discourages eating insects. The fraction of subjects choosing the discouraging video drops by over a third, from 18.3% to 11.3%, as incentives rise from \$3 to \$30. The effect size of 7.01% (s.e. 3.60) is significant at the 10% level.²⁶

Hence, subjects attempt to self-persuade by selecting the information to acquire. Do they succeed? If so, the video treatment should affect both, their willingness to engage in the transaction at the promised incentive as well as their reservation prices. The choice of a video is not the only mechanism that could cause such effects. Incentives may also change the extent to which subjects let themselves be persuaded by the chosen video. This is particularly plausible since the videos contain a variety of arguments, each of which one may deem more or less convincing, depending on incentives.

I intentionally do *not* condition the following analyses on the video a subject has watched. The reason is that the choice of information is one possible mechanism behind the self-persuasion effect this experiment is designed to document. Statistically muting this mechanism would defeat the experiment’s purpose.²⁷

Result 2: The supply of participants responds more strongly to incentives when endogenous information acquisition is possible. I now study supply curves, the potential smoking gun for self-persuasion. What fraction of subjects is willing to eat insects for the promised incentive? In the *no video* condition, 37.17% of subjects are willing to eat insects in exchange for \$3 (averaged across the five species), as Panel A of Table 3 shows. This number rises to a significantly higher 59.57% if the incentives are raised to \$30, an increase of 22.39 percentage points. This is no surprise. Incentives work, because a higher incentive exceeds the reservation price of a larger number of subjects.

In the *video* condition, in which selective information acquisition is possible, incentives can affect participation through a second mechanism. They may now also *change* reservation prices by affecting how subjects acquire and interpret information. If subjects successfully self-persuade, this second mechanism will cause participation rates to respond more strongly to incentives. Indeed, in the *video* condition, the rise in the incentive from \$3 to \$30 increases the participation rate from 37.69% to

²⁶Linear regression, $n = 400$.

²⁷Additionally, section 3.4 shows that the choice of the video alone neither suffices to explain the difference in the slope of the supply curves, nor the effect on reservation prices. Hence, incentives must also cause subjects to interpret the same video differently.

70.53%. This is an increase of 32.84 percentage points, 10.44 percentage points more than the same change in incentives had caused in the *no video* condition. The supply response is nearly 150%, as Panel A of Figure 1 illustrates. Subjects successfully self-persuade.

Incentive	A. Supply curves			B. Reservation prices		
	\$3	\$30	<i>Difference</i>	\$3	\$30	<i>Difference</i>
Information						
<i>no video</i>	37.17 (3.39)	59.57 (3.41)	22.39*** (4.78)	20.86 (1.76)	26.83 (1.59)	5.97*** (2.26)
<i>video</i>	37.69 (3.01)	70.53 (2.55)	32.84*** (3.93)	20.79 (1.60)	19.51 (1.45)	-1.28 (2.10)
Difference	0.52 (4.56)	10.96** (4.33)	10.44* (6.19)	-0.07 (2.32)	-7.32*** (1.99)	-7.25** (3.06)

Table 3: Panel A shows the percentage of participants who agree to eat the food item for the incentive amount they were promised, by treatment, and averaged over the five food items. Panel B shows estimates of mean reservation prices in dollars elicited in stage 3 by treatment, amongst those who are in principle willing to eat some insect in exchange for \$60 or less. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively. Asterisks are suppressed for levels.

Result 3: Incentives change reservation prices. I now analyze the effect of incentives on reservation prices directly, rather than through their implication on supply curves. The results show that incentives do not only change choices, but also subjects’ expectations of how unpleasant insect eating will be.

I face two challenges, one experimental, the other statistical. Experimentally, the fact that subjects are first presented with an incentive amount, and are then asked to reveal their reservation price in a multiple decision list potentially invites anchoring effects (Maniadis et al., 2014). To illustrate, the offer of \$15 that subjects encounter in the multiple decision lists may look quite attractive if one was initially promised \$3 in exchange for eating an insect, and quite puny if one was initially promised \$30. Through anchoring, incentives raise reservation prices.

Through self-persuasion, by contrast, incentives lower reservation prices. A subject may not even start thinking about eating insects when offered \$3, and \$15 might not make her change her mind. But if she is promised \$30 at the beginning of the experiment, she might persuade herself that eating insects is not all that bad. Once she faces the \$15 offer in the multiple price list, she might then also accept that amount. After all, she has already persuaded herself.

The anchoring mechanism relies merely on the presence of different promised incentives that precede the elicitation of reservation prices. It should thus be present to comparable extents in both the *video* and *no video* conditions. By contrast, persuading oneself is easier if one can selectively acquire and interpret information about the transaction. Hence, it should be more pronounced in the *video*

than in the *no video* condition. Therefore, I will measure the effect of incentives on expectations by the *difference* in their effect on reservation prices across the *video* and *no video* treatments.

The econometric challenge is twofold. First, while the vast majority of subjects are willing to eat insects once the incentive becomes sufficiently large, a minority are opposed to the idea of eating any insect for any price in the experiment, including the maximum of \$60. For those subjects, neither \$3 nor \$30 is a high incentive. One can therefore not reasonably expect these people to respond to the treatment. Second, even if a subject’s reservation price is less than \$60 for some species, it may exceed that threshold for others, and will thus be censored.

I estimate the effect of incentives on those subjects’ reservation prices who are, in principle, willing to eat insects in exchange for money. I use a version of Cragg’s (1971) double hurdle model which endogenously determines the incidence of the two types of subjects and simultaneously accounts for censoring. I account for the panel structure of the experimental data (Dong, Chung and Kaiser, 2004; Dong and Kaiser, 2008; Engel and Moffatt, 2014).²⁸

The results are in Panels B of Table 3 and Figure 1. In the *no video* treatment, there is a sizable anchoring effect. The increase in the incentive from \$3 to \$30 leads to an \$6.13 increase in reservation prices, from \$10.84 to \$26.97. If incentives do not lead subjects to form different expectations about the experience of bug-eating, then they should lead to a similar total change in the *video* treatment. The data, however, differ greatly.

For subjects in the *video* treatment, the increase in the incentive leads to a decrease in the average reservation price, from \$21.03 to \$19.63. The difference in the effects across the *video* and *no video* treatments is a significant \$7.53. Hence, incentives change subjects’ expectations by influencing how they acquire and interpret information.

These effects persist beyond the distribution of the insects. Receiving the insects lowers the mean reservation price in the (*video*, \$30)-treatment by \$2.52, raises it in the (*no video*, \$3)-treatment by \$1.35, and leaves it close to unchanged in the remaining conditions. The difference-in-difference across treatments remains virtually unchanged.²⁹

Result 4: Subjects are not fully aware that incentives cause self-persuasion As a proxy for rationality, I now study whether subjects are aware that incentives cause self-persuasion, both in others and in themselves.³⁰ I use the data from stage 7 of the experiment, in which subjects guess other subjects’ reservation prices and are incentivized for accuracy. They make separate guesses for those in the \$3 and \$30 conditions, respectively. Implicitly, subjects thus reveal their beliefs about how

²⁸See Appendix A.3 for details. The Cragg (1971) double hurdle model nests the Tobit model. Both models consist of two terms, one that determines the probability that an observation is censored, and one that determines the distribution of uncensored observations. The Tobit model imposes the restriction that any explanatory variable enters each of these terms with the same coefficient, the Craggit model does not. I test and reject this restriction in Appendix A.3.

²⁹See Appendix A.5 for details.

³⁰Lack of awareness does not disprove the rationality hypothesis, however, as Friedman and Savage’s (1948) famous example illustrates: An expert billiard player may be unaware of the laws of Newtonian mechanics, but he still strikes balls as if he were.

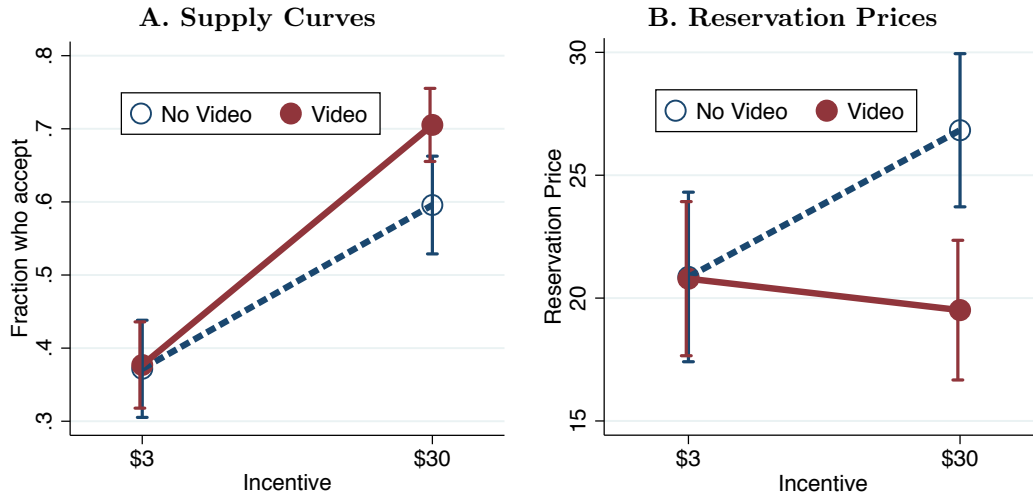


Figure 1: Panel A shows how the treatments affect the fraction of subjects willing to eat insects for the promised incentive. Panel B shows how the treatments affect reservation prices. Data in both panels taken from Table 3.

incentives affect others' reservation prices. I compare the actual effect of incentives to these beliefs (see Appendix A.4 for details). To distinguish between anchoring and self-persuasion, recall that subjects in the *no video* (*video*) condition guess the reservation prices of others who were also in the *no video* (*video*) condition. Data from the former subjects reveal whether they anticipate anchoring; data from the latter shows whether they correctly predict the combined effect of self-persuasion and anchoring.

Focusing on predictions about others in the *no video* condition, I find that subjects very accurately predict anchoring. As shown in Result 3, higher incentives raise reservation prices in the *no-video* condition. Subjects' do not only correctly predict the sign of this effect, they also accurately guess its magnitude ($p > 0.5$ for the test that predicted and actual effect sizes are equal).

In stark contrast, the predictions about others in the *video* condition are far off from the actual effect and directionally mistaken ($p < 0.01$). In fact, subjects predict the same effect of incentives regardless whether they concern others in the *video* condition or others in the *no video* condition. Hence, they correctly predict that others are subject to the anchoring effect, but they entirely fail to anticipate that incentives also cause others to self-persuade when they have the opportunity to selectively acquire information.

Subjects are not only unaware that incentives cause self-persuasion in others, they are similarly unaware that it affects themselves. To see this, recall that subjects in the *video* condition who face the \$30 incentive successfully self-persuade, and thus arrive at a lower reservation price than those who are given the \$3 incentive. Having just persuaded themselves that insect-eating is not that bad, they then predict that other subjects will also have low reservation prices, in spite of incentives for accuracy. Specifically, within the *video* condition, those offered the \$30 incentive predict reservation

prices that are, on average, \$2.62 lower than those offered the \$3 incentive ($p < 0.05$). (For those who could not see a video, this effect is just 40% as large, and not statistically significant.) Hence subjects are unable to net out the effect that incentives had on their own beliefs when making predictions about others.³¹ This shows that they lack awareness of the self-persuasion effect.³²

These findings call into question whether self-persuasion in this experiment is entirely consistent with rationality.

3.4 Alternative explanations and mechanisms

This experiment paints a coherent picture that subjects engage in self-persuasion: When incentives increase, people change how they acquire and interpret information in a way that is more favorable to participation. Hence, the supply of people willing to eat insects responds more strongly to increasing incentives in the *video* than in the *no video* condition, and expectations about the experience of insect-eating change accordingly. For each separate of these results, there are alternative explanations that I address here. None of them naturally explains the collection of findings. Nonetheless, because this experiment involves a physical transaction about which subjects have their own prior beliefs, the rigor with which I can separate mechanisms is limited. It motivates the experiment in section 5 that provides a precise delineation of mechanisms in a more stylized setting.

***Ex post* rationalization and cognitive dissonance.** It is conceivable that a subject decides to accept the transaction immediately after learning that his incentive is \$30 (or to refuse it immediately after learning that it is \$3), and merely chooses between videos as a means to *ex post* rationalize this choice, possibly as a way to resolve cognitive dissonance. Such *ex post* rationalization can explain both the results on video choice, and those on reservation prices. It is inconsistent, however, with the fact that supply is more responsive to incentives in the *video* treatment than in the *no video* treatment. According to *ex post* rationalization, the information in the video is not instrumental—it will not change the subjects’ decision whether to eat bugs for money—since that decision has already been made. The stronger response of supply to incentives in the *video* treatment, however, shows that at least some subjects’ participation decision is affected by the videos.³³

Confirmation bias and positive testing bias. A large literature in psychology on confirmation bias and positive testing bias shows that people often seek and interpret information in a way that

³¹Relatedly, [Chance et al. \(2011, 2015\)](#) find that people who are allowed to cheat on an IQ test convince themselves that their innate ability is higher, and are unable to correct these biased beliefs when incentivized to accurately predict their performance in a subsequent test in which opportunities to cheat are absent.

³²Self-persuasion is stronger, however, for subjects who expect this effect in others. This suggests that subjects tend to be aware of the direction in which incentives affect their beliefs, but are off about the magnitude.

³³The data are, however, consistent with a more involved account of *ex post* rationalization in which subjects anticipate how easily they will be able to *ex post* rationalize their choice, and are more likely to accept if it is easier. This mechanism is closely related to the idea that subjects might select particular sources of information not because they merely affect beliefs about the experience of eating insects, but because they directly affect that experience. The experiment in section 5 precludes this explanation.

tends to confirm the hypotheses they currently hold (Klayman, 1995; Rabin and Schrag, 1999). The results of this experiment can be interpreted as another manifestation of this tendency, but they go deeper by showing what the hypothesis is that people are testing. If incentives are high (low), individuals start from the hypothesis that participation (abstention) is the optimal choice. This is not trivial; one could imagine that high incentives trigger heuristics that make people more interested in exploring the potential downsides of the incentivized transactions, leading to the opposite effect.

Does information weaken anchoring? In this experiment, incentives change reservation prices differently when endogenous information acquisition is possible than when it is not. A potential explanation is that information eliminates anchoring. Several aspects of the data, however, do not naturally fit with this alternative.³⁴ First, it does not easily explain why the supply is more responsive to incentives in the *video* than in the *no video* condition.³⁵ Second, it does not account for the fact that subjects in the video treatment predict lower reservation prices for others if they receive a higher incentive themselves. Third, access to the videos changes reservation prices in the \$30 but not in the \$3 condition. It is not obvious why information would eliminate anchoring in only one of them.

Causal effect of the videos. In principle, the experimental results may obtain because the videos were unexpectedly (un)convincing, and hence might be an artifact of a design choice. If so, the behavioral effects should be explained entirely by subjects' video choice. Alternatively, incentives could lead to self-persuasion if they change how convincing a subject deems a given video. This second effect can less obviously arise as an artifact of the videos. It does play a substantial role, however, as two back-of-the-envelope calculations reveal. First, higher incentives lead to an additional increase in participation of 10.44 percentage points in the *video* treatment compared to the *no video* treatment. This exceeds the 7.01 percentage points increase in the number of subjects who choose the encouraging video, which therefore falls short of explaining the effect on participation. Second, suppose that the entire difference-in-difference estimate of the effect on reservation prices of \$7.25 is wholly due to subjects' choice of the video (rather than a differential interpretation of it). Since subjects in the \$30-condition chose the encouraging video only 7.01 percentage points more often, to explain the effect of \$7.25, the effect of watching the encouraging rather than the discouraging video on reservation prices would need to be on the order of $\$7.25 / 0.07 = \103.55 , much more than the maximal difference that could be measured by the price lists (\$60). In fact, the reservation prices of those who watched the discouraging video exceed those of other subjects in the *video* treatment by

³⁴In light of previous literature, this is not surprising. The participants in Ariely et al. (2003) were all given a sample of the aversive stimulus (obnoxious noise) before they were subjected to the anchor and revealed their reservation price to listen to more of the same noise. Hence, they had arguably complete information about that experience. Substantial anchoring effects were nonetheless present.

³⁵To explain the change in participation at \$30, the \$30 anchor would have to increase some subjects' reservation price from a value below \$30 to a value above \$30. The anchoring hypothesis, however, merely says that valuations will be drawn towards an anchor, not that they will overshoot.

only \$10.60 (s.e. 3.59).³⁶ Additionally, Appendix A.8 analyzes the data of only those subjects who opted for the encouraging video, and finds qualitatively unchanged results.³⁷

3.5 Discussion

This experiment shows that higher incentives do not merely exceed a larger fraction of reservation prices. They also cause subjects to acquire and interpret information in a way that is more favorable to participation. Hence, they change people’s expectations about what the transaction entails, in a way that amplifies the effect of incentives on participation. Subjects are not entirely aware of these effects, neither in themselves, nor in others.

These effects of incentives closely resemble the conjectures in the ethics literature that have given rise to policies that restrict incentives (see section 2). The extent to which these worries are justified depends on whether self-persuasion truly is an indication of deficient decision making, or whether it is merely an expression of Bayes-rational behavior. Policy makers and ethicists have typically assumed the former, and thus regarded self-persuasion as a cause for worry. The model in the next section, however, shows that self-persuasion can be fully consistent with Bayesian rationality. The experiment in section 5 then studies decision making in a stylized setting in which a rigorous test for rationality can be conducted.

4 Model: Can self-persuasion be rational?

The experiment in the previous section shows how an increase in incentives can lead to behavior that might look worrisome to ethicists and policy makers—it causes subjects to demand and interpret information in a way that is more favorable to participation, and to alter their expectations about what the transaction entails. But is this behavior *necessarily* a cause for concern? In this section, I show that with costly information acquisition, even a perfectly rational Bayesian will act in this way.³⁸ All proofs are in Appendix section B.4.

Intuition. To see the intuition, consider a prospective kidney donor. Kidney donation is a vastly complex transaction with potentially lifelong consequences. Most prospective donors, however, do not obtain an advanced degree in nephrology (kidney science) before their donation decision; this would be far too costly. Instead, they decide based upon incomplete information. This introduces the possibility that they make a different decision than they would have made under complete information, akin to

³⁶Clustered by subjects, estimated using university and species fixed effects. This statistic reflects a choice. If there is selection in the direction that more easily disgusted participants select the discouraging video more often (e.g. in order to justify their decision to not eat the insect for money, or because they anticipate that that video may contain off-putting imagery), the true effect of that video on reservation prices is even smaller.

³⁷That analysis ignores the endogeneity of the video choice.

³⁸There are other domains in which behavior at first sight appears non-rational, even though it is consistent with Bayesian updating. An example in which 60% of all drivers rationally consider themselves as above-average drivers is in Benoit and Dubra (2011).

type I and type II errors in classical statistics.³⁹ Hence, the question is: What *kind* of incomplete information should a rational prospective donor acquire, and how does this depend on the incentives offered?

For the sake of the (highly stylized) argument, suppose a prospective donor can talk to one of two groups of people.⁴⁰ One is composed of presumably happy previous donors, the other is a group of presumably unhappy previous donors. To capture the idea that information is costly, assume that she can consult only one of the two groups, but not both. Which of the two groups will she consult?

First, suppose she consults the group of presumably happy donors. They will likely encourage her to donate. Hence, *if* her decision whether to donate is mistaken, the mistake will likely be that she participates, even though the costs of participation exceed the benefits. This is a type I, or *false positive*, error. Second, suppose she consults the group of presumably unhappy donors. Those will likely discourage her from donating. Now, *if* her decision whether to donate is mistaken, the mistake will likely be the opposite kind—she abstains, even though the benefits of participation would have exceeded the costs. This is a type II, or *false negative*, error.

What kind of information the decision maker prefers to obtain (which group of previous donors she wants to consult) depends on whether false positive or false negative errors are more costly. Higher incentives for participation change these costs, and thus the preferred source of information. They make mistaken participation less costly, since even an undesirable outcome is less bad if one receives more compensation for it. At the same time, (mistakenly) abstaining is more costly, since one forgoes more money. Hence, with higher incentives and costly information acquisition, even a rational Bayesian will demand more encouraging information about the transaction, and therefore acquire different beliefs about what it entails.

4.1 Setting

I formally capture this intuition in the following setting. An agent decides whether or not to participate in a transaction in exchange for a material incentive m . The agent is uncertain about the (utility) consequences of participation, which depend on an unknown state of the world $s \in \{G, B\}$. The state is good ($s = G$) with prior probability μ . If so, and if the agent participates, he obtains utility π_G such that net utility is positive, $\pi_G + m > 0$. Otherwise, the state is bad ($s = B$). In that case, participation leads to utility π_B such that net utility is negative $\pi_B + m < 0$. If the agent does not participate, he receives utility 0.

Before the agent decides whether or not to participate, he can acquire information about the state, and decide how to act on that information. Implicitly, choosing what information to observe, and how to act on it, is a choice of a false negative and a false positive probability. Here, I let the agent choose

³⁹A decision that is mistaken in this sense is suboptimal only compared to the hypothetical benchmark of full information; given the costs of information it might be optimal.

⁴⁰Appendix B.1 contains a Balls-and-Urns example that parallels the argument here in a mathematically rigorous manner. It is a special case of the model in this section.

these probabilities directly (see Section 4.3 for how this choice can be made in practice). Specifically, the agent chooses a probability p_G of participating in the transaction if the state of the world is good, and a probability p_B of participating if it is bad. p_B and $(1 - p_G)$ are the false positive and false negative probabilities, respectively. An agent who is perfectly informed and participates if and only if the state is good, $(p_G, p_B) = (1, 0)$.

In this model, information acquisition is costly. Specifically, the cost of the information associated with a pair of state-contingent choice probabilities (p_G, p_B) is given by the increasing, convex, real valued, differentiable function $\lambda \cdot c(p_G, 1 - p_B)$, where $\lambda > 0$ is a constant that parametrizes the marginal cost of information. To capture the idea that the agent can implement any choice probability that does *not* depend on the state of the world without acquiring any information, I assume that $c(q, 1 - q) = 0$ for all $q \in [0, 1]$. Finally, I ensure an interior solution by assuming that $\lim_{p_G \rightarrow 1} c(p_G, 1 - p_B) = \infty$ for all $p_B < 1$ and $\lim_{p_B \rightarrow 0} c(p_G, 1 - p_B) = \infty$ for all $p_G > 0$. These conditions encompass Shannon mutual information costs (Sims, 2003, 2006; Caplin and Dean, 2013b; Matějka and McKay, 2015).⁴¹ Section 4.3 discusses the interpretation of this setting within the context of real world transactions, as well as within the experiment in section 3.

4.2 Analysis

I first derive the agent's objective function. If he selects state-dependent choice probabilities (p_G, p_B) , he obtains the upside payoff $\pi_G + m > 0$ with probability $\mu \cdot p_G$, and the downside payoff $\pi_B + m < 0$ with probability $(1 - \mu) \cdot p_B$. With the remaining probability he does not participate in the transaction and obtains 0. Hence, his *ex ante* expected utility, excluding costs of information, is $U(p_G, p_B; m) = \mu p_G(\pi_G + m) + (1 - \mu)p_B(\pi_B + m)$. The decision maker chooses the pair of probabilities (p_G, p_B) to solve the following problem.

$$\max_{p_G, p_B} U(p_G, p_B; m) - \lambda c(p_G, 1 - p_B) \quad (1)$$

How does the solution to this problem depend on the monetary incentive m ? The answer is most easily seen graphically. Figure 2 depicts (a part of) the agent's choice set.⁴² The vertical axis depicts p_G , the horizontal axis depicts $(1 - p_B)$. Both of these are goods, they are the probabilities of avoiding a mistaken decision, conditional on the state. Hence, the further up the bundle the agent chooses, the smaller is the probability of false negatives. The further right the bundle he chooses, the smaller is the probability of false positives. I separately plot the level curves of U and those of the cost of information function c on this space. The level curves of U are straight and parallel lines, since U is a

⁴¹With Shannon mutual information, the function c takes the following form. For state-dependent participation probabilities p_G, p_B let $p = \mu p_G + (1 - \mu)p_B$ denote the *ex ante* participation probabilities, and $\gamma_G = \frac{p_G \mu}{p}$ and $\gamma_B = \frac{p_B(1 - \mu)}{p}$ the agent's posterior belief about the event $\{s = G\}$ if he has observed a signal that makes him participate and abstain, respectively. Let h denote the binary entropy function, $h(x) = x \log(x) + (1 - x) \log(1 - x)$. Then, c is given by $c(p_G, p_B) = h(\mu) - E[h(\gamma_s)]$.

⁴²The agent's choice set is $\{(p_G, p_B) | 1 \geq p_G \geq p_B \geq 0\}$. For ease of exposition only a subset is depicted.

linear combination of p_G , $(1 - p_B)$ and a constant. U increases towards the upper right of the graph. The level curves of c are concave, since c is convex.

The agent chooses whether to acquire any information at all, and if so, what information, and hence which bundle (p_G, p_B) , to select. Throughout, I consider the case in which he does acquire information. Hence, for an initial level of material compensation m , the agent's optimal choice may be a vector such as point A in the figure.

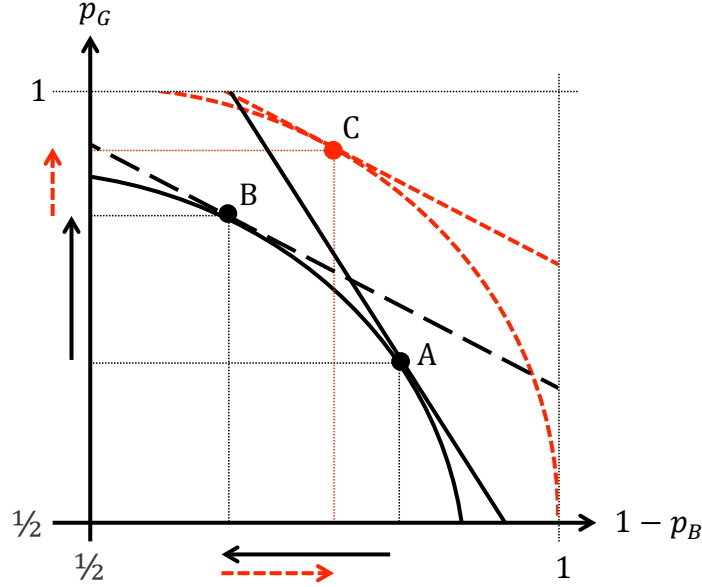


Figure 2: Effects of an increase in the incentive amount m . The horizontal axis plots $1 - p_B$, the probability that the agent rejects if the state is bad, the vertical axis plots p_G , the probability that the agent accepts if the state is good. The choice set is $\{(p_G, p_B) \in [0, 1]^2 : p_G \geq p_B\}$. For better visibility only a part is plotted here. Straight lines represent indifference curves of a Bayesian decision maker. Curved lines are iso-cost functions. The solid, black arrows indicate the substitution effect. The dashed, red arrows indicate the stakes effect.

The total effect of an increase in m derives from a substitution effect and a stakes effect. We obtain the former by temporarily interpreting problem (1) as the Lagrangian to the maximization of U subject to a constraint on the costs of information acquisition, $c(p_G, 1 - p_B) = \bar{c}$ for some fixed \bar{c} . An increase in m raises the weight of the good p_G in the utility function U and lowers that of $(1 - p_B)$. Intuitively, the increase in the weight on p_G reflects the increased opportunity cost of non-participation, whereas the decrease in the weight on $(1 - p_B)$ reflects the fact that higher incentives partially insure against adverse outcomes. Hence, the indifference curves tilt to the left, and the constrained optimum shifts to the northwest; for instance to a bundle such as point B . The agent now takes greater care avoid false negatives and is more tolerant of false positives. He acquires a different *kind* of information.

An increase in m not only changes the relative cost of false negatives and false positives, it also changes the total stakes of this decision. Hence, the agent may choose to spend a different *amount* of resources on information acquisition. If the agent chooses to acquire a larger *amount* of information, his optimal bundle will move to the northeast, for instance to a one such as point C .⁴³ This further decreases the incidence of false negatives, and counteracts the increase in false positives that arose through the substitution effect. The stakes effect will not outweigh the substitution effect as long as the cost of information function c has increasing differences (a positive cross-derivative). This is the case, for instance, with Shannon mutual information costs. The following proposition characterizes the total effect.

Proposition 1. *Consider an increase in the incentive from m to m' , with $m' > m$. Let (p_G, p_B) and (p'_G, p'_B) denote the associated optimal state-contingent participation probabilities. Then, if $p_G \neq p_B$ and $p'_G \neq p'_B$, the following hold.*

- (i) $p'_G > p_G$ and $p'_B > p_B$ if $\frac{\partial^2 c}{\partial p_G \partial p_B} > 0$ everywhere.
- (ii) A sufficient condition for part (i) is that the costs of information are proportional to Shannon mutual information.⁴⁴

The condition that the participation probabilities differ across the states means that the agent bases his choice upon a positive amount information, rather than on the prior alone. Part (i) shows that the increase in the incentive decreases the false negative rate and increases the false positive rate if a condition on the cost function is satisfied. Part (ii) shows that this condition is satisfied for the most popular cost-of-information function in the rational inattention literature.

This proposition explains why the same increase in incentives can lead to a larger increase in participation when endogenous information acquisition is possible than when it is not, as was the case in the insect experiment. Without endogenous information acquisition, subjects' reservation prices are given by their prior beliefs, which are independent of the incentive amount. Hence, incentives can increase participation only because they exceed the reservation prices of a larger number of subjects. With endogenous information acquisition, by contrast, there is an additional channel: Incentives may now directly affect the distribution of reservation prices.⁴⁵

The proposition also explains why the decision maker demands information that is more favorable to participation as incentives rise. To see this, think of the decision maker as following a signal that either tells him to participate or to abstain. The *ex ante* probability of participation (or, equivalently,

⁴³Whether this stakes effect is positive or negative depends on parameters.

⁴⁴More generally, it suffices that c is posterior-separable in the sense of [Caplin and Dean \(2013b\)](#), see Appendix Section B.4.

⁴⁵The model is consistent with the finding that the same increase in incentives leads to a larger increase in participation when endogenous information is possible than when it is not. It is, however, also consistent with the opposite. The experiment in section 5 tests a case in which the supply response must necessarily be stronger.

of receiving a “participate” signal) is given by

$$p = P(\text{participate}) = \mu p_G + (1 - \mu) p_B \quad (2)$$

If the condition in (i) holds, higher incentives cause the decision maker to opt for an information structure that is *ex ante* more likely to recommend participation both in the state in which this is the optimal action (higher p_G) and in the state in which it is not (higher p_B)—to demand information that is more encouraging. This also shows why perfectly Bayesian behavior might look worrisome to a policy maker or ethicist who mistakenly interprets the rational choice of more encouraging information as an irrational attempt at self-deception.

When to expect rational self-persuasion? The model also suggests we should expect self-persuasion whenever the stakes of the decision are comparable in magnitude to the costs of information. More specifically, self-persuasion will occur if full information is prohibitively costly, but small amounts of information are inexpensive, so that some information is demanded.

Proposition 2. *If the costs of information are proportional to Shannon mutual information, then there exists $\bar{\lambda} > 0$ such that the effect of an increase in incentives on participation due to different information acquisition, given by $\frac{\partial p}{\partial m}$, is increasing in the marginal cost of information λ for $\lambda \leq \bar{\lambda}$ and is equal to zero for $\lambda > \bar{\lambda}$.*

Hence, one cannot expect self-persuasion to disappear merely because the stakes in a given setting are large, as long as information acquisition is comparably costly. As an example, consider kidney donation. The fact that its consequences potentially extend over the remainder of a lifetime means both that the stakes are large, and that acquiring full information about its consequences on subjective well-being is extremely costly.⁴⁶ Hence, self-persuasion should be expected.

4.3 Interpretation

How to choose information in practice. The main assumption of this model is that the decision maker can separately choose false positive and false negative probabilities. In practice, this can be achieved in multiple ways.

Formally, consider a decision maker who gathers information sequentially, so his posterior beliefs whether participation or abstention is the optimal action evolve over time. Suppose he follows the decision rule illustrated in figure 3. He decides to participate as soon as his posterior that the state is good is sufficiently high, and he decides to abstain as soon as the posterior is sufficiently low. Otherwise, he continues searching for information. The choice of these thresholds corresponds to a

⁴⁶Moreover, this proposition helps assess the external validity of the experiments reported in this paper. It suggests that the empirical findings will generalize to settings with larger stakes, as long as the costs of information are also larger, and to a comparable extent.

choice of p_G and p_B . By choosing an upper threshold that is far from his prior, for instance, the decision maker implements a low false positive probability. If information about a transaction is rich and multifaceted, the agent may use such a rule to decide when to stop searching through the information and reach a decision.⁴⁷

There are other ways in which decision makers might select between different kinds of information in practice. An individual contemplating whether to donate a kidney, for instance, may deem certain kinds of information more credible than others, in a way that depends on the incentive. To illustrate, consider two recent studies on the consequences of a kidney graft on longevity. Mjøen et al. (2014) find that 20 years after donating a donor is 5 percentage points *more* likely dead than a comparable non-donor. Ibrahim et al. (2009), by contrast, find he is 2 percentage points *less* likely dead than a comparable non-donor. There are countless explanations for this difference. A decision maker may decide which of them to pay attention to, and thus which of the studies to believe in.

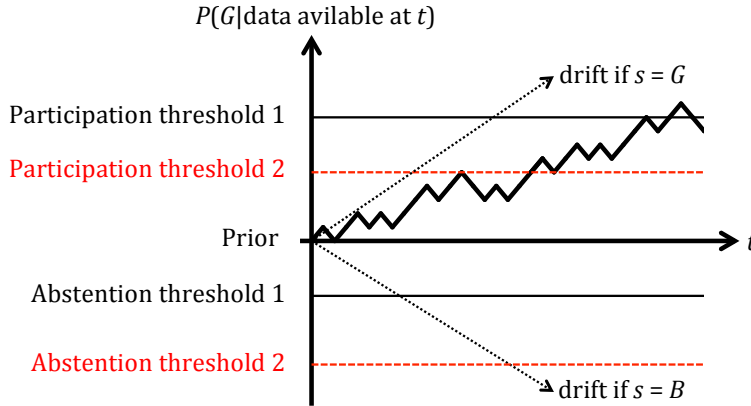


Figure 3: Selecting (p_G, p_B) with sequential information acquisition. The decision maker gathers information sequentially, until $P(G|\text{data})$ hits an upper threshold, in which case he participates, or until $P(G|\text{data})$ hits a lower threshold, in which case he abstains. If $s = G$, then $P(G|\text{data})$ drifts upwards, otherwise it drifts downwards. The further a threshold is from the prior, the less likely the subject makes the corresponding decision in error. In this example, both p_G and p_B are larger for the dashed, red thresholds than for the solid, black ones.

Learning about magnitudes. In this model, the state-dependent payoffs π_G and π_B are fixed and known; learning only concerns their probabilities. In any applied setting, by contrast, people are likely to lack and acquire information about the magnitudes of possible desirable and undesirable outcomes as well. Appendix B.3 extends this model to include uncertainty about magnitudes by allowing the unknown state s to take any value on the real line. The qualitative conclusions remain unchanged.

⁴⁷Hence, an agent may have significant leeway over choosing false positive and false negative rates, even if he is not explicitly presented with a menu of different sources of information. This idea is related to a Wald (1947) sequential probability ratio test and the drift-diffusion model (see Fehr and Rangel (2011) and Bogacz et al. (2006) for reviews). Che and Mierendorff (2016) more generally characterize how a Bayesian in a dynamic setting optimally selects informative signals to attend in order to arrive at an optimal distribution of posterior beliefs.

Interpreting the model in the context of real world transactions. The consequences of a choice like donating a kidney may take a lifetime to fully realize. How should we think about the state-contingent payoffs π_G and π_B in such a case? In this model, a decision maker commits a false positive or false negative error if he takes an action that differs from the one he would have taken under full information. Hence, we may interpret π_G (π_B) as the expected lifetime utility for a prospective participant for whom participation (abstention) is the optimal action, conditional on all information about its utility consequences for the particular decision maker that is potentially available at the time of the decision. By extension, medical information alone is not sufficient to make a well-informed choice. The agent is interested in encompassing information about how donation affects his life, for instance through restraining the professional and leisure activities he may want to pursue. It is also up to him to determine, for instance, how much utility he would lose from fatigue that may arise as a side effect of donation (Tellioglu et al., 2008; Beavers et al., 2001).

Interpreting the insect experiment. This model usefully aids interpreting the insect experiment. Due to that experiments' focus on ecological validity, however, relating it to this model requires additional assumptions. (For a more direct test of the model, see section 5.) The first concerns the interpretation of the videos. I assume the encouraging and discouraging videos correspond to pairs (p_G^+, p_B^+) and (p_G^-, p_B^-) of state-dependent participation probabilities, with $1 > p_G^+ > p_G^- > 0$ and $1 > p_B^+ > p_B^- > 0$. This requires that people cannot perfectly predict how convincing either video will be—if they could, the video would not be informative.⁴⁸ It also requires that people expect the encouraging video to cause a lower false negative rate, and a higher false positive rate than the discouraging video, as their titles suggest.⁴⁹ Under these assumptions, the model explains why subjects are more likely to choose the encouraging video when incentives are higher. Further, depending on the distribution of prior beliefs, it also explains the stronger supply response in the *video* than in the *no video* condition.^{50,51}

⁴⁸Based on priors alone, a subject will either participate regardless of the state, or abstain, regardless of the state, and thus have $(p_{G,B}) \in \{(1, 1), (0, 0)\}$. Intuitively, a subject could, for instance, decide based on a rule such as the following: "I have two worries about eating insects, and three potential hopes. If the video confirms at least two more worries than hopes, I will abstain; otherwise I will eat the insect in exchange for the incentive payment".

⁴⁹It is possible that someone who elects the encouraging video finds it so unconvincing that he arrives at more pessimistic beliefs about insect eating than if he had watched the discouraging video. This is precisely what one should expect from a Bayesian if the encouraging video decreases the false negative rate $1 - p_G$ sufficiently much relative to the increase in the false positive rate p_B , since the pessimistic realization of the Bayesian posterior, $P(s = G | \text{abstain}) = \frac{\mu(1-p_G)}{\mu(1-p_G) + (1-\mu)(1-p_B)}$ is decreasing in p_G (and increasing in p_B).

⁵⁰In this model, for a single subject who has to decide based upon prior information alone, a change in incentives that is effective increases the participation probability from 0 to 1. For a subject who can acquire information, by contrast, a change in incentives will increase the participation probability by an amount smaller than 1, but importantly, it may do so even if it were entirely ineffective for a subject deciding based on priors alone. The stronger supply response in the *video* condition of the insect experiment is consistent with this model if the latter effect outweighs the former.

⁵¹In the insect experiment, higher incentives suppress reservation prices. This is consistent with the modal outcome of the model. Specifically, the model predicts that with higher incentives, the agent chooses an information structure that more often increases his posterior (but more slightly), and more rarely decreases his posterior (but more heftily). Suppose the probability that the encouraging (discouraging) video increases posteriors is larger (smaller) than 0.5. Then the modal outcome over all possible signal realizations is that reservation prices decrease (increase) in the high (low) incentive treatment. If reservation prices are a linear function of beliefs, the model is not consistent, with any change in the expected reservation price. This is because by the law of iterated expectations, the expected posterior equals the

4.4 Discussion

The model in this section shows that when information about a transaction is costly, Bayesians will rationally engage in self-persuasion. Higher incentives make false positives cheaper and false negatives more expensive. Hence, the Bayesian will opt to commit more false positive errors and fewer false negative errors by demanding information that is more likely to convince him to participate, both when this is the right choice (fewer false negatives) and when it is not (more false negatives). He will rationally demand more encouraging information, and arrive at different posterior beliefs when the incentive is higher.

The model also highlights a crucial point regarding the external validity of the experiments in this paper. Whether self-persuasion will occur depends not simply on the magnitude of the stakes. Rather, it will occur whenever the costs of information and the stakes of the decision are comparable in magnitude.

The contribution of the model to the discussion of policies that restrict incentives is twofold. First, it produces seemingly worrisome behavior within a fully rational framework, and hence, from a standard welfare economics standpoint does not imply any worrisome behavior, as long as we are concerned about *ex ante* expected utility. (Section 6 discusses when policy makers and ethicists might nonetheless be concerned about the mechanisms identified here.) Second, it clarifies what kind of behavior might be a cause for concern. Directional changes in information acquisition or posterior beliefs are insufficient; one needs to document that incentives cause subjects to violate Bayesian rationality. The experiment in the next section performs this test in a stylized setting that maps directly to this model.

5 Experiment: How rational is self-persuasion?

The model in section 4 has shown that self-persuasion, the effect of incentives that may appear worrisome to ethicists and policy makers, can be entirely consistent with rational behavior. But is *empirical* behavior consistent with Bayesian rationality? Predictions of Bayesian self-persuasion and non-Bayesian self-*deception* are directionally similar, so that testing for rationality requires the extent of control an experiment in a stylized setting provides.

This experiment also permits a rigorous test of the model in section 4; alternative explanations that might apply to the insect experiment are excluded. In particular, the only good involved is money, so that the treatments can only affect beliefs about outcomes, but not the outcomes *per se*. In the insect experiment, by contrast, it is conceivable that watching an unpleasant video about insects makes the

prior, which, by linearity, implies that any agent's expected posterior reservation price must equal their prior reservation price. Because the insect experiment presents subjects with only two out of a very large space of conceivable videos about insect-eating, reservation prices are not averaged across all possible signal realizations, as the law of iterated expectations would require. Hence, the modal outcome of the model is arguably a more appropriate prediction about the insect experiment.

experience of ingesting them more revolting. In addition, I experimentally ensure that subjects have correct prior beliefs. In the insect experiment, it is possible that subjects began with implausible, or directionally distorted prior beliefs about bug-eating, in a way that might lead to behavior that would not be observed for a Bayesian with reasonable priors.⁵²

5.1 Design

The structure of this experiment parallels (a part of) the insect experiment in section 3. The difference is that instead of eating insects, subjects are incentivized to take a bad gamble. Someone who takes it and gets lucky loses nothing, someone who gets unlucky loses \$3.50. Each happens with known prior probability 0.5. Subjects can freely decide whether to take this gamble in exchange for a monetary incentive after they have acquired costly information about the likelihood of a loss, as detailed below.

The experiment follows a 2×2 design, and each subject participates in each condition exactly once, in random order.⁵³ The first dimension varies the incentive for participating in the gamble, which is either high (\$3) or low (\$0.50). Hence, a subject in the low incentive condition decides whether to take a win \$0.50 / lose \$3 gamble, whereas a subject in the high incentive condition decides about a win \$3 / lose \$0.50 gamble, and sees her options presented in this way.⁵⁴

Before subjects decide whether to take the gamble for the promised incentive, they view information about whether taking the gamble will lead to a loss. It is presented as a picture consisting of 450 randomly ordered letters, such as in Figure 4. Subjects know that if the lottery leads to a net gain, the picture contains 50 letters *G* and 40 letters *B* (for “good” and “bad”, respectively), and that these numbers are reversed if the lottery leads to a loss. Subjects can examine that picture as long as they like. Hence, somebody willing to put in the time and effort can know with certainty whether taking the gamble will lead to a gain or a loss.⁵⁵

The second treatment dimension varies the point in time at which a subject learns her incentive for taking the gamble. In the *incentive first* condition, she first learns the incentive, and then views the picture. In the *picture first* condition, this is reversed; she first examines the picture, and then learns her incentive. Before examining the picture, she only learns that the net upside of the bet will be \$3 or \$0.5 with equal chance, and that the net downside will be \$3 or \$0.5 with equal chance. Hence,

⁵²I do not address the issue of what it means for prior beliefs to be distorted or implausible. I merely contend that in the experiment in section 5 the experimentally induced prior beliefs represent the true data generating process, whereas no such assertion can be made for the experiment in section 3.

⁵³In addition to the four main treatments, each subject participates in two ancillary treatments that differ from the main treatments only regarding how much money can be lost or won. See Appendix C.2 for details.

⁵⁴I refrained from presenting the incentive amount separately from the win \$0 / lose \$3.50 gamble to minimize confusion.

⁵⁵To help subjects orient themselves about which round they are in, each picture was presented within a colored frame, and referred to by that color. For each subject these colors were selected randomly.

incentives may influence her information acquisition in the *incentive first* but not in the *picture first* condition.⁵⁶

To permit a test for Bayesian rationality, subjects reveal their subjective posterior belief that they have seen a good picture. They do so after deciding whether to take the gamble by selecting one of 12 bins, corresponding to 0, 5, 15, 25, ..., 85, 95, or 100% certainty that the state in that round was good. Payment according to the binarized scoring rule (Hossain and Okui, 2013) incentivizes truth-telling.⁵⁷ At any point during the belief elicitation stage, subjects can return to the previous stage of the experiment, in case the belief elicitation leads them to reassess whether to take the gamble. One may nonetheless be concerned that elicited beliefs are merely an *ex post* rationalization of the betting decision. I test and reject this hypothesis in section 5.2.

In the *incentives first* condition, incentives may affect information acquisition. To see how, consider the example of a subject with the following strategy. He scans through the picture, and keeps track of the number of *G*s and *B*s he has encountered. He does so until he has either seen two more *G*s than *B*s, in which case he takes the gamble; or until he has seen six more *B*s than *G*s, in which case he rejects. This subject's criterion for accepting the gamble is less stringent than that for rejecting it, and hence is more likely satisfied by mere chance. Hence, he will more likely commit a false positive than a false negative error. By choosing different criteria of when to stop searching and accept or reject the gamble, he can change the false positive and false negative probabilities (see Figure 3).⁵⁸

Implementation. I conducted this experiment on the Amazon Mechanical Turk online labor market with 450 subjects in March and October 2015 and an additional 503 subjects in April 2016. Data on subjective beliefs are available for the latter.⁵⁹ All instructions were presented on screen. A subject could proceed only if they correctly marked each of 11 statements about the instructions as true or false.⁶⁰

Each subject participated in each treatment once, in individually randomized order. For each subject and each decision, a state of the world was drawn according to the uniform prior probability, and a new picture with state-dependent, randomly ordered letters was generated. To prevent subjects

⁵⁶Caplin and Dean (2013a) and Caplin and Dean (2013b) use a similar way of presenting information. Babcock and Loewenstein (1997) and Gneezy et al. (2015) use a similar treatment variation to prevent the dependence of information acquisition and interpretation on situational factors.

⁵⁷This rule is incentive compatible regardless of the shape of a subject's utility for money, as long as preferences are linear in probabilities. Briefly, an agent is incentivized for his report r about his subjective probability of an event A as follows. The computer independently draws probability q from a uniform distribution. If $q > r$, the agent receives a prize with probability q . If $q \leq r$, the agent receives the prize if event A occurs. Since subjective probabilities were elicited in bins, I used the midpoint of each bin to determine payments, and informed subjects of this.

⁵⁸In principle, this interpretation has testable implications regarding response times. Measured response times in this experiment, however, are extremely variable.

⁵⁹The former sessions employed a confounded belief elicitation mechanism (a version of the quadratic scoring rule).

⁶⁰In case of a mistake, the subject only learns that at least one of the statements is marked incorrectly. Hence, it is extremely unlikely that participants completed this task by chance. The instructions and test questions are reproduced in Appendix D.

w q t f B G r p o w t o r q p r q d s s s d w y G p o G t d
f s B B f w B p q G o d t o r t q r t q t y B o t B o f o t
s s r r s s s d p s r B G B w f w t t r d q f y B w B f w s
d w f f p s q o y y w d w r B p G s G p s q o w r G o G p o
s y q B w d r w t G f r f p G p B w o r y q w r o G r p y y
q y r f w f q r G t w f y d f G y y p B f t r r f o G G t s
f y q o y B B r y r p d B o t f d B t y p p o r q y d G s p
p f y t B t y y y t d G o G B s p B p s o f s B w r y G r t
r G y G B r B r r w q G G B w t q d G f w w q y q t B t f G
y G f B f d G d w f r y t s p t f t r d w d d r t G G s d r
q w y w G G w t B G f s o B r r G y w p B w t f s p G p w q
G o B t G w B B G s r t f t s y f t y d s p q p t r f y p f
G r d p s o t o w o d G f G r w q y B G d q s q s B y B p y
s t q t d G o w G p G t d t o q G q t t f d t B q o y o d w
s G B f d B G w G q o r w o y f s p B d o w o s r p f f G q

Figure 4: Presentation of information about the state.

in the *picture first* condition from skewing their attention allocation according to the incentive, subjects could not return to the picture once they had decided to continue.⁶¹

Payment. Participants are paid for one randomly selected decision of one randomly selected round and thus have an incentive to reveal their genuine preferences in each decision. Subjects are aware that they will be paid according to a betting decision with 80% probability, and according to a belief elicitation decision with 20% probability. The larger weight on the betting decision serves to increase the effect of the incentive condition on information acquisition. Losses are discounted from a completion payment of \$6, gains are added. By comparison, laborers on Amazon Mechanical Turk typically earn an hourly wage around \$5 (Horton, Rand and Zeckhauser, 2011; Mason and Suri, 2012).

Preliminary analysis. On average, subjects take 33 minutes to complete the study. They pay attention to the pictures. Averaged over all treatments, they decide to bet 36.94 percent of the time if the state is bad, and a significantly higher 64.54 percent of the time if it is good.⁶² 5.47% of participants revise their decision about the bet during the belief elicitation stage in at least one round. They do so infrequently, only 1.01% of all decisions are changed.⁶³

⁶¹Subjects could not use a text editor to automatically count the letters because they were presented in a picture format (HTML5 Canvas). In principle, subjects might have taken screenshots to refer to the picture once incentives are revealed in the *picture first* condition. If they did, my results underestimate the effect of endogenous attention allocation.

⁶²The time spent examining a picture is right skewed, with a mean of 59 seconds per picture, and a median of 30 seconds. Response times are also highly dispersed, with a standard deviation of 87 seconds.

⁶³Due to a coding error, these decisions were not recorded in the October 2015 session; the reported numbers are based on the remaining 658 participants.

5.2 Analysis

I first test whether incentives cause self-persuasion also in this stylized setting. Do incentives change how subjects acquire and interpret information in a way that makes them more likely to participate? I document that they do, and then study how the beliefs elicited from subjects compare to the Bayesian posteriors. Do subjects choose what they think their choices entail, or is there a significant disconnect between elicited and Bayesian posteriors? In particular, do incentives cause subjects' behavior deviate from rationality in the way policy makers and ethicists are concerned about?

Result 1: The supply of participants willing to take the gamble responds more strongly to the incentive when the incentive can affect information acquisition. As in the insect experiment (section 3), I test for self-persuasion by documenting that the supply of people willing to take the gamble responds more strongly to incentives when information acquisition can depend on the incentive than when it cannot. Unlike in the insect experiment, the amount of potentially available information is held constant across conditions. Therefore, the model in section 4 *necessitates* this effect if information is instrumental in all four treatment conditions (see Appendix B.2 for a formal derivation).

Indeed, the increase in incentives from \$0.50 to \$3 increases participation by 39.61 percentage points (from 23.16% to 67.80%) in the *picture first* condition, as Panels A of Table 4 and Figure 5 show. In the *incentive first* condition, the effect of incentives is larger by a highly significant 9.44 percentage points. There, the same increase in incentives raises participation by 49.04 percentage points (from 21.11% to 71.15%). This confirms the theoretical prediction, and shows that incentives cause self-persuasion also in this stylized setting. Because the available information available was (stochastically) the same in all conditions, this result shows that incentives cause subjects to draw different conclusions from exactly the same information.

What leads to this result? Is the supply response in the *incentive first* condition stronger because subjects adjust the false positive rate or the false negative rate? By analyzing participation probabilities in the two states separately, we see that both play a significant role. When the incentive is low, the false positive rate is 4.36 percentage points lower in the *incentive first* than in the *picture first* condition, as Panel B of Table 4 shows. But it is a full 8.02 percentage points *higher* when the incentive is high. This combines to an increase in the false positive rate (relative to the *picture first* condition) of a highly significant 12.38 percentage points. A similar result obtains for correct positives (that is, avoiding false negatives). As Panel C shows, the correct positive rate is 5.58 percentage points lower in the *incentive first* than in the *picture first* condition when the incentive is low, but a (statistically insignificant) 2.47 percentage points higher when the incentive is high. Combined, the correct positive rate increases by a significant 8.05 percentage points (relative to the *picture first* condition) as incentives increase. All of these comparative statics are predicted by the model of section 4 if the cost of information has a globally positive cross-derivative (see Appendix B.2).

State	A. Supply curves			B. False positives			C. Correct positives		
	Both			Bad only			Good only		
Incentive	\$0.50	\$3	<i>Difference</i>	\$0.50	\$3	<i>Difference</i>	\$0.50	\$3	<i>Difference</i>
Information									
<i>picture first</i>	26.89 (1.59)	66.50 (1.68)	39.61*** (2.40)	12.29 (1.60)	44.74 (2.55)	32.44*** (3.00)	41.87 (2.55)	87.08 (1.64)	45.21*** (3.05)
<i>incentive first</i>	22.11 (1.35)	71.15 (1.50)	49.04*** (2.13)	7.93 (1.24)	52.75 (2.39)	44.81*** (2.68)	36.29 (2.21)	89.56 (1.35)	53.27*** (2.58)
Difference	-4.78** (1.86)	4.66** (2.14)	9.44*** (2.89)	-4.36** (1.90)	8.02** (3.21)	12.38*** (3.77)	-5.58* (3.01)	2.47 (2.00)	8.05** (3.66)

Table 4: Percentage of subjects willing to take the lottery. Panel A shows participation rates pooled over states. Exactly half the total weight is given to observations in which the state is good, and half to those in which the state is bad. It features 3502 observations from 953 subjects. Panels B and C separately show participation rates in the bad and good states, respectively. Panel B features 1923 observations from 893 subjects, and Panel C features 1889 observations from 892 subjects. Standard errors are clustered by subject. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively. Asterisks are suppressed for levels.

Result 2: Higher incentives make subjects more optimistic in a way that violates Bayesian rationality. I begin to test for rationality by studying whether elicited posterior beliefs satisfy the law of iterated expectations, one of the hallmarks of Bayesian reasoning. Formally, the law states that a Bayesian’s expected posterior must equal his prior. Intuitively, a Bayesian may not expect to become more optimistic, on average, after observing information—if he did, he should be more optimistic already. Because this law holds regardless of the information structure a Bayesian chooses to observe, it is well applicable in the current experiment.

Mean elicited posteriors in the *incentive first* condition are displayed in Panel B of Figure 5. They exceed the prior of 50% in the high incentive condition, and fall short of 50% in the low incentives condition. Hence, an increase in incentives makes subjects more optimistic in a way that violates the law of iterated expectations and thus Bayesian rationality. Moreover, this effect is independent of the information condition. In the *incentive first* condition, mean posteriors increase from 46.52% (s.e. 1.25) to 53.15% (s.e. 1.38) as incentives increase. The difference to the *picture first* condition is statistically indistinguishable; the respective magnitudes are 47.22% (s.e. 1.32) and 53.13% (s.e. 1.38).⁶⁴

While these deviations from rationality are statistically significant, their magnitude is rather small. Indeed, the following, more granular analysis of the elicited beliefs data shows that Bayesian rationality is a surprisingly good approximation to subjects’ behavior.

⁶⁴Standard errors clustered by subject.

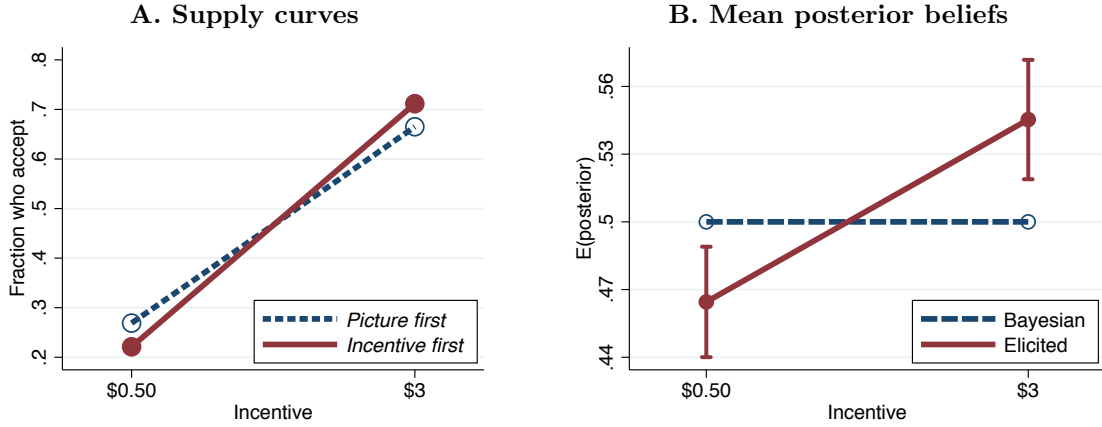


Figure 5: Choice probabilities and posterior beliefs that the state is good. Panel A displays the fraction of subjects accepting the gamble by treatment condition. Panel B shows mean elicited posterior beliefs in the *incentive first* condition and compares them to the Bayesian benchmark.

Result 3: Elicited and Bayesian posterior beliefs are close. In this experiment, Bayesian posteriors cannot be calculated algebraically. Nonetheless, the difference between elicited and Bayesian posteriors can be estimated. A subject's reported posterior p coincides with the Bayesian posterior if the objective probability that this subject has seen a good picture is p . Empirically, I fix a level of elicited posterior belief p and count the number of times the state has actually been good amongst all subjects who report that posterior.

It is far from obvious that elicited and objective posteriors will coincide. Consider how a subject might arrive at a subjective posterior belief. She examines the picture for a while, and at some point decides that she is sufficiently certain to make a decision about the bet. She then reports a subjective posterior that most likely does not result from any kind of calculation. How likely has she actually seen a good picture?

Elicited posteriors track objective posteriors surprisingly closely, as Figure 6 shows (data are averaged across all treatments). Amongst subjects who report subjective posteriors of 65%, for instance, 67.7% have actually seen a good picture. A similar result holds for many other elicited posteriors. Deviations rarely exceed a handful of percentage points and statistically significant differences are rare, as Column 1 of Table 5 shows. Consistent with result 2, when deviations do occur, objective posteriors tend to fall short of subjective ones, indicating slight overoptimism. This phenomenon is concentrated primarily at high elicited posteriors.

The fact that elicited and objective posteriors are close has the following implication. If a subject is more confident that the state is good, this is mainly because she possesses better information, not because she is more overoptimistic. To see this, notice that if higher subjective posteriors were predominantly a consequence of more pronounced overoptimism, increases in subjective confidence

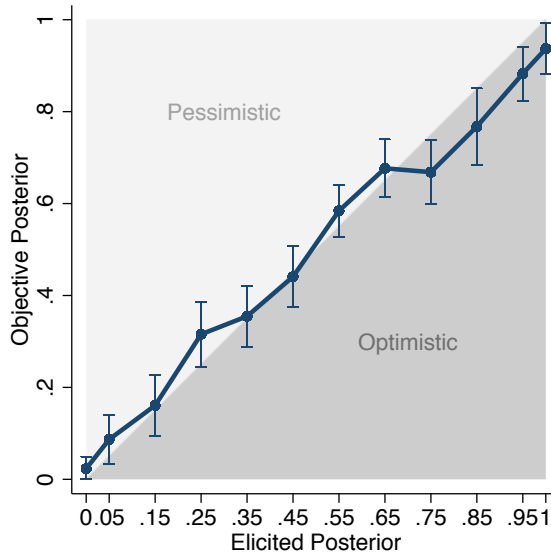


Figure 6: Elicited and objective posteriors. This figure plots estimates of the Bayesian posterior for each level of elicited posterior. Whiskers indicate 95% confidence intervals. Standard errors are clustered by subject.

would not be accompanied by corresponding increases in objective posteriors. Hence, the graph in Figure 6 would have a slope less than 1, whereas in fact it closely tracks the diagonal.⁶⁵

Result 4: Elicited posteriors are not an *ex post* rationalization of the decision to bet.

The results about subjective posterior beliefs would be difficult to interpret if elicited beliefs were simply an *ex post* rationalization of the decision to bet. To test whether this is the case, I again consider the difference between elicited and objective posteriors. If elicited beliefs were an *ex post* rationalization of the betting decision, then subjects who took the bet should appear more optimistic than those who refused it, *for any objective posterior*. That is, the graph in figure 6 should shift to the right if we consider only subjects who took the bet, and should shift to the left if we only consider those who refused it. By contrast, if beliefs inform the choice to bet, rather than *ex post* rationalize it, then objective and elicited beliefs should track each other similarly closely, regardless of whether subjects took the bet. Hence, I estimate Column 1 of Table 5 separately for the cases in which subjects decided to bet, and for the cases in which they abstained; Columns 2 and 3 of Table 5 display the results. In contrast to the *ex post* rationalization hypothesis, whenever the estimates differ at the 5%-level, it is the subjects who decided to reject the gamble who are more overly optimistic (Column 4). Hence, beliefs are *not* an *ex post* rationalization of the betting choice.

⁶⁵A parallel argument applies for subjects who are more confident that the state is bad.

Variable	(1)	(2)	(3)	(4)
	Objective posterior			<i>Difference</i>
Decision	Both	Accept	Reject	
Elicited posterior				
0	0.023*	0.000	0.024*	-0.024*
	(0.013)	(0.000)	(0.014)	(0.014)
0.05	0.087	0.473	0.080	0.393
	(0.027)	(0.355)	(0.027)	(0.356)
0.15	0.161	0.285	0.154	0.131
	(0.034)	(0.146)	(0.034)	(0.151)
0.25	0.316*	0.282	0.319*	-0.037
	(0.036)	(0.108)	(0.038)	(0.115)
0.35	0.355	0.393	0.349	0.043
	(0.034)	(0.100)	(0.036)	(0.107)
0.45	0.441	0.531	0.420	0.111
	(0.034)	(0.080)	(0.039)	(0.092)
0.55	0.584	0.576	0.596	-0.02
	(0.029)	(0.038)	(0.047)	(0.061)
0.65	0.677	0.669	0.708	-0.039
	(0.032)	(0.037)	(0.064)	(0.074)
0.75	0.668**	0.699	0.507**	0.192**
	(0.035)	(0.038)	(0.085)	(0.093)
0.85	0.768*	0.805	0.472**	0.332**
	(0.042)	(0.040)	(0.140)	(0.141)
0.95	0.883**	0.922	0.591**	0.331**
	(0.030)	(0.027)	(0.143)	(0.146)
1	0.937**	0.959**	0.564	0.395
	(0.028)	(0.018)	(0.277)	(0.278)
Observations	2,012	930	1,082	-
#Subj	503	470	484	-

Table 5: All regressions pool across both information conditions (*incentive first* and *picture first*). Standard errors clustered by subject. Asterisks indicate the level of statistical significance by which objective and elicited posteriors differ. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

5.3 Discussion

This experiment replicates the main empirical conclusion from the insect experiment. Incentives do not only change what people choose, they also lead people to self-persuade, and thus change what they believe their choices entail. The same increase in incentives has a significantly stronger effect on participation if the acquisition and interpretation of information can depend on the incentives than if it cannot.

This experiment allows for a rigorous test of rationality. On the one hand, higher incentives cause subjects to become more optimistic in a way that violates Bayesian rationality. This is the kind of violation that might appear concerning to ethicists and policy makers. On the other hand, these deviations are not large. In fact, elicited posteriors track objective posteriors surprisingly closely, regardless of the distribution of posteriors chosen by the subject, or the betting decision ultimately made.⁶⁶

The fact that elicited and objective posteriors are close even though different incentives substantially change their distribution suggests that subjects are well aware of the effect of incentives on information acquisition and beliefs. In the insect experiment, by contrast, subjects are unable to predict how incentives affect reservation prices when endogenous information acquisition is possible. While there are numerous differences between the two experiments, one salient reason for this divergence is the fact that the insect experiment involves a much more visceral transaction than the experiment in this section.⁶⁷

The next section discusses the welfare and policy implications of these findings, and suggests further applications.

6 Welfare, policy, and further applications

Overall, this paper shows that incentives may not only change what people choose, but also what they think their choices entail. When people can acquire costly information about the uncertain consequences of a transaction, then an increase in incentives will cause them to gather and interpret information in a way that is more favorable to participation, and hence to arrive at different beliefs about what the transaction entails. While this behavior has been hypothesized and taken as a cause for concern by ethicists and policy makers, I have shown that it should be expected from a Bayesian rational decision maker. Moreover, while higher incentives make people more optimistic in a way that violates the law of iterated expectations, empirical behavior adheres to the Bayesian benchmark rather closely. In this section, I first discuss the implications of these findings for the kind of transactions for which incentives are restricted by law (see section 2). I then suggest applications to various other subfields of economics.

⁶⁶Appendix C.1 shows that it neither depends on whether the incentive is high or low.

⁶⁷An alternative explanation is the fact that each subject participates in only one treatment in the insect experiment, but in all of them in the experiment in this section. Repeated participation in a similar setting potentially allows for learning, which might help subjects become aware of the effects of incentives. The data, however, reject this hypothesis, as evidenced by the fact that the graph in Figure 6 remains virtually unchanged if only the first of each subject's elicited posteriors is included.

6.1 Transactions with restricted incentives

When should we expect ethicists and policy makers to be concerned about self-persuasion? The phenomenon can arise from both rational and irrational causes, and these give rise to different sets of concerns.

Rational self-persuasion. When self-persuasion is rational, introducing or increasing incentives for participation in a transaction cannot make people *ex ante* worse off.⁶⁸ But as incentives rise, so does the false positive probability, and hence the number of people who *ex post* regret participation. Therefore, introducing or raising incentives will generally not be an *ex post* Pareto improvement. People may regard this as concerning for three reasons. First, a policy that reduces some individuals' *ex post* welfare potentially raises inequality, even if it increases *ex ante* expected utility—particularly so if it predominantly affects the poor. In this case, the question whether to implement the policy is no different from any other equity-efficiency tradeoff, which economics alone famously cannot advice. Second, both professional ethicists (Satz, 2010; Kanbur, 2004) and broader populations (Andreoni et al., 2015) sometimes consider *ex post* undesirable outcomes worrisome *per se*, even if the actions that led to them were entirely voluntary, and if they do not lead to an increase in *ex post* inequality.⁶⁹ These concerns arise especially with outcomes that are irreversible, or may lead to destitution. Policies as diverse as veteran service, personal bankruptcy laws, and emergency medical services mitigate adverse outcomes no matter whether they result from voluntary decisions. Third, purely selfish reasons (such as concerns about reelection) may render a politician wary of a policy that may cause significant *ex post* regret, even if it is *ex ante* beneficial.⁷⁰

Irrational self-persuasion (self-deception). If introducing or raising incentives leads to irrational reactions, then people who participate when a rational individual would have abstained will suffer from the policy even from an *ex ante* point of view. Obviously, prevention of incentives may have large costs that cannot be neglected—somebody who fails to find a donor kidney in time may die. Nonetheless, if increased incentives *ex ante* harm to those they target, these voluntary transactions are no longer Pareto-improvements, and standard welfare economic arguments no longer imply they should necessarily be supported.

Other reasons for concerns, or lack thereof. There are other reasons why ethicists disapprove of incentives for particular transactions; discussing those is outside of the scope of this paper.⁷¹

⁶⁸Unless people suffer from other biases that cause deficient decision making. See footnote 71.

⁶⁹Kanbur (2004) is explicit: “Extreme negative situations for individuals that leave them destitute attract our sympathy, no matter that the actions which led to them were freely undertaken.” Here, the concern is not about equity, but more narrowly about the welfare of the worst-off.

⁷⁰To illustrate, the documentary *Eggsploitation* (2011) criticizes incentives for egg donation based on unfavorable *ex post* outcomes to a handful of donors.

⁷¹See Kanbur (2004); Satz (2010); Grant (2011); Sandel (2012). Most closely related to the economics literature is the concern that incentives may hurt individuals with time-inconsistent preferences (Frederick, Loewenstein and

The motivation for restricting incentives, however, is not simply to prevent people from making potentially large mistakes; if that were the case, then also non-incentivized participation would need to be prevented. Rather, the concern is about *inducing* people to make decisions they might regret. This is particularly apparent when incentives are called *coercive* (Macklin, 1981; National Bioethics Advisory Commission, 2001; McGregor, 2005; Ripley, 2006).

Conversely, there are many transactions for which incentives do not raise concerns, although costly information acquisition, and thus the mechanisms identified in this paper, potentially play a role. One example is the decision to accept a new job in exchange for a higher wage. While, again, a comprehensive discussion is outside the scope of this paper, concerns appear less often for transactions whose consequences are easily reversible, or can be halted quickly and with only minor consequences if the decision to participate turns out to be mistaken. These properties limit the downside a person may suffer from the decision.

Policies. Two classes of policies can curtail self-persuasion without restricting incentives. They focus, respectively, on increasing the information people possess when they decide whether or not to enter the transaction, and on compensating undesirable outcomes *ex post*. Information, in turn, can be increased in two ways. On the one hand, policy can simply obligate people to be well informed before they participate, for instance through stringent informed consent requirements.⁷² In some cases, participation in a transaction can be limited to those who have undergone a similar experience before, and can thus better predict the consequences of participation. For example, commercial surrogate motherhood is legal in Russia, but only for women who have a child of their own (Svitnev, 2010). On the other hand, policy can decrease the costs of information acquisition, for instance by making it more easily accessible. The practicality of information-based policies is limited, however, by the fact that much pertinent information is lacking. For example, there is no consensus on whether or by how much living kidney donation affects life expectancy.⁷³ Similarly, human egg donors are not tracked, so that the only information relating multiple egg donation to reproductive cancers is a handful of case studies (Bodri, 2013).

The second approach focuses on compensating participants for *ex post* undesirable outcomes. In principle, this may render the difficult equity-efficiency tradeoff outlined above into a simple question of efficiency. Such a policy, however, has to overcome at least three obstacles. First, welfare losses from

o'Donoghue, 2002). Such concerns can be addressed by instituting cooling-off periods (Becker and Elias, 2007) or by making the disbursement of incentive payments coincide with the incidence of the utility costs of participating in the transaction over time.

⁷²According to Thiessen et al. (2013), for instance, current informed consent practices for living kidney donors leave much to desire. Informed consent documents often do not supply all required pieces of information, and are highly variable in items as crucial as complications related to surgery, health problems following surgery, and payment for routine follow-up care. Moreover, informed consent for living kidney donors is currently thin on psychological and emotional preparedness. Prospective living donors spend 45 minutes with a social worker or other mental health professional and equal time with a living donor advocate (livingdonor101.com, 2013).

⁷³Of two recent studies on the long term effects of kidney donation, one estimates that donors are 5 percentage points more likely to be dead 20 years after donation than comparable non-donors (Mj  en et al., 2014), whereas the other estimates they are 2 percentage points *less* likely to be dead after that time (Ibrahim et al., 2009).

undesirable *ex post* outcomes may not be perfectly observable, and incentivizing truthful revelation is hardly possible after the fact. Hence, determining adequate compensation and eligibility may be difficult.⁷⁴ Second, a significant part of the ethics literature maintains that it is fundamentally impossible to compensate people for the loss of certain goods such as bodily integrity (Anderson, 1995; Sandel, 2012; Sen, 1985, 1999; Nussbaum, 1995). Third, *ex post* undesirable outcomes are the reason why subjects acquire costly information about a transaction. A policy that insures against such outcomes therefore dilutes incentives for information acquisition, a form of moral hazard discussed below.

6.2 Further applications

Incentives and experts’ beliefs. Most fundamentally, this paper shows that if a person has incentives for entertaining certain beliefs, then they will adjust their acquisition and interpretation of information accordingly; a mechanism that is at least in part driven by rational motives. This finding has implications for many situations in which experts charged with acquiring and interpreting information are subject to incentives. In the domain of personal finance, for instance, Linnainmaa, Melzer and Previtero (2016) present evidence that financial advisors’ personal portfolios are often shaped by the same misguided investment strategies they recommend to their clients. For advisers who are incentivized to promote particular products, the present paper suggests that they will decide how to search for and interpret information in a way that supports giving this advice. Because this is driven by information acquisition, and potentially entirely rational, they will harbor the corresponding beliefs themselves, and follow the advice they give out. In a similar vein, this paper suggests that lobbying will not only change how politicians vote, but because incentives affect how people acquire and interpret information, it will also change their sincere beliefs about policies. The same mechanism provides further explanation as to why pharmaceutical gifts to doctors may be effective (Campbell et al., 2007; Morgan et al., 2006; Wazana, 2000).⁷⁵ The results also explain why even perfectly honest accountants may find it in their interest to allocate their scarce resources such as time and funds in a way that ultimately leads to skewed reporting (Bazerman et al., 2002).

Bait and switch marketing. This paper also helps better understand marketing techniques reminiscent of bait-and-switch, as have been employed, for instance, in US army recruiting. Briefly, prospective recruits learn at the beginning of the recruitment interview that they may be eligible for a sign-up bonus of up to several tens of thousands of dollars in value. They then proceed through the entire interview and take a battery of tests. Only just before signing the contract do they learn the actual bonus they are eligible for; for most it is far lower than the maximum that got their attention (Cave, 2005; McCormick, 2007). This technique may be effective because the prospect of the high

⁷⁴See Mitchell and Moro (2006) for a related issue regarding compensating those who lose from trade liberalization.

⁷⁵Malmendier and Schmidt (2012) identify a complementary reason in reciprocal gift-giving.

bonus causes candidates to favorably interpret the information given in the recruiting interview, making them more willing to enroll even at a lower incentive.⁷⁶ In the context of the model in section 4, this procedure works because the optimal posterior chosen by a rational Bayesian that makes him accept is more optimistic than the posterior for which he would just be indifferent between participation and abstention.⁷⁷

Informational moral hazard. Insuring participants against adverse outcomes is a frequently discussed policy in domains such as living kidney donation (Rosenberg, 2015a; Eyal et al., 2014). This paper suggests the potential side effect that insurance increases the fraction of participants who choose to participate but later regret this decision. This is because a Bayes-rational agent acquires costly information *only* in order to avoid false positives. Insurance against undesirable outcomes decreases the cost of false positives, and thus the incentive for information acquisition.⁷⁸ Unlike moral hazard that arises, for instance, when car insurance leads someone to drive more riskily, informational moral hazard can be mitigated by providing better information to the insured.

Bargaining impasse. Finally, this paper informs a literature finds that two parties who expect to take opposite stands in a bargaining situation interpret the same information differently (Babcock et al., 1995a,b; Loewenstein et al., 1993). They do so in an apparently self-serving way that may lead to costly bargaining impasse. This paper shows that even rational Bayesians may interpret the same information differently as incentives vary. To the extent that empirical behavior is driven by such considerations, merely making subjects aware of the bias will not eliminate it (as Babcock and Loewenstein (1997) have found empirically).

7 Conclusion

Much of economics assumes that higher incentives increase participation in a transaction *only* because they exceed the reservation price of a larger number of people. In this paper I have shown both experimentally and theoretically that when the consequences of participating in a transaction are uncertain and people can acquire costly information, incentives will increase participation also because they *change* reservation prices. People demand and interpret information in a way that is more

⁷⁶Generally, the bait lures a decision maker to make some investment. The switch is successful because when it occurs, the investment costs are sunk. Here, the investment is information acquisition. Existing work on bait-and-switch, by contrast, has focused on the case in which the investment consists in traveling to the point of sale (Lazear, 1995; Gerstner and Hess, 1990; Hess and Gerstner, 1998; Wilkie et al., 1998).

⁷⁷The reason is the following. A decision maker who rejects a transaction will not be better off. Hence, investing into costly information acquisition is justified only if conditional on having obtained the (optimally chosen) posterior that leads the decision maker to accept the transaction, the expected utility from doing so is strictly positive. This allows the counterparty to drop the incentive below the promised level after information acquisition has happened in a way that makes the decision maker still want to accept the transaction.

⁷⁸Relatedly, a literature starting with Doherty and Thistle (1996) consider how endogenous information acquisition interacts with adverse selection. That literature abstracts from the kind of moral hazard that arises due to endogenous information acquisition.

favorable to participation, attain different beliefs about the transaction, and become more likely to participate, both in states of the world in which participation is the optimal course of action and in those where it is not. Incentives do not only change what people choose, but also what they believe their choices entail.

This finding is particularly relevant for the many laws and regulations that restrict whether and how people can be incentivized to participate in transactions like organ donation, surrogate motherhood, medical trial participation, and others. Fears that incentives would lead to deficient decision making are one important motivation behind such laws (amongst others). I have shown that apparent indications for such deficiencies may be fully consistent with Bayesian rationality, and are therefore not a reason for concern *per se*. I have presented an experiment in which the effect of higher incentives on information acquisition and subsequent behavior is remarkably close to the predictions of Bayesian rationality (although higher incentives do make subjects slightly more optimistic in a way that is inconsistent with Bayesian rationality). Finally, I have discussed the kind of welfare objectives that may nonetheless justify concerns about incentives, and what kind of policy interventions may address them.

This paper is relevant for many applied fields in economics in which opportunities for costly information acquisition coexist with material incentives, such as health economics (pharmaceutical gifts to doctors, and informational moral hazard), marketing (bait-and-switch strategies), accounting, and political economy (lobbying). The experiments also provide an empirical validation of an important comparative static of the literature on rational inattention.

Most broadly, this paper bridges a gap between economics on the one hand, and the applied ethics and policy literatures on the other. It uses standard economic methodology to inform a prominent, but vaguely formulated concern in the latter literature both empirically and conceptually. Generally, using the powerful toolbox of economics to study moral intuitions and the behavioral assumptions on which they rely is an important direction for future research. For otherwise our laws will be based on intuitions alone.

References

- Ambuehl, Sandro, Muriel Niederle, and Alvin E. Roth**, “More Money, More Problems? Can High Pay be Coercive and Repugnant?,” *American Economic Review, Papers & Proceedings*, 2015, 105 (5), 357–60.
- American Society for Reproductive Medicine**, “Financial compensation of oocyte donors,” *Fertility and Sterility*, 2007, 88 (2), 305–309.
- Anderson, Elizabeth**, *Value in ethics and economics*, Harvard University Press, 1995.
- Andreoni, James, Deniz Aydin, Blake Barton, B. Douglas Bernheim, and Jeffrey Naecker**, “When Fair Isn’t Fair: Sophisticated Time Inconsistency in Social Preferences,” *unpublished manuscript*, 2015.
- Ariely, Dan, George Loewenstein, and Drazen Prelec**, “‘Coherent Arbitrariness’: Stable Demand Curves Without Stable Preferences,” *The Quarterly Journal of Economics*, 2003, 118 (1), 73–105.
- , **Uri Gneezy, George Loewenstein, and Nina Mazar**, “Large stakes and big mistakes,” *The Review of Economic Studies*, 2009, 76 (2), 451–469.
- Babcock, Linda and George Loewenstein**, “Explaining bargaining impasse: The role of self-serving biases,” *The Journal of Economic Perspectives*, 1997, 11 (1), 109–126.
- , – , and **Xianghong Wang**, “The relationship between uncertainty, the contract zone, and efficiency in a bargaining experiment,” *Journal of Economic Behavior & Organization*, 1995, 27 (3), 475–485.
- , – , **Samuel Issacharoff, and Colin Camerer**, “Biased judgments of fairness in bargaining,” *The American Economic Review*, 1995, pp. 1337–1343.
- Basu, Kaushik**, “The economics and law of sexual harassment in the workplace,” *The Journal of Economic Perspectives*, 2003, 17 (3), 141–157.
- , “Coercion, Contract and the Limits of the Market,” *Social Choice and Welfare*, 2007, 29 (4), 559–579.
- Bazerman, Max H, George Loewenstein, and Don A Moore**, “Why good accountants do bad audits,” *Harvard business review*, 2002, 80 (11), 96–103.
- Beavers, Kimberly L, Robert S Sandler, Jeffrey H Fair, Mark W Johnson, and Roshan Shrestha**, “The living donor experience: donor health assessment and outcomes after living donor liver transplantation,” *Liver transplantation*, 2001, 7 (11), 943–947.

- Becker, Gary S and Julio J Elias**, “Introducing incentives in the market for live and cadaveric organ donations,” *The Journal of Economic Perspectives*, 2007, *21* (3), 3–24.
- Bénabou, Roland**, “The Economics of Motivated Beliefs,” *Revue d’économie politique*, 2015, *125* (5), 665–685.
- **and Jean Tirole**, “Self-confidence and personal motivation,” *Quarterly Journal of Economics*, 2002, *117* (3), 871–915.
- **and —**, “Incentives and prosocial behavior,” *The American Economic Review*, 2006, *96* (5), 1652–1678.
- **and —**, “Identity, morals, and taboos: Beliefs as assets,” *The Quarterly Journal of Economics*, 2011, *126* (2), 805–855.
- **and —**, “Mindful Economics: The Production, Consumption, and Value of Beliefs,” *The Journal of Economic Perspectives*, 2016, *30* (3), 141–164.
- Benoît, Jean-Pierre and Juan Dubra**, “Apparent overconfidence,” *Econometrica*, 2011, *79* (5), 1591–1625.
- Bodri, Daniel**, “Risk and complications associated with egg donation,” in “Principles of Oocyte and Embryo Donation,” Springer, 2013, pp. 205–219.
- Bogacz, Rafal, Eric Brown, Jeff Moehlis, Philip Holmes, and Jonathan D Cohen**, “The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks,” *Psychological Review*, 2006, *113* (4), 700–765.
- Boo, Michael, Suzanna M van Walraven, Jeremy Chapman, Brian Lindberg, Alexander H Schmidt, Bronwen E Shaw, Galen E Switzer, Edward Yang, Torstein Egeland, and World Marrow Donor Association**, “Remuneration of hematopoietic stem cell donors: principles and perspective of the World Marrow Donor Association,” *Blood*, 2011, *117* (1), 21–25.
- Campbell, Eric G, Russell L Gruen, James Mountford, Lawrence G Miller, Paul D Cleary, and David Blumenthal**, “A national survey of physician–industry relationships,” *New England Journal of Medicine*, 2007, *356* (17), 1742–1750.
- Caplin, Andrew**, “Measuring and Modeling Attention,” *Annual Reviews of Economics*, forthcoming.
- **and Daniel Martin**, “A testable theory of imperfect perception,” *The Economic Journal*, 2014, *125* (582), 184–202.
- **and John Leahy**, “Psychological expected utility theory and anticipatory feelings,” *Quarterly Journal of Economics*, 2001, *116* (1), 55–79.

- **and Mark Dean**, “Rational inattention and state dependent stochastic choice,” *unpublished manuscript*, New York University, 2013.
 - **and –**, “Rational Inattention, Entropy, and Choice: The Posterior-Based Approach,” *unpublished manuscript*, New York University, 2013.
 - **and –**, “Revealed Preference, Rational Inattention, and Costly Information Acquisition,” *American Economic Review*, 2015, *105* (7), 2183–2203.
- Cave, Damien**, “Critics Say It’s Time to Overhaul The Army’s Bonus System,” *New York Times*, August 15 2005.
- Center for Bioethics and Culture**, “Eggsploitation,” www.eggsploitation.com 2011.
- Chance, Zoe, Francesca Gino, Michael I Norton, and Dan Ariely**, “The slow decay and quick revival of self-deception,” *Frontiers in psychology*, 2015, *6* (1075).
- Chance, Zoë, Michael I Norton, Francesca Gino, and Dan Ariely**, “Temporal view of the costs and benefits of self-deception,” *Proceedings of the National Academy of Sciences*, 2011, *108*, 15655–15659.
- Che, Yeon-Koo and Konrad Mierendorff**, “Optimal Sequential Decision with Limited Attention,” *unpublished manuscript*, 2016.
- Choi, Stephen J, Mitu Gulati, and Eric A Posner**, “Altruism Exchanges and the Kidney Shortage,” *Law & Contemporary Problems*, 2014, *77*, 289–322.
- Clark, Alan C.**, “The Challenge of Proselytism and the Calling to Common Witness,” *Ecumenical Review*, 1996, *48* (2), 212–221.
- Council of Europe**, “Medically assisted procreation and the protection of the human embryo. Comparative study on the situation in 39 states,” Strasbourg June 1998.
- Cragg, John G**, “Some statistical models for limited dependent variables with application to the demand for durable goods,” *Econometrica*, 1971, *39* (5), 829–844.
- Cryder, Cynthia E., Alex John London, Kevin G. Volpp, and George Loewenstein**, “Informative inducement: Study payment as a signal of risk,” *Social Science & Medicine*, 2010, *70*, 455–464.
- Danna, Daniela**, “Report on prostitution laws in the European Union,” *unpublished manuscript*, 2014.
- Doherty, Neil A and Paul D Thistle**, “Adverse selection with endogenous information in insurance markets,” *Journal of Public Economics*, 1996, *63* (1), 83–102.

- Dong, Diansheng and Harry M Kaiser**, “Studying household purchasing and nonpurchasing behaviour for a frequently consumed commodity: two models,” *Applied Economics*, 2008, *40* (15), 1941–1951.
- , **Chanjin Chung, and Harry M Kaiser**, “Modelling milk purchasing behaviour with a panel data double-hurdle model,” *Applied Economics*, 2004, *36* (8), 769–779.
- Elias, Julio J, Nicola Lacetera, and Mario Macis**, “Markets and Morals: An Experimental Survey Study,” *PloS one*, 2015, *10* (6).
- Elias, Julio, Nicola Lacetera, and Mario Macis**, “Sacred Values? The Effect of Information on Attitudes Toward Payments for Human Organs,” *American Economic Review, Papers & Proceedings*, 2015, *105* (5), 361–365.
- Eliaz, Kfir and Ran Spiegler**, “Can anticipatory feelings explain anomalous choices of information sources?,” *Games and Economic Behavior*, 2006, *56*, 87–104.
- Emanuel, Ezekiel J**, “Ending concerns about undue inducement,” *The Journal of Law, Medicine & Ethics*, 2004, *32* (1), 100–105.
- Engel, Christoph and Peter G Moffatt**, “dhref, xtdhref, and bootdhref: Commands to implement double-hurdle regression,” *Stata Journal*, 2014, *14* (4), 778–797.
- Epley, Nicholas and Thomas Gilovich**, “The Mechanics of Motivated Reasoning,” *Journal of Economic Perspectives*, 2016, *30* (3), 133–40.
- Eyal, Nir, Julio Frenk, Michele B. Goodwin, Lori Gruen, Gary R. Hall, Douglas W. Hanto, Frances Kissling, Ruth Macklin, Steven Pinker, Lloyd E. Ratner, Harold T. Shapiro, Peter Singer, Andrew W. Torrance, Robert D. Truog, and Robert M. Veatch**, “An Open Letter to President Barack Obama, Secretary of Health and Human Services Sylvia Mathews Burwell, Attorney General Eric Holder and Leaders of Congress,” 2014.
- Farley, Melissa**, “Prostitution, Liberalism, and Slavery,” *Logos: A Journal of Modern Society & Culture*, 2013, *46* (3).
- Fehr, Ernst and Antonio Rangel**, “Neuroeconomic foundations of economic choice—recent advances,” *The Journal of Economic Perspectives*, 2011, *25* (4), 3–30.
- Feltman, Rachel**, “You can earn \$13,000 a year selling your poop,” *Washington Post*, January 2015.
- Frederick, Shane, George Loewenstein, and Ted o’Donoghue**, “Time discounting and time preference: A critical review,” *Journal of Economic Literature*, 2002, *40* (2), 351–401.
- Frey, Bruno S and Reto Jegen**, “Motivation Crowding Theory,” *Journal of Economic Surveys*, 2001, *15* (5), 589–611.

- Friedman, Milton and Leonard J Savage**, “The utility analysis of choices involving risk,” *Journal of Political Economy*, 1948, 56 (4), 279–304.
- Gentzkow, Matthew and Emir Kamenica**, “Bayesian persuasion,” *American Economic Review*, 2011, 101 (6), 2590–2615.
- Gerstner, Eitan and James D Hess**, “Can bait and switch benefit consumers?,” *Marketing Science*, 1990, 9 (2), 114–124.
- Gino, Francesca, Michael I Norton, and Roberto A Weber**, “Motivated Bayesians: Feeling Moral While Acting Egoistically,” *Journal of Economic Perspectives*, 2016, 30 (3), 189–212.
- Gneezy, Uri, Silvia Saccardo, Marta Serra-Garcia, and Roel van Veldhuizen**, “Motivated self-deception, identity, and unethical behavior,” *unpublished manuscript*, 2015.
- Grant, Ruth W**, *Strings attached: Untangling the ethics of incentives*, Princeton University Press, 2011.
- Greene, William**, “The behaviour of the maximum likelihood estimator of limited dependent variable models in the presence of fixed effects,” *The Econometrics Journal*, 2004, 7 (1), 98–119.
- Held, P.J., F. McCormick, A. Ojo, and J.P. Roberts**, “A Cost-Benefit Analysis of Government Compensation of Kidney Donors,” *American Journal of Transplantation*, 2016, 16 (3), 877–885.
- Hess, James D and Eitan Gerstner**, “Yes, ‘bait and switch’ really benefits consumers,” *Marketing Science*, 1998, 17 (3), 283–289.
- Holt, Charles A. and Susan K. Laury**, “Risk Aversion and Incentive Effects,” *American Economic Review*, 2002, 92(5), 1644 – 1655.
- Horton, John J., David G. Rand, and Richard J. Zeckhauser**, “The online laboratory: Conducting experiments in a real labor market,” *Experimental Economics*, 2011, 14, 399–425.
- Hossain, Tanjim and Ryo Okui**, “The binarized scoring rule,” *The Review of Economic Studies*, 2013, 80 (3), 984–1001.
- Ibrahim, Hassan N, Robert Foley, LiPing Tan, Tyson Rogers, Robert F Bailey, Hongfei Guo, Cynthia R Gross, and Arthur J Matas**, “Long-term consequences of kidney donation,” *New England Journal of Medicine*, 2009, 360 (5), 459–469.
- Kahneman, Daniel, Jack L Knetsch, and Richard Thaler**, “Fairness as a constraint on profit seeking: Entitlements in the market,” *The American Economic Review*, 1986, 76 (4), 728–741.
- Kamenica, Emir**, “Contextual inference in markets: On the informational content of product lines,” *The American Economic Review*, 2008, 98 (5), 2127–2149.

- Kanbur, Ravi**, “On obnoxious markets,” in Stephen Cullenberg and Prasanta Pattanaik, eds., *Globalization, Culture and the Limits of the Market: Essays in Economics and Philosophy*, Oxford University Press, 2004.
- Khaleeli, Homa**, “The hair trade’s dirty secret,” *The Guardian*, 28 October 2012.
- Klayman, Joshua**, “Varieties of confirmation bias,” *Psychology of learning and motivation*, 1995, 32, 385–418.
- Knaub, Kelly**, “Egg Donors Get Pay Limits Axed With Antitrust Settlement,” *Law360*, February 2016.
- Koszegi, Botond**, “Ego Utility, Overconfidence, and Task Choice,” *Journal of the European Economic Association*, 2006, 4 (4), 673–707.
- Kunda, Ziva**, “The case for motivated reasoning,” *Psychological bulletin*, 1990, 108 (3), 480.
- Lazear, Edward P**, “Bait and switch,” *Journal of Political Economy*, 1995, 103 (4), 813–830.
- Leider, Stephen and Alvin E Roth**, “Kidneys for sale: Who disapproves, and why?,” *American Journal of Transplantation*, 2010, 10 (5), 1221–1227.
- Linnainmaa, Juhani T, Brian T Melzer, and Alessandro Previtero**, “The Misguided Beliefs of Financial Advisors,” *Unpublished manuscript*, 2016.
- livingdonor101.com**, “Living Donor Informed Consent,” <http://www.livingdonor101.com/informedconsent.shtml> November 30 2013.
- Loewenstein, George, Samuel Issacharoff, Colin Camerer, and Linda Babcock**, “Self-serving assessments of fairness and pretrial bargaining,” *The Journal of Legal Studies*, 1993, pp. 135–159.
- Macklin, Ruth**, “‘Due’ and ‘Undue’ Inducements: On Paying Money to Research Subjects,” *IRB*, 1981, 3 (5), 1–6.
- Malmendier, Ulrike and Klaus Schmidt**, “You owe me,” *NBER working paper*, 2012, 18543.
- Maniadis, Zacharias, Fabio Tufano, and John A List**, “One swallow doesn’t make a summer: New evidence on anchoring effects,” *The American Economic Review*, 2014, 104 (1), 277–290.
- Martin, Daniel**, “Strategic pricing and rational inattention to quality,” *unpublished manuscript*, 2014.
- Mason, Winter and Siddarth Suri**, “Conducting behavioral research on Amazon’s Mechanical Turk,” *Behavioral Research*, 2012, 44, 1–23.

- Matějka, Filip and Alisdair McKay**, “Rational inattention to discrete choices: A new foundation for the multinomial logit model,” *American Economic Review*, 2015, 105 (1), 272–98.
- McCormick, Patrick T**, “Volunteers and Incentives: Buying the Bodies of the Poor,” *Journal of the Society of Christian Ethics*, 2007, 27 (1), 77–93.
- McGregor, Joan**, “‘Undue Inducement’ as Coercive Offers,” *The American Journal of Bioethics*, 2005, 5 (5), 24–25.
- Mitchell, Matthew F and Andrea Moro**, “Persistent distortionary policies with asymmetric information,” *The American Economic Review*, 2006, 96 (1), 387–393.
- Mjøen, Geir, Stein Hallan, Anders Hartmann, Aksel Foss, Karsten Midtvedt, Ole Øyen, Anna Reisæter, Per Pfeffer, Trond Jenssen, and Torbjørn Leivestad**, “Long-term risks for kidney donors,” *Kidney International*, 2014, 86 (1), 162–167.
- Morgan, MA, Jason Dana, George Loewenstein, S Zinberg, and J Schulkin**, “Interactions of doctors with the pharmaceutical industry,” *Journal of Medical Ethics*, 2006, 32 (10), 559–563.
- National Bioethics Advisory Commission**, “Ethical and policy issues in research involving human participants,” Rockville, MD 2001.
- National Collegiate Athletic Association**, “Division 1 Manual,” Indianapolis, IN July 2015.
- National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, Bethesda, MD.**, *The Belmont report: Ethical principles and guidelines for the protection of human subjects of research* 1978.
- Niederle, Muriel and Alvin E Roth**, “Philanthropically Funded Heroism Awards for Kidney Donors,” *Law & Contemporary Problems*, 2014, 77, 131–144.
- Nussbaum, Martha C**, “Objectification,” *Philosophy & Public Affairs*, 1995, 24 (4), 249–291.
- Rabin, Matthew and Joel L Schrag**, “First impressions matter: A model of confirmatory bias,” *Quarterly Journal of Economics*, 1999, pp. 37–82.
- Raven, John C**, *Guide to the standard progressive matrices: sets A, B, C, D and E*, Pearson, 1960.
- Ripley, Elizabeth BD**, “A review of paying research participants: It’s time to move beyond the ethical debate,” *Journal of Empirical Research on Human Research Ethics*, 2006, 1 (4), 9–19.
- Rosenberg, Tina**, “It’s Time to Compensate Kidney Donors,” *The New York Times* August 7 2015.
- , “Need a Kidney? Not Iranian? You’ll Wait.” *The New York Times* July 31 2015.

- Roth, Alvin E**, “Repugnance as a Constraint on Markets,” *Journal of Economic Perspectives*, 2007, 21 (3), 37–58.
- Sandel, Michael J**, *What money can't buy: the moral limits of markets*, Macmillan, 2012.
- Satz, Debra**, *Why Some Things Should Not Be For Sale: The Moral Limits of Markets*, Oxford University Press, 2010.
- Sen, Amartya**, “Well-being, agency and freedom: the Dewey lectures 1984,” *The journal of philosophy*, 1985, 82 (4), 169–221.
- Sen, Amartya**, *On ethics and economics*, Oxford University Press, 1999.
- Sims, Christopher A**, “Implications of rational inattention,” *Journal of Monetary Economics*, 2003, 50 (3), 665–690.
- , “Rational inattention: Beyond the linear-quadratic case,” *American Economic Review, Papers & Proceedings*, 2006, 96 (2), 158–163.
- Suen, Wing**, “The Self-Perpetuation of Biased Beliefs,” *The Economic Journal*, 2004, 114 (495), 377–396.
- Svitnev, Konstantin**, “Legal regulation of assisted reproduction treatment in Russia,” *Reproductive Biomedicine Online*, 2010, 20 (7), 892–894.
- Tellioglu, G, I Berber, I Yarkin, B Yigit, T Ozgezer, S Gulle, G Isitmangil, M Caliskan, and I Titiz**, “Quality of life analysis of renal donors,” *Transplantation proceedings*, 2008, 40 (1), 50–52.
- Thiessen, C, YA Kim, R Formica, M Bia, and S Kulkarni**, “Written informed consent for living kidney donors: practices and compliance with CMS and OPTN requirements,” *American Journal of Transplantation*, 2013, 13 (10), 2713–2721.
- Toplak, Maggie E, Richard F West, and Keith E Stanovich**, “Assessing miserly information processing: An expansion of the Cognitive Reflection Test,” *Thinking & Reasoning*, 2014, 20 (2), 147–168.
- Vatican Radio**, “Pope Francis meets a group of transplant surgeons; including the mayor of Rome,” http://en.radiovaticana.va/news/2014/09/20/pope_francis_meets_a_group_of_transplant_surgeons/1106931 September 20 2014.
- Wald, Abraham**, *Sequential Analysis*, New York: Wiley, 1947.
- Wazana, Ashley**, “Physicians and the pharmaceutical industry: is a gift ever just a gift?,” *Journal of the American Medical Association*, 2000, 283 (3), 373–380.

- Wilkie, William L, Carl F Mela, and Gregory T Gundlach**, “Does ‘bait and switch’ really benefit consumers?,” *Marketing Science*, 1998, *17* (3), 273–282.
- Woodford, Michael**, “Inattentive valuation and reference-dependent choice,” *Unpublished manuscript*, 2012.
- , “An Optimizing Neuroeconomic Model of Discrete Choice,” *National Bureau of Economic Research*, 2014, *19897*.
- Yang, Ming**, “Coordination with flexible information acquisition,” *Journal of Economic Theory*, 2014, *158* (B), 721–738.

ONLINE-APPENDIX

Sandro Ambuehl

“An Offer You Can’t Refuse? Incentives Change How We Think”

Table of Contents

A Insect experiment: additional analysis	1
A.1 Randomization check	1
A.2 Choice consistency	1
A.3 Estimating reservation prices	2
A.4 Awareness	5
A.5 Reservation prices after distribution of insects	7
A.6 Choice of video clips	8
A.7 Robustness checks	9
A.8 Subjects who watched the encouraging video	12
B Model: additional materials	13
B.1 Balls and urns example	13
B.2 Predictions for the experiment in section 5	15
B.3 General state space	17
B.4 Proofs	20
C Stylized experiment: additional analysis	25
C.1 Additional results	25
C.2 Analysis of ancillary treatments	26
D Experimental materials	28
D.1 Instructions	33
References	49

A Insect experiment: additional analysis

A.1 Randomization check

The four treatments are balanced across demographic characteristics. Table A.6 displays summary statistics of these variables by treatment. For each variable, the table reports the p -value of an F -test for differences in the mean value of the variable across treatments. Of 24 tests conducted, one is significant at the 5% level, and an additional three are significant at the 10% level. This is within the expected range.

A.2 Choice consistency

A participants choices are inconsistent if she rejects a transaction at price p in the multiple price list (MPL) in step 3 of the experiment, but accepts the same transaction in step 4, or vice versa.¹ Table A.7 details the fraction of each of these types of inconsistencies by treatment. It shows that subjects in the low incentive treatments tend to state reservation prices that are too high relative to their behavior in their \$3-treatment decision. No such directional bias is evident for subjects in the high incentive condition. This does *not* point to a difference across treatments, as the decisions that reveal inconsistencies differ across the incentive treatments.

The fraction of inconsistent decisions is somewhat higher than is usually found in the literature on decision making under explicit risk, in which inconsistencies are identified by means of multiple switching in a price list (e.g. Holt and Laury 2002). This may be because the decisions that reveal inconsistencies in this experiment are temporally separated, whereas in the risky decision making literature they are typically presented simultaneously.

Condition	<i>video</i>	<i>no video</i>
Incentive \$3		
Reservation price > \$3 in MPL, accept \$3 in treatment decision	15.03%	16.32%
Reservation price < \$3 in MPL, reject \$3 in treatment decision	1.42%	3.82%
<i>Total</i>	16.45%	20.15%
Incentive \$30		
Reservation price > \$30 in MPL, accept \$30 in treatment decision	4.24%	8.00%
Reservation price < \$30 in MPL, reject \$30 in treatment decision	5.71%	6.37%
<i>Total</i>	9.95%	14.37%

Table A.7: Choice inconsistencies across steps 3 and 4 of the experiment.

¹The choices subjects' made in step 6 of the experiment cannot reveal any inconsistencies, as they are made with different information about the transaction than the treatment decisions in step 4.

<i>Treatment condition</i>					
Incentive	\$30	\$3	\$30	\$3	
Video	Yes	Yes	No	No	
Variable	Mean				<i>p</i> -value
Male	0.55	0.53	0.54	0.54	1.00
Age	21.43	22.01	21.37	21.30	0.34
<i>Ethnicity</i>					
African-American	0.05	0.06	0.07	0.07	0.70
Caucasian	0.57	0.51	0.59	0.56	0.28
East Asian	0.19	0.26	0.19	0.23	0.22
Hispanic	0.07	0.08	0.04	0.04	0.99
Indian	0.03	0.04	0.04	0.07	0.56
Other	0.08	0.05	0.07	0.04	0.70
Monthly spending in USD	251.72	301.40	289.07	288.42	0.44
Year of study ^a	3.50	3.60	3.61	3.47	0.32
Graduate student	0.13	0.15	0.13	0.05	0.07
<i>Field of study</i>					
Arts and humanities	0.16	0.09	0.13	0.11	0.04
Business or economics	0.27	0.36	0.34	0.43	0.09
Engineering	0.20	0.16	0.11	0.12	0.49
Science	0.21	0.23	0.27	0.23	0.47
Social science (excluding business and economics)	0.17	0.17	0.15	0.11	0.59
Political orientation ^b	0.50	0.32	0.27	0.09	0.08
Raven's score ^c	14.77	14.76	14.69	14.68	1.00
CRT score ^d	3.76	3.80	3.50	3.22	0.08
<i>Experience with insects as food</i>					
Has intentionally eaten insects before	0.19	0.22	0.19	0.20	0.71
Grown up in culture that practices entomophagy	1.30	1.27	1.25	1.31	0.92
Grown up eating mostly western foods	0.81	0.73	0.82	0.78	0.06
Had a pet that fed on store-bought insects	0.25	0.25	0.21	0.26	0.68
Knew that this study concerns insect eating	2.64	2.46	2.54	2.49	0.31

Table A.6: Summary statistics and randomization check. The last column displays the *p*-value of the test of joint significance of a regression of the indicated variable on treatment dummies.

^aYear of study only includes undergraduate students.

^bPolitical orientation is measured on a scale of -2 (conservative) to 2 (liberal).

^cRaven's score is measured on a scale of 0 to 24.

^dCRT score refers to performance on the extended version of the test (Toplak et al., 2014) and is measured on a scale of 0 to 6.

A.3 Estimating reservation prices

In section 3.2, I report the estimates of the second stage of a double hurdle model. Here, I describe the estimating equation, present the full set of estimated coefficients, and test the restriction that the coefficients in the selection and amount equations are equal (as the Tobit model imposes).

Double hurdle model There are n subjects, each of whom provide measures of reservation prices for T species each. I use r_{it} to denote individual i 's reservation price for species t and set $y_{it} = 60 - r_{it}$

(\$60 is the highest price offered in any decision in this experiment). The two hurdles are defined as follows:

1. Participation decision (first hurdle)

$$\begin{aligned} d_i^* &= z_i' \alpha + \epsilon_{1,i} \\ d_i &= \begin{cases} 1 & \text{if } d_i^* > 0 \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

2. Amount decision (second hurdle)

$$\begin{aligned} y_{it}^{**} &= x_{it}' \beta + u_i + \epsilon_{2,it} \\ y_{it}^* &= \max(y_{it}^{**}, 0) \end{aligned}$$

The vector of error terms $(\epsilon_{1,i}, u_i, \epsilon_{2,i})$ is normally distributed with mean zero, and variance-covariance matrix

$$\Sigma = \begin{pmatrix} 1 & \rho\sigma_u & 0 \\ \rho\sigma_u & \sigma_u^2 & 0 \\ 0 & 0 & \sigma_\epsilon^2 \end{pmatrix}$$

The observed variable is $y_{it} = d_i y_{it}^*$. The second hurdle contains a subject-specific random-effect term u_i that allows between-subject heterogeneity.

Implementation I estimate this model using the user-written Stata command `xtbreg` by [Engel and Moffatt \(2014\)](#). I estimate two specifications. In the first, I include demographic controls (gender, ethnicity, age, age²) as well as university and species dummies in the amount equation. I also allow for correlation in the error terms across the hurdles (that is, ρ is endogenous).

The results from this specification are reported in the main text. In the second, I do not include any controls, and in the third I include only the demographic controls. I also demonstrate that a double hurdle model is the appropriate specification. To do so, for each of the two specifications I also estimate a version in which I impose the restriction the Tobit model makes, namely that the treatments affect the participation decision and the amount decision in the same way. Formally, I require $\alpha_k = \beta_k$ for all variables k that are included in both x_i and z_i . For each of the specifications, I use a likelihood ratio test to test that restriction. Finally, I also estimate a specification including further controls (see section [A.7](#) for details).

Results Table [A.8](#) displays the results. A comparison of Columns 1 and 3 shows that the estimates are not significantly affected by the inclusion of demographic control variables, and one of Columns

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Reservation price elicited in stage 3						
Restriction Equation	Unrestricted Particip. Amount	Restricted Particip. Amount	Unrestricted Particip. Amount	Restricted Particip. Amount	Unrestricted Particip. Amount	Restricted Particip. Amount	Unrestricted Particip. Amount
High incentive	-0.04 (0.19)	5.97*** (2.26)	-0.06 (0.19)	-0.06 (0.19)	-0.03 (0.19)	6.39*** (2.26)	0.06 (0.19)
Video	-0.05 (0.18)	-0.07 (2.32)	-0.05 (0.19)	-0.05 (0.19)	-0.03 (0.18)	0.55 (2.18)	0.05 (0.19)
High inc. \times Video	0.11 (0.26)	-7.25*** (3.06)	0.13 (0.26)	0.13 (0.26)	0.09 (0.26)	-9.05*** (3.09)	-0.14 (0.26)
Constant	-1.00*** (0.13)	-	-1.00*** (0.13)	-1.00*** (0.13)	-1.01*** (0.13)	-	-1.00*** (0.14)
Controls							
Demographic	No	Yes	No	No	No	Yes	No
Other	No	No	No	No	No	No	No
University dummies	No	Yes	No	No	No	No	No
Species dummies	No	Yes	No	No	No	No	No
σ_u	19.33	19.57	19.69	19.81	19.14	19.66	18.66
σ_ϵ	9.97	9.98	10.66	10.67	10.67	10.67	9.97
ρ	0.46	0.45	0.43	0.40	0.46	0.47	0.48
Log likelihood	-9512.23	-9519.58	-9663.75	-9670.19	-9644.43	-9650.30	-9492.52
Observations	3,276	3,276	3,276	3,276	3,276	3,276	3,276
Subjects	671	671	671	671	671	671	671

Table A.8: Estimation results of the double hurdle model specifications. Each estimation result consists of two equations labeled *particip.* and *amount*, corresponding to the parameters in the participation and amount equations, respectively. Whenever control variables are included, the constant is not easily interpretable, and thus not reported. The model in each even numbered Column is the same as the one immediately to the right, with the restriction that the coefficients of the treatment variables are the same across the two equations. The model in Column 7 parallels that in Column 1 but includes an additional set of control variables described in Section A.7. The number of observations falls short of $671 \times 5 = 3355$ by 79 since the first 79 subjects did not make any decisions about field crickets.

1 and 5 shows that the qualitative results are not substantially affected by the inclusion of university and species fixed effects. Comparing Columns 1 and 7 further reveals that the adding further control variables to the amount equation does not substantially change any qualitative results either.

For the first, second, and third specification, the value of the likelihood ratio test statistic is 14.54, 12.88 and 11.74, respectively. Each exceeds the 5%-critical value of the $\chi^2(4)$ distribution of 9.49. Hence, the restricted versions are rejected in favor of the unrestricted ones. The interpretation is that subjects who do not consume any insect for any price in this experiment are differently affected by the treatment interventions than the remaining subjects, as one would expect given the treatment incentives of \$3 and \$30 are less than half of those subjects' reservation prices.

A.4 Awareness

I obtain the results on subjects' awareness of self-persuasion discussed in Section 3.3 by estimating the following regression model, separately for subjects in the *video* condition and for subjects in the *no video* condition. (Recall that subjects in the *video* (*no video*) condition predicted the reservation price of other subjects in the *video* (*no video*) only.)

$$\widehat{WTA}_{cs}^i = \beta_0 + \beta_1 \mathbb{1}(\text{incentive}_i = \text{high}) + \beta_2 \mathbb{1}(c = \text{high}) + \epsilon_{cs}^i \quad (3)$$

Here, \widehat{WTA}_{cs}^i is subject i 's prediction of the the mean reservation price of subjects in incentive condition c for species s . In words, I regress subject i 's prediction about others' reservation price on a dummy that indicates whether the prediction concerns a previous subject facing $c = \$3$ or $c = \$30$ low incentives. Hence, β_2 is the amount by which subjects' predict incentives change reservation prices for others. I let the intercept vary depending on whether the subject making the prediction was herself offered the high or the low incentives.

I compare subjects' predictions of the effect of incentives on reservation prices to their actual effect. To simplify the comparison, I estimate the following linear model.

$$WTA_s^i = \gamma_0 + \gamma_1 \mathbb{1}(\text{incentive}_i = \text{high}) + \eta_s^i \quad (4)$$

I jointly estimate equations (3) and (4), jointly across both video conditions. I control for gender, ethnicity, age and age², and include university and species fixed effects.

Column 1 of table A.9 displays the estimated parameters of equation (3) for subjects in the *no video* condition. It shows that these subjects predicted that other subjects in the *no video* condition would demand an additional \$4.07 to eat an insect when offered the high rather than the low incentive. Column 2 displays the estimate of the effect of incentives on actual reservation prices, and shows an effect of \$4.45. Subjects' predictions deviate from the measured effect of \$4.45 by a statistically and economically insignificant \$0.37.

Columns 3 and 4 show the respective data for the *video* condition. Subjects predict that the effect of incentives on other subjects in the *video* condition is \$5.16, and thus predict the anchoring effect of \$4.45 (as measured in the *no video* condition) with reasonable accuracy. In reality, however, that effect is countervailed by a sizable self-persuasion effect. These two effects sum to a *negative* \$0.55. Hence, the predictions of subjects in the *video* treatment are wildly off, by a highly significant \$5.70. On average, subjects lack awareness of the self-persuasion effect.

Dependent variable	(1) <i>no video</i>	(2)	(3) <i>video</i>	(4)
	Reservation price predicted	actual	Reservation price predicted	actual
Level with \$3 incentive	20.35 (2.16)	22.31 (2.22)	18.76 (1.22)	20.63 (1.37)
Effect of increase in incentive	4.07*** (0.72)	4.53* (2.55)	5.16*** (0.63)	- 0.63 (2.05)
Effect of predictor's own incentive on prediction	-0.95 (1.60)	-	-2.72** (1.34)	-
Difference predicted vs. actual effect of increase in incentive	-0.46 (2.65)	-	5.79*** (2.03)	-
Observations	2,710	1,355	3,520	1,921
Number of subjects	271	127	352	195

Table A.9: Demographic controls are gender, ethnicity, age and age². Levels are displayed for the mean participant. The first 48 participants at Stanford did not predict others' reservation prices, and are therefore not included in the regression in this table. Estimated using university and species fixed effects. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

Even though subjects do not predict the self-persuasion effect, their predictions are affected by it. Apparently, subjects who have both an incentive and the opportunity to persuade themselves do so. But because they lack awareness of this effect, they project their own lower willingness to accept onto others. Column 3 of Table A.9 shows that amongst subjects in the *video* condition, those who were given the high incentive predict significantly lower reservation prices for others. For those who could not see a video, this effect is just 40% as large, and not statistically significant.

Subjects' beliefs about the effect of incentives on others are heterogenous. Roughly one third of subjects predict that higher incentives decrease reservation prices, and roughly two-thirds predict the opposite. These fractions does not substantively differ across the four treatment cells ($p = 0.79$, test for joint significance of treatment dummies). Hence, while the above results show that subjects on average fail to predict the self-persuasion effect, it is still possible, for instance, that subjects who are

themselves more prone to self-persuasion are more likely to predict a negative effect of incentives on reservation prices. To address this issue, I split the sample into those who predict that incentives lower reservation prices, and those who predict the opposite. I can do so because each subject separately predicted the reservation prices of previous participants in the high and low incentive conditions. These predictions cannot be an ex-post rationalization of subjects' own behavior, since each subject was in only one treatment.

Indeed, subjects who expect self-persuasion in others are indeed more prone to self-persuasions themselves. Amongst the 24.25% of subjects in the *video* condition who believe that self-persuasion outweighs anchoring, the increase in incentives indeed leads to a \$5.83 *decrease* in reservation prices. Amongst the 69.5% of subjects in the *video* condition with the opposite beliefs, by contrast, the same increase in incentives leads to a \$2.49 increase in reservation prices. The difference between these effects is significant at the 10%-level. Hence, subjects are *partially* aware of the effect of incentives on reservation prices.²

	Low Incentive	High Incentive	Difference High - Low
Information			
<i>no video</i>	22.10 (1.82)	26.68 (1.81)	4.58* (2.55)
<i>video</i>	19.64 (1.72)	16.99 (1.69)	-2.65 (2.35)
Difference	-2.46 (2.42)	-9.69*** (2.44)	-7.23** (3.43)

Table A.10: Reservation prices after distribution of insects, in dollars, estimated by the second equation of the double hurdle model using 3,226 observations from $n = 671$ subjects. Standard errors clustered by subjects. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively. Asterisks are suppressed for levels.

A.5 Reservation prices after distribution of insects

In step 5 of the experiment, subjects receive all the insects. This potentially changes reservation prices, due to a multitude of factors. On the one hand, the additional information one can gather by observing the insects is still imperfect about what it is like to ingest them, so that skewed interpretation is still possible. On the other hand, the additional information might lead to some convergence of beliefs towards the true disutility from consuming the insects. Moreover, because the decision that will be carried out with the highest likelihood has already been made, *ex post* rationalization may

²Amongst subjects in the *no video* condition, 31% believe that self-persuasion outweighs anchoring, and 63.1% have the opposite beliefs. The effect of the increase in incentives on those subjects' reservation prices are \$8.39 and \$2.01, respectively. The p -value of the difference is $p = 0.102$.

now affect behavior. Whatever the mechanism, table A.10 displays the mean reservation prices after the distribution of the insects as estimated by the second stage of the double hurdle model.³

A.6 Choice of video clips

Subjects in the *video* condition also select at least four out of a selection of 9 video clips, grouped in bins of three named “Reasons for eating insects”, “Reasons against eating insects”, “Other information about eating insects”. This reveals whether incentives affect the *amount* of information demanded.

Subjects know that they will either watch the selected 6-minute video, or all the clips they selected, but not both. They also know that the chance of the former is 97%. This probability is chosen for reasons of statistical power. There are many possible selections of video clips, each of which could potentially differently affect behavior. By contrast, there are only two selections of 6-minute videos, thus leading to a potentially more pronounced treatment effect.

Empirically, subjects’ choice of video clips reinforces the finding that incentives cause them to demand information in a way that is more favorable to participation. Table A.11 shows that subjects in the \$30-condition select significantly fewer discouraging clips, and significantly more encouraging clips, while the number of other clips is unaffected. Incentives do not affect the total number of clips selected, because most subjects opt for the minimum number of four clips.

	(1)	(2)	(3)	(4)
	Number of clips chosen			
	Encouraging	Discouraging	Other	Total
Effect of higher incentive	0.17*	-0.24**	0.01	-0.07
	(0.10)	(0.11)	(0.08)	(0.11)
<i>Levels</i>				
\$30 incentive	2.29	0.99	1.15	4.42
	(0.06)	(0.08)	(0.06)	(0.11)
\$3 incentive	2.12	1.23	1.14	4.49
	(0.07)	(0.08)	(0.06)	(0.09)
Observations	321	321	321	321

Table A.11: Video clips chosen by incentive condition. The number of observations is smaller for the video clips than for the six minute video as the first 79 participants at Stanford could not choose any clips. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively. Asterisks are suppressed for levels.

³I use the same set of control variables as in the main text. Other estimated parameters: $\sigma_u = 19.25$, $\sigma_\epsilon = 10.87$. $\rho = 0$ is enforced, as the likelihood function is flat when this parameter is endogenous.

A.7 Robustness checks

Video choice and participation probabilities Table A.13 reports estimates of alternative regression specifications of the results on information choice and participation probabilities. First, I estimate all results with a wider range of control variables. In addition to gender, ethnicity, and age, I also control for monthly spending, score in the CRT test, score in the IQ test, college major, as well as variables relating to experience with insects. The latter are whether the subject has voluntarily eaten insects before, whether the subject has grown up eating mostly western foods, and whether the subject has a background in a culture in which entomophagy is common. Second, I estimate Probit and Logit specifications. (For these regressions, I do not include university and species fixed effects, see [Greene 2004](#).)

Variable	(1)	(2)	(3)	(4)	(5)	(6)
	Choose encouraging video					
Effect of high incentive	0.07* (0.04)	0.07** (0.03)	0.32** (0.16)	0.35** (0.16)	0.56* (0.29)	0.58* (0.30)
Method	OLS	OLS	Probit	Probit	Logit	Logit
Demographic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Other Controls	No	Yes	No	Yes	No	Yes
University FE	Yes	Yes	No	No	No	No
Species FE	Yes	Yes	No	No	No	No
Observations	2,000	2,000	1,920	1,920	1,920	1,920
Subjects	400	400	384	384	384	384

Table A.12: Robustness checks for treatment effects on the probability of choosing the encouraging rather than the discouraging video. *Other controls* are monthly spending, score in the CRT test, score in the IQ test, college major, as well as variables relating to experience with insects. The latter are whether the subject has voluntarily eaten insects before, whether the subject has grown up eating mostly western foods, and whether the subject has a background in a culture in which entomophagy is common. The number of observations for the probit and logit models is reduced since for some realizations of control variables, the outcome is predicted perfectly (for instance subjects identifying as neither male nor female). Column 1 replicates the respective specification reported in the main text. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

Reservation prices The estimates of reservation prices with the additional control variables described in the previous paragraph are displayed in Column 5 of Table A.8. None of the results are relevantly different.

In addition to the double hurdle specifications, I also use linear specifications, both with and without additional statistical control variables, and interval regression. I estimate each of these specifications twice. The first time, I only include the 90.61% *complier* subjects. These are the subjects

Variable	(1)	(2)	(3)	(4)	(5)	(6)
	Accept transaction at promised incentive					
Effect in <i>no video</i>	0.22*** (0.05)	0.23*** (0.05)	0.60*** (0.13)	0.61*** (0.13)	0.98*** (0.22)	1.02*** (0.22)
Effect in <i>video</i>	0.33*** (0.04)	0.33*** (0.04)	0.91*** (0.11)	0.93*** (0.12)	1.48*** (0.19)	1.55*** (0.20)
Difference in effects	0.10* (0.06)	0.10* (0.06)	0.31* (0.18)	0.32* (0.18)	0.51* (0.29)	0.53* (0.29)
Method	OLS	OLS	Probit	Probit	Logit	Logit
Demographic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Other Controls	No	Yes	No	Yes	No	Yes
University FE	Yes	Yes	No	No	No	No
Species FE	Yes	Yes	No	No	No	No
Observations	3,307	3,307	3,288	3,288	3,288	3,288
Subjects	671	671	667	667	667	667

Table A.13: Robustness checks for treatment effects on the supply curves. See Table A.13 for explanation of variables. Column 1 replicates the respective specification reported in the main text. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

who accept at least one out of the 10 offers in the experiment to eat an insect in exchange for \$60. The second time, I also include the remaining 9.39% of subjects who categorically refuse to eat insects, even for an amount that is twice the ‘high’ incentive of \$30. These subjects cannot reasonably be expected to be affected by the incentive treatment. Because of their extreme reservation prices (set to \$60 in the linear specifications), however, their presence substantially lowers the statistical precision of the estimates.

Table A.14 reports the results. Considering first the results on the *complier* subjects alone, these specifications replicate the significant interaction between the incentive and video conditions, as in the main text. The estimated coefficient is attenuated in the linear specifications that ignore the censoring issue (Columns 1 and 3). The inclusion of additional control variables only marginally alters the estimated coefficients. Considering the results on the sample including all subjects we find coefficients that are slightly attenuated, and standard errors that are slightly larger, as one would expect from adding what essentially amounts to noise to the data. While the estimated coefficients on the interaction between the incentive and video condition loses statistical significance, they remain largely similar in magnitude to the case in which only *complier* subjects are included, across all three specifications.

Variable	(1)	(2)	(3)	(4)	(5)	(6)
	Reservation price					
Method	OLS	OLS	OLS	OLS	iReg	iReg
Sample	Compliers	All	Compliers	All	Compliers	All
High incentive \times video	-6.55** (3.11)	-5.15 (3.22)	-5.94** (3.00)	-4.89 (3.13)	-7.36* (3.79)	-5.33 (4.38)
High incentive	5.30** (2.42)	4.56* (2.49)	5.04** (2.36)	4.52* (2.44)	5.90** (2.94)	4.68 (3.42)
Video	1.05 (2.16)	-0.65 (2.27)	0.88 (2.15)	-0.59 (2.24)	1.14 (2.65)	-1.19 (3.09)
Constant	11.09 (11.63)	17.27 (12.60)	14.66 (11.65)	19.31 (12.69)	9.16 (14.14)	15.81 (16.67)
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Other controls	No	No	Yes	Yes	No	No
University dummies	Yes	Yes	Yes	Yes	Yes	Yes
Species dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,961	3,276	2,961	3,276	2,961	3,276
Subjects	608	671	608	671	608	671

Table A.14: Robustness checks for treatment effects on reservation prices. See Table A.13 for explanation of variables. iReg refers to interval regression. Standard errors clustered by subject. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

A.8 Subjects who watched the encouraging video

Here I replicate Table 3 comparing subjects in the *video* condition who watched the encouraging video to subjects in the *no video* treatment. This provides supporting evidence for the finding that incentives affect participation probabilities and reservation prices not only through their effect on video choice, but also by changing how subjects interpret the information in the encouraging video. These estimates need to be interpreted with caution, as they are subject to endogeneity and selection bias because the video a subject watched reflects a choice. The estimated coefficients remain largely comparable to those I obtain when all subjects are included.

Incentive	A. Supply curves			B. Reservation prices		
	\$3	\$30	<i>Difference</i>	\$3	\$30	<i>Difference</i>
Information						
<i>no video</i>	37.08 (3.39)	59.43 (3.39)	22.34*** (4.77)	20.11 (1.72)	24.90 (1.86)	4.79** (2.31)
<i>video</i>	40.94 (3.41)	74.38 (2.53)	33.44*** (4.21)	17.64 (1.73)	17.58 (1.45)	-0.05 (2.18)
Difference	3.86 (4.83)	14.95** (4.31)	11.10* (6.37)	-2.47 (2.25)	-7.32*** (2.25)	- 4.85 (3.23)

Table A.15: Replication of Table 3 including only subjects who either had chosen to watch the encouraging video, or were in the control treatment. Standard errors clustered by subjects. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively. Asterisks are suppressed for levels.

B Model: additional materials

B.1 Balls and urns example

This section presents a numerical example of the model in section 4 in which the decision maker can select only one of two information structures.

Suppose a decision maker is offered the opportunity to participate in a win 10 / lose 20 lottery with a 50 / 50 prior (that is, $\pi_G = 10, \pi_B = -20, \mu = 0.5$). Before he decides whether to take this gamble, he can select a source of information and observe a realization of a signal. A source of information produces either “good news” or “bad news”, suggesting participation or abstention. There are two sources of information the individual can choose between, an encouraging and a discouraging one, as depicted in figure B.16. The encouraging source is *ex ante* more likely to produce good news than the discouraging one.

We consider the decision maker’s choice as described, and compare it to the case in which he is offered a \$4 incentive for taking the gamble (that is, $m = 4$). The incentive is paid whenever he accepts to take the gamble, regardless of whether he wins or loses. Hence, in this case, he decides whether to take a win 14 / lose 16 lottery. Which information structures will subjects choose to observe, and what will their resulting beliefs be?

Consider first the condition without incentives. Suppose that the agent participates after good news, and abstains after bad news.⁴ If he chooses the discouraging information source, the chance that the state is good and he will receive good news (‘correctly’) is $\frac{1}{2} \cdot \frac{3}{9}$. The chance that the state is bad, and he will still receive good news (‘mistakenly’) is $\frac{1}{2} \cdot \frac{1}{9}$. Hence, in expectation, observing a signal from the discouraging information source and following it leads to an expected payoff of $\frac{1}{2} \cdot \frac{3}{9} \cdot 10 + \frac{1}{2} \cdot \frac{1}{9} \cdot (-20) = \frac{5}{9}$. By contrast, observing a signal from the encouraging information source leads to an expected payoff of $\frac{1}{2} \cdot \frac{8}{9} \cdot 10 + \frac{1}{2} \cdot \frac{4}{9} \cdot (-20) = 0$. Hence, without an additional incentive for taking the gamble, the decision maker prefers to observe news from the discouraging rather than from the encouraging information source, and participates (only) after receiving good news. This calculus changes once the decision maker is paid an additional \$4 for participation (and hence faces a win 14 / lose 16 gamble). The expected payoff from observing a signal from the discouraging and encouraging information structures are now $\frac{1}{2} \cdot \frac{3}{9} \cdot 14 + \frac{1}{2} \cdot \frac{1}{9} \cdot (-16) = \frac{13}{9}$ and $\frac{1}{2} \cdot \frac{8}{9} \cdot 14 + \frac{1}{2} \cdot \frac{4}{9} \cdot (-16) = \frac{24}{9}$, respectively. Hence, the decision maker now prefers the encouraging information source to the discouraging one.

Thus the increase in incentives cause the decision maker to opt for a different source of information. As a result, he is more likely to observe good news, and becomes more likely to believe winning is more likely than losing. Moreover, without the possibility of information acquisition, he would not participate neither with nor without the \$4 incentive.

⁴The parameters in this example are chosen such that regardless of the chosen information source, participation is always weakly optimal after good news, and abstention is optimal after bad news.

Setting		
State	Good	Bad
Prior	50%	50%
Information sources		
Discouraging		
Encouraging		
Payoff if bet		
No Incentive	10	-20
\$4 incentive	(10 + 4)	(-20 + 4)
Choice		
Incentive	\$0	\$4
Information source	discouraging	encouraging
% who believe $P(\text{win}) > P(\text{lose})$	22%	67%

Table B.16: Example of the model in section 4 with only two information structures to choose from. In this example, with higher incentives for participation, a Bayesian demands to see more encouraging information about the transaction, and thus becomes more likely to believe that winning is more likely than losing.

B.2 Predictions for the experiment in section 5

The experiment in section 5 maps directly to the model in section 4. Here, I formally derive the predictions I test with that experiment.

Slope of the supply curve Consider incentive amounts \bar{m} and \underline{m} with $-\pi_B > \bar{m} > \underline{m} > 0$. Consider two treatments. In the *incentive first* treatment, the agent first learns the incentive $m \in \{\bar{m}, \underline{m}\}$, then gathers information, and finally decides whether to take the bet. In the *picture first* treatment, the agent first only learns that the incentive will be \bar{m} or \underline{m} with probability α and $(1 - \alpha)$, respectively. He then gathers information, then learns the realization of m , and finally decides whether participate. I say that information is instrumental if the agents' betting decision depends on the realization of the informative signal.

Proposition 3. *Suppose the cost-of-information function c has increasing differences. Let $p_b(m)$ and $p_a(m)$ denote the ex ante participation probabilities in the two treatments if the incentive is m . Suppose the chosen information is instrumental for both incentive amounts in both the incentive first and the picture first condition. Then, $p_b(\bar{m}) > p_a(\bar{m})$ and $p_b(\underline{m}) < p_a(\underline{m})$. In particular, the effect of incentives is stronger in the incentive first condition: $p_b(\bar{m}) - p_b(\underline{m}) > p_a(\bar{m}) - p_a(\underline{m})$.*

The proof is in section B.4. Intuitively, the result can easily be seen in Figure B.7. The subjects' objective function in the *picture first* condition regarding information acquisition is simply $\alpha U(p_G, p_B; \bar{m}) + (1 - \alpha)U(p_G, p_B; \underline{m}) = \mu p_G(\pi_G + \hat{m}) + (1 - \mu)p_B(\pi_B + \hat{m})$, where $\hat{m} = \bar{m} + \underline{m}$. Hence, the decision maker chooses the same information structure he would choose if the incentive amount were equal to the expected incentive. This leads to a choice of false positive and false negative probabilities that are between those he would choose if he knew the realization of the incentive amount before acquiring information.

The condition of this proposition is that information is instrumental for both incentive amounts in both information conditions. To see that it is satisfied empirically, compare panels B and C of Table 4. For both incentive amounts in both information conditions, the participation probability is higher by a factor of about 2 or more if the state is good than if it is bad. Hence, the information about the state does affect subjects' participation probability.

If information acquisition cannot depend on the incentive, and the subject follows the signal, this model predicts that the increase in the incentive has no effect at all. A variety of extensions allow incentives to affect participation in the *picture first* condition even if information is instrumental. The most plausible one is that a fraction of subjects remember, perhaps vaguely, the picture they have seen. Once they learn which incentive amount has realized, they mentally revisit the picture, and decide whether to take the lottery. In this case, the *picture first* condition does not entirely preclude incentive-dependent information acquisition. (A related possibility is that some subjects take a screenshot, and further examine the picture once they know what incentive has realized.)

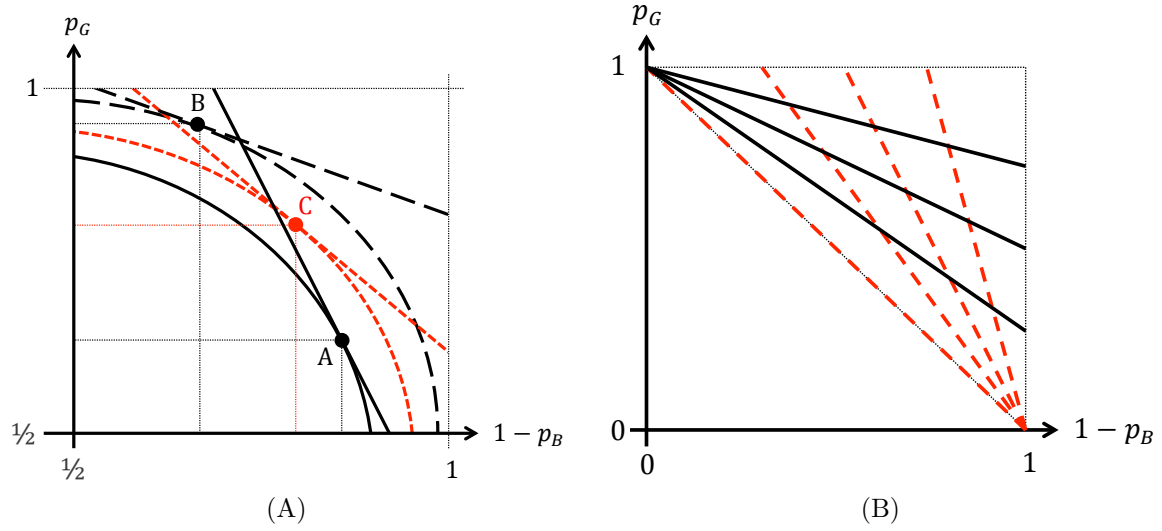


Figure B.7: Panel A. Effects of having to choose an information structure before the incentive amount is known. As in figure 2, points A and B represent optimal information demand with low (\underline{m}) and high (\bar{m}) incentives, respectively. Point C represents the optimal information demand if information is chosen knowing solely that the incentive amount will be \underline{m} or \bar{m} , each with equal probability. Panel B. Contour curves of the posterior beliefs $P(s = G|\text{participate})$ (dashed red lines) and $P(s = G|\text{abstain})$ (solid black lines) for prior $\mu = 0.5$. Both posteriors increase to the north-east. Note that this panel displays the full choice set of state-contingent participation probabilities, $\{(p_G, p_B) \in [0, 1]^2 : p_G \geq p_B\}$.

Optimal posterior beliefs In the experiment in section 5, posterior beliefs are observable. If the cost of information function is proportional to Shannon entropy, the model predicts how they should change across the incentive conditions.

Proposition 4. *If c is proportional to Shannon mutual information, then the increase in the incentive m decreases the posterior probability that the state is good conditional on participating, $P(s = G|\text{participate}, m') < P(s = G|\text{participate}, m)$. It also decreases the probability that the state is good conditional on abstaining, $P(s = G|\text{abstain}, m') < P(s = G|\text{abstain}, m)$.*

The intuition is that as the upside of a bet grows and the downside becomes less severe, a Bayesian decision maker with Shannon mutual information costs will demand to be less sure that the state is good before deciding to participate.

This proposition is different from part (i) of proposition 1 which considers state-contingent participation probabilities rather than optimal posteriors. That the two differ can easily be seen graphically. Figure B.7 plots the level curves of $P(s = G|\text{participate})$, given by $\mu p_G / (\mu p_G + (1 - \mu)p_B)$, and of $P(s = G|\text{abstain})$, given by $\mu(1 - p_G) / (\mu(1 - p_G) + (1 - \mu)(1 - p_B))$. It is evident that one can increase the false positive rate p_G and decrease the false negative rate $1 - p_B$ and achieve either an increase or a decrease in $P(s = G|\text{participate})$.

B.3 General state space

In this section, I show that the assumption of two states of the world made in section 4 is inessential to the qualitative predictions of the model. I extend the model to a general state space $\Omega \subseteq \mathbb{R}$. This allows a decision maker to learn not only about the likelihood that the consequence of a transaction will be good or bad, but also the distribution of the magnitudes of the gains and losses that can be incurred.

I still find that higher incentives increase the demand for information about the upside of the transaction, and decrease the demand for information about the downside. I also find that if the costs of information are proportional to Shannon mutual information, then an increase in incentives increases the probability an agent *ex post* regrets participating *conditional on having participated*. In addition, I show that posterior-separability of the cost of information function (in the sense of Caplin and Dean (2013b)) is sufficient for higher incentives to increase the false positive rate.

Setup An agent whose preferences are quasilinear in money can decide whether or not to participate in a transaction. If he abstains, he receives utility 0. If he accepts, he receives a monetary payment $m \geq 0$, and stochastic, non-monetary utility $u(\omega)$ with $u : \Omega \rightarrow \mathbb{R}$ increasing, and $u(0) = 0$, where Ω is a measurable subset of \mathbb{R} that represents the set of states of the world. The agent is imperfectly informed about ω and thus about his utility from accepting the transaction. His prior distribution

of ω is given by a probability density function $\mu(\omega)$. Before deciding whether or not to accept the transaction, the agent can obtain information about ω . As in section 4, I directly model the agent as choosing state-dependent participation probabilities $p_\omega = P(\text{accept}|\omega)$. The cost of a vector of state dependent acceptance probabilities $(p_\omega)_\omega$ is given by $c((p_\omega)_\omega) \in \mathbb{R}$. Hence, the setting differs from that in section 4 only to the extent that the state space Ω is more general.

Analysis The agent's utility from state dependent acceptance probabilities $(p_\omega)_\omega$ is

$$V = E[(\omega + m)p_\omega] - c((p_\omega)_\omega) \quad (5)$$

where the expectation is taken with respect to the agent's prior beliefs. To illustrate, notice that with incentives m , a perfectly informed agent would accept if $\omega + m \geq 0$ and reject otherwise.

I consider an increase in the incentive for participation from m to $m' > m$. In contrast to the two-states model, such a change now does not only affect the costs of false positives and false negatives, but also changes the set of states ω in which the agent would participate under full information. Nonetheless, an increase in the incentive still leads to the substitution and stakes effects outlined in section 4. On the one hand, higher incentives change the stakes of the decision, and thus lead the agent to acquire a different amount of information. If it causes the agent to purchase more information, he will increase p_ω for those ω in which accepting is optimal, and decrease p_ω for the ω for which rejection is optimal. On the other hand, higher incentives make false positives cheaper, and they make false negatives more expensive. Hence, the agent will acquire a different kind of information; p_ω will increase for all ω , including those for which rejection is optimal. Which of these effects outweighs depends on the cost of information function. Figure B.8 illustrates.

I now show that *posterior separability* (Caplin and Dean, 2013b) of the cost of information function is a sufficient condition for the substitution effect to outweigh (higher incentives increase false positives). To define this property, I use the following notation. I write $p = E(p_\omega) = \int p_\omega \mu_\omega d\omega$ for the unconditional probability that the agent participates if his state-contingent participation probabilities are given by $(p_\omega)_\omega$. Moreover, I write $\gamma_G^\omega = \frac{p_\omega \mu}{p}$ for the value of the posterior density at ω if the agent participates, and $\gamma_B^\omega = \frac{(1-p_\omega)\mu}{1-p}$ if he abstains. The cost function c is *posterior separable* if there exists a strictly convex function $f : [0, 1] \rightarrow \mathbb{R}$ such that c can be written in the following form

$$c((p_\omega)_\omega) = -E[f(\mu)] + pE[f(\gamma_G^\omega)] + (1-p)E[f(\gamma_B^\omega)]$$

The Shannon mutual information cost function is of this form. It arises if the function f in the above expression is the negative of the binary entropy function, $f(x) = x \log(x) + (1-x) \log(1-x)$.

Part (i) of the following proposition formally shows that an increase in incentives increases the false positive rate if the cost function is posterior separable. Specifically, for all $\omega < -m'$, the agent will expect regret if he participates if the state is ω . Because under posterior separability all p_ω increase, this

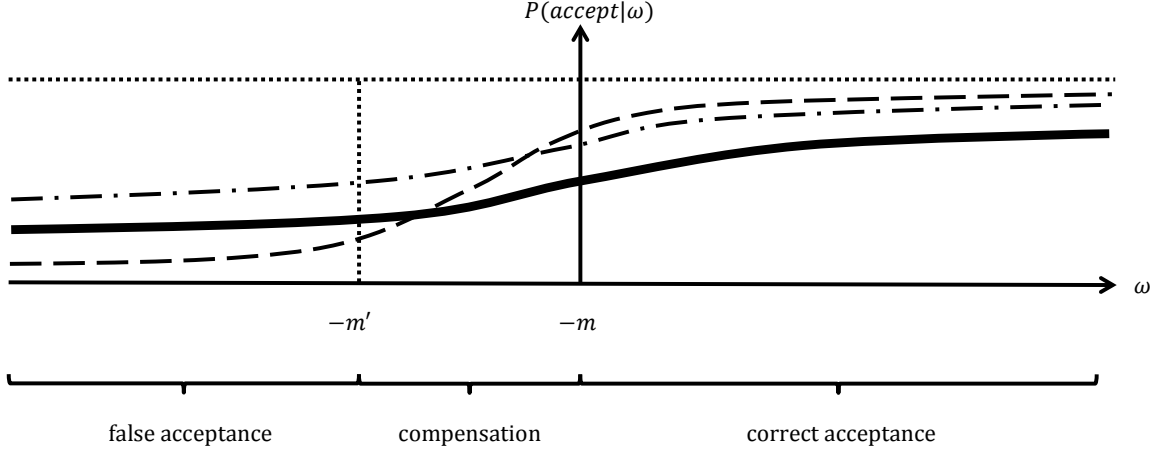


Figure B.8: Effects of a change in incentives on the agents' information with $u(\omega) = \omega$. The figure plots, for each state of the world ω the probability that the agent accepts the offer given his optimal information demand. With incentives $m = 0$, optimal information demand could, for instance, lead to a schedule of acceptance probabilities $P(\text{accept}|\omega)$ depicted by the bold line. An increase in incentives to $m' > m$ increase $P(\text{accept}|\omega)$ for all $\omega > -m$, if the income and stakes effects have the same direction. Depending on the cost of information function, the stakes effect or the substitution effect may dominate for $\omega < -m'$. These cases are illustrated by the dashed and dot-dashed schedules, respectively. In case of a posterior-separable cost of information function, the substitution effect dominates.

means that in particular the false positive probability increases. Note also that posterior separability is also a sufficient condition for the required sign of the cross-derivative in the special version of this model outlined in section 4.

Part (ii) of the proposition helps to better understand the implications of Shannon mutual information costs, by applying a result from Matějka and McKay (2015). It shows that with Shannon costs, the shape of the function $\omega \mapsto p_\omega$ depends on prior beliefs only through the unconditional participation probability p . Hence, it is independent of the shape of the prior belief distribution. Consequently, Shannon mutual information costs limit the effectiveness of certain information provision policies. With Shannon information costs, for instance, it is not possible to decrease the false positive rate without changing the false negative rate merely by externally providing information that disproportionately emphasizes potential downsides.

Finally, part (iii) of the proposition shows that with Shannon mutual information costs, higher incentives increase the probability that the agent *ex post* regrets participation *conditional on having participated*.

Proposition 5.

- (i) If c is posterior separable, and if $m' > m$, then for all $\omega \in \Omega$, $p_\omega(m') \geq p_\omega(m)$.

- (ii) If c is proportional to Shannon mutual information with factor of proportionality λ , then p_ω is strictly increasing in ω , and for any m , $p_{\omega+m}$ is a function only of p and λ .
- (iii) If c is proportional to Shannon mutual information, and $P(\omega \in [-m', m])$ is sufficiently small, then an increase in incentives from m to $m' > m$ increases the probability the agent regrets conditional on having participated.

B.4 Proofs

Proof of proposition 1

- (i) For convenience, let $q_G = p_B$, $q_B = 1 - p_B$, and let $q = (q_G, q_B)$. Without loss of generality, we set $m = 0$ (π_G and π_B can be adjusted accordingly) and $\lambda = 1$. The agent solves

$$\max_q \mu q_G \pi_G + (1 - \mu)(1 - q_B) \pi_B - c(q)$$

We derive the comparative statics regarding the payoffs π_G , and π_B . By the limit assumptions on c , whenever the optimal solution involves a positive amount of information acquisition, the optimal participation probabilities are interior. Since the cost function is strictly convex, any interior solution q^* to the maximization problem is characterized by the first order condition

$$\begin{bmatrix} \mu & 0 \\ 0 & -(1 - \mu) \end{bmatrix} \begin{bmatrix} \pi_G \\ \pi_B \end{bmatrix} = \nabla c(q^*)$$

Let $D^2c = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}$ denote the Hessian of the cost function. Then, totally differentiating the foregoing condition yields

$$\begin{bmatrix} \mu & 0 \\ 0 & -(1 - \mu) \end{bmatrix} \begin{bmatrix} d\pi_G \\ d\pi_B \end{bmatrix} = D^2c(q^*) \begin{bmatrix} dq_G^* \\ dq_B^* \end{bmatrix}$$

We obtain $dq_s^*/d\pi_{s'}$ for $s, s' \in \{G, B\}$ by setting either $d\pi_B = 0$ or $d\pi_G = 0$ and solving for the derivatives of interest. Specifically, left-multiplying the above expression by $(D^2c(q^*))^{-1}$, setting $d\pi_B = 0$ and dividing by $d\pi_G$ yields

$$(D^2c(q^*))^{-1} \begin{bmatrix} \mu & 0 \\ 0 & -(1 - \mu) \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} dq_G^*/d\pi_G \\ dq_B^*/d\pi_G \end{bmatrix}$$

If alternatively, we set $d\pi_G = 0$ and divide by $d\pi_B$, we obtain

$$(D^2c(q^*))^{-1} \begin{bmatrix} \mu & 0 \\ 0 & -(1-\mu) \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} dq_G^*/d\pi_B \\ dq_B^*/d\pi_B \end{bmatrix}$$

Stacking these sideways, and substituting the explicit expression for $(D^2c(q^*))^{-1}$, we thus obtain

$$\frac{dq^*}{d\pi} = \begin{bmatrix} \frac{dq_G^*}{d\pi_G} & \frac{dq_G^*}{d\pi_B} \\ \frac{dq_B^*}{d\pi_G} & \frac{dq_B^*}{d\pi_B} \end{bmatrix} = \frac{1}{\det(D^2c(q^*))} \begin{bmatrix} \mu c_{22} & (1-\mu)c_{12} \\ -\mu c_{21} & -(1-\mu)c_{11} \end{bmatrix}$$

By strict convexity of c , $\det(D^2c(q^*)) > 0$. By Schwarz's theorem, $c_{12} = c_{21}$, which, by assumption, is positive. Hence, $\frac{dp_G}{dm} = \frac{dq_G}{dm} = \frac{\mu c_{22} + (1-\mu)c_{12}}{\det(D^2c(q^*))} > 0$, and $\frac{dp_B}{dm} = -\frac{dq_B}{dm} = \frac{\mu c_{21} + (1-\mu)c_{11}}{\det(D^2c(q^*))} > 0$, which directly implies the claim.

- (ii) This statement is true not only for Shannon mutual information costs, but more generally for posterior-separable cost functions. Moreover, it holds not only within the two-states version of the model in section 4, but for any measurable state-space $\Omega \subseteq \mathbb{R}$, as considered in Appendix B.3. Here, I prove the general statement (part (i) of proposition 5). The two-states case obtains for $\Omega = \{\pi_G, \pi_B\}$.

I use Topkis' theorem to prove the claim. The agent maximizes expression (5) by choosing the family of state-contingent participation probabilities $(p_\omega)_\omega$. We must show that the objective function has increasing differences both in $(p_\omega, p_{\omega'})$ and in (p_ω, m) for all $\omega, \omega' \in \Omega$. As m enters the objective function additively, the latter is trivially true.

To show the former, write $p := E(p_\omega)$, where the expectation is taken regarding the prior μ . By posterior separability of the cost function, taking derivatives of the objective function with respect to p_ω yields

$$\begin{aligned} \frac{\partial V}{\partial p_\omega} &= \omega + m - \mu_\omega [f(\gamma_a^\omega) + f(\gamma_r^\omega)] - pf'(\gamma_a) \left(\frac{\mu_\omega}{p} - \frac{p_\omega \mu_\omega^2}{p^2} \right) \\ &\quad - (1-p)f'(\gamma_r^\omega) \left(\frac{-\mu_\omega}{1-p} + \frac{(1-p_\omega)\mu_\omega^2}{(1-p)^2} \right) \\ &= \omega + m - \mu_\omega [f(\gamma_a^\omega) + f(\gamma_r^\omega)] - f'(\gamma_a) \mu_\omega (1 - \gamma_a^\omega) + f'(\gamma_r) \mu_\omega (1 - \gamma_r^\omega) \end{aligned} \quad (6)$$

where $\gamma = (\gamma_a, \gamma_r)$ denotes the distribution of posterior beliefs under consideration, For $\omega' \neq \omega$, $p_{\omega'}$ enters the above expression only through p . Therefore, we have that $\frac{\partial^2 V}{\partial p_\omega \partial p_{\omega'}} = \frac{\partial^2 V}{\partial p_\omega \partial p} \frac{\partial p}{\partial p_{\omega'}} =$

$\frac{\partial^2 V}{\partial p_\omega \partial p} \mu_{\omega'}$, so that it suffices to show that $\frac{\partial^2 V}{\partial p_\omega \partial p} > 0$. Indeed,

$$\frac{\partial^2 V}{\partial p_\omega \partial p} = -\mu_\omega f''(\gamma_a^\omega)(1 - \gamma_a^\omega) \frac{\partial \gamma_a^\omega}{\partial p} + \mu_\omega f''(\gamma_r^\omega)(1 - \gamma_r^\omega) \frac{\partial \gamma_r^\omega}{\partial p}$$

Since $\frac{\partial \gamma_a^\omega}{\partial p} = -\frac{\gamma_a^\omega}{p}$ and $\frac{\partial \gamma_r^\omega}{\partial p} = \frac{\gamma_r^\omega}{1-p}$, by the definition of γ_a^ω and γ_r^ω we obtain

$$\frac{\partial^2 V}{\partial p_\omega \partial p} = f''(\gamma_a^\omega) \gamma_a^\omega (1 - \gamma_a^\omega) + f''(\gamma_r^\omega) \gamma_r^\omega (1 - \gamma_r^\omega)$$

which is nonnegative due to $f'' > 0$ and $\gamma_a^\omega, \gamma_r^\omega \in [0, 1]$ for all $\omega \in \Omega$.

Proof of proposition 2 A direct application of theorem 1 in [Matějka and McKay \(2015\)](#) shows that for all $\omega \in \Omega$, the state-contingent acceptance probabilities p_ω for $\omega \in \{G, B\}$ in case of Shannon mutual information costs are given by the following equations.

$$\forall \omega \in \{G, B\} : p_\omega = \left[1 + \left(\frac{1}{p} - 1 \right) \exp \left\{ -\frac{1}{\lambda} (\pi_\omega + m) \right\} \right]^{-1} \quad (7)$$

Moreover, a direct application of lemma 2 of [Matějka and McKay \(2015\)](#) shows that the *ex ante* acceptance probability $p = \mu p_G + (1 - \mu) p_B$ satisfies

$$0 = \frac{\mu}{p + (\exp \{ \frac{1}{\lambda} (\pi_G + m) \} - 1)^{-1}} + \frac{1 - \mu}{p + (\exp \{ \frac{1}{\lambda} (\pi_B + m) \} - 1)^{-1}} \quad (8)$$

Letting $\gamma = 1/\lambda$ and rearranging expression (8) shows that the *ex ante* acceptance probability p is given by

$$p = -\mu f(\pi_B + m, \gamma) - (1 - \mu) f(\pi_G + m, \gamma)$$

where $f(x, \gamma) = \frac{1}{e^{\gamma x} - 1}$. We see that if $\lambda \rightarrow \infty$, p grows negative, which means the agent decides based on the prior alone, without acquiring information. The cross-derivative of f is given by

$$\frac{\partial^2 f}{\partial x \partial \gamma}(x, \gamma) = \frac{e^{\gamma x} (\gamma x + 1 + e^{\gamma x} (\gamma x - 1))}{(e^{\gamma x} - 1)^3}$$

Setting $h(\gamma x) = \gamma x + 1 + e^{\gamma x} (\gamma x - 1)$, we see that $h(0) = 0$ and $h' > 0$, and therefore that $h(\gamma x) \geq 0$ if and only if $\gamma x \geq 0$. Because the denominator also has this property, it follows that $\frac{\partial^2 f}{\partial x \partial \gamma}(x, \gamma) > 0$ for all x, γ (except if $\gamma x = 0$ in which case the cross-derivative is not defined). Consequently, $\frac{\partial^2 p}{\partial x \partial \gamma}(x, \gamma) < 0$ for all x, γ , and thus, because $\gamma = 1/\lambda$, we have that $\frac{\partial p}{\partial m}$ is increasing in λ everywhere, as was to be shown.

Proof of proposition 3 Incentives are \bar{m} with probability α and \underline{m} with probability $1-\alpha$, with $\bar{m} > \underline{m}$. Since $U(p_G, p_B; m)$ is linear in m , the agent's expected utility from state-contingent acceptance probabilities (p_G, p_B) is $U(p_G, p_B; \alpha\bar{m} + (1-\alpha)\underline{m}) - \lambda c(p_G, 1-p_B)$. This has increasing differences in $(p_G, 1-p_B, \alpha)$ because $U(p_G, p_B; m) - \lambda c(p_G, 1-p_B)$ has increasing differences in $(p_G, 1-p_B, m)$. Consequentially, by Topkis' theorem, (p_G, p_B) is increasing in α . Hence, the fraction of participants, given by $p = \mu p_G + (1-\mu)p_B$, is therefore also increasing in α .

Proof of proposition 4 This proposition is a special case of the corresponding claim in part (iii) of proposition 5 in Appendix section B.3. Hence, I prove the general claim. A direct application of theorem 1 in Matějka and McKay (2015) shows that for all $\omega \in \Omega$, the state-contingent acceptance probabilities p_ω are given by

$$p_\omega = \left[1 + \left(\frac{1}{p} - 1 \right) \exp \left\{ -\frac{1}{\lambda} \omega \right\} \right]^{-1} \quad (9)$$

where p is the *ex ante* participation probability, $p = E(p_\omega)$.

I assume $P(\omega \in [-m'', -m]) = 0$ for some $m'' > m'$ (this condition is satisfied in the two-state model of section 4 by the assumption that $\pi_G + m' > \pi_G + m > 0 > \pi_B + m' > \pi_B + m$). The extension to the case in which $P(\omega \in [-m'', -m])$ is sufficiently small follows by continuity. Moreover, without loss of generality, I set $m = 0$ (the prior distribution μ can be changed accordingly). We have $P(\text{regret}|\text{participate}) = P(\omega + m' \leq 0|\text{participate}) = \int_{-\infty}^{-m'} \frac{p_\omega}{p} d\mu(\omega)$. By equation (9) and the assumption that $P(\omega \in [-m'', -m]) = 0$, this is equivalent to

$$P(\text{regret}|\text{participate}) = \int_{-\infty}^{-m'} \frac{1}{p + (1-p) \exp \left[-\frac{\omega + m'}{\lambda} \right]} d\mu(\omega)$$

The denominator is decreasing in m' and decreasing in p (the latter is true because the integral is taken only over ω for which $\omega < -m'$, so that $\exp \left[-\frac{\omega + m'}{\lambda} \right] > 1$). Because by the proof of part (ii) of proposition 1, an increase in m' increases p_ω for all ω , it increases p . Consequently, a marginal increase in m' increases $P(\text{regret}|\text{participate})$, as was to be shown.

Specializing to the two states case, the above shows that the posterior $\gamma_G^a = P(G|\text{participate})$ is decreasing in m . The claim about the optimal posterior that leads to rejection, γ_G^r , now directly derives from the Invariant Likelihood Ratio property of optimal information acquisition with Shannon mutual information costs, as characterized in Theorem 1, part (i) of Caplin and Dean (2013b). According to that result, $\frac{\gamma_G^a}{\frac{\pi_G^a + m}{e^\lambda}} = \frac{\gamma_G^r}{\frac{\pi_G^r}{e^\lambda}}$, and hence, $\gamma_G^r = \gamma_G^a e^{-\frac{\pi_G^a + m}{\lambda}}$. We have just shown that γ_G^a is decreasing in m , as is $e^{-\frac{\pi_G^a + m}{\lambda}}$. Hence, γ_G^r is decreasing in m , as was to be shown.

Proof of proposition 5 (Continuous model)

- (i) See the proof to proposition 1, part (ii).

- (ii) The state-contingent acceptance probabilities p_ω are characterized in equation (9). That p_ω is strictly increasing in ω can be read off directly from that equation. To see that for any m , $p_{\omega+m}$ is a function only of p and λ , note that a change in incentives can simply be modeled as a shift in the prior μ , and consider equation (9). In particular, this shows that the shape of the function p_ω is independent of the shape of the prior μ .
- (iii) See the proof to proposition 4.

C Stylized experiment: additional analysis

C.1 Additional results

Result A1: With higher incentives, subjects participate at lower posteriors. Proposition 4 shows that if the costs of information are proportional to Shannon mutual information, then as incentives for participation rise, subjects will change their information acquisition such that if they participate, they will do so at less confident posterior beliefs that the state is good. The posteriors of those taking the gamble are of interest because it is precisely the fraction of subjects who will regret their decision amongst those who decided to take the bet. The experiment in section 5 allows me to test this proposition. We observe the posteriors at which subjects decide to take or refuse the gamble, for each treatment condition.

In the experiment, incentives affect this posterior for two reasons. First, as predicted by the model in section 4 they affect the optimal acquisition of information, and hence the chosen distribution of posterior beliefs. Second, and independently of information acquisition, if subjects' risk preferences are heterogenous, then more subjects will participate if the incentive is high than if it is low for any given posterior. The *picture first* condition, in which incentives cannot affect information acquisition, identifies the latter effect alone. The *incentive first* condition identifies the combined effect.

Table C.17 displays the distributions of objective posterior beliefs. In the *picture first* condition, the posterior at which subjects decide to participate is 11.56 percentage points lower if the incentive is high. In the *incentive first* condition, this number is 19.13. The 7.57 percentage points difference in these effect sizes shows how incentives affect information acquisition—when the incentive is high, subjects halt their information acquisition at lower levels of confidence .

By contrast, no statistically significant difference between the *incentive first* and *picture first* conditions is apparent for the posteriors at which subjects decide to refuse the gamble, although such an effect is predicted by proposition 4. Note, however, that the coefficient does have the predicted sign.

Result A2: Higher incentives do not change the deviation of elicited from Bayesian posteriors. Result 2 in section 5.2 has shown that higher incentives lead to more overoptimism. Why is that the case? There are two mechanisms through which this could happen. On the one hand, with higher incentives, subjects will rationally elect a different distribution of posterior beliefs. If the magnitude of the deviation of elicited from Bayesian posteriors varies with the Bayesian posterior, this will translate into varying overoptimism overall. (Graphically, this idea corresponds to a movement *along* the curve in figure 5.) On the other hand, it is conceivable that incentives alter how much subjective beliefs deviate from Bayesian beliefs, for any posterior. (This is represented by a *change* of the curve in figure 5.) To distinguish between these mechanisms, Columns 2 and 3 of Table C.18 display the relation between subjective and Bayesian posterior beliefs separately for the high and the

Bayesian posteriors by action						
	Incentive \$0.50	Incentive \$3	Difference High - Low	Incentive \$0.50	Incentive \$3	Difference High - Low
	A. Bet taken			B. Bet refused		
Information						
<i>picture first</i>	77.96 (2.58)	66.40 (1.89)	-11.56*** (3.10)	39.64 (1.86)	18.67 (2.16)	-20.98*** (2.76)
<i>incentive first</i>	82.06 (2.64)	62.93 (1.89)	-19.13*** (3.20)	40.90 (1.81)	18.10 (2.27)	-22.80*** (2.79)
Difference before - after	4.10 (3.60)	3.46 (2.56)	-7.57* (4.53)	1.26 (2.54)	-0.56 (0.30)	-1.82 (3.93)

Table C.17: Bayesian posteriors depending on betting decision. Based on 3502 observations from 953 subjects. Standard errors clustered by subject. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively. Asterisks are suppressed for levels.

low incentive condition. They are similarly close across these two conditions (Column 4 shows that a statistically significant difference at the 5% level occurs only for the elicited posterior 0.25), suggesting that incentives change the overall extent of overoptimism because they lead to a different distribution of Bayesian posteriors, not because they induce people to distort a given Bayesian posterior to a different extent.

C.2 Analysis of ancillary treatments

There were two ancillary treatments. In one, subjects decided whether to take a win \$3 / lose \$3 bet. In the other, they decided whether to take a win \$0.5 / lose \$0.5 bet. In both cases, they knew what their choice would be before they examined the picture. Table C.19 shows how often subjects choose to take the bet in each treatment, both averaged over states, and separated. Subjects take the gamble more often as stakes increase. This is because the false negative rate drops by a significant 7.02 percentage points (s.e. 2.92) whereas the false positive rate is not statistically significantly different (it increases by 3.47 percentage points (s.e. 2.72) as stakes increase). While the decrease in the false negative rate is consistent with the model in section 4, that model would also predict a decrease in the false positive rate as stakes increase.

Moreover, Figure C.9 plots estimates of Bayesian posteriors against elicited posteriors for decisions in the ancillary treatments only, as Figure 5 had done for the main treatments. Again, elicited posteriors track Bayesian posteriors remarkably closely. The only statistically significant deviation is at the elicited posterior 0.25.

Variable	(1)	(2)	(3)	(4)
Incentive	Objective Posterior			<i>Difference</i>
	Both	High	Low	
Elicited Posterior				
0	0.023* (0.013)	0.014 (0.014)	0.034 (0.024)	-0.02 (0.028)
0.05	0.087 (0.027)	0.089 (0.036)	0.084 (0.036)	0.005 (0.046)
0.15	0.161 (0.034)	0.137 (0.046)	0.180 (0.049)	-0.042 (0.069)
0.25	0.316* (0.036)	0.214 (0.047)	0.383** (0.046)	-0.169*** (0.061)
0.35	0.355 (0.034)	0.415 (0.068)	0.331 (0.040)	0.084 (0.081)
0.45	0.441 (0.034)	0.498 (0.059)	0.413 (0.042)	0.084 (0.073)
0.55	0.584 (0.029)	0.543 (0.038)	0.635* (0.044)	-0.093 (0.058)
0.65	0.677 (0.032)	0.671 (0.039)	0.687 (0.053)	-0.016 (0.064)
0.75	0.668** (0.035)	0.625** (0.046)	0.750 (0.057)	-0.125* (0.074)
0.85	0.768* (0.042)	0.773 (0.051)	0.760 (0.064)	0.014 (0.076)
0.95	0.883** (0.030)	0.906 (0.039)	0.853* (0.051)	0.053 (0.066)
1	0.937** (0.028)	0.918** (0.041)	0.955* (0.025)	-0.037 (0.037)
Observations	2,012	1,006	1,006	-
#Subj	503	503	503	-

Table C.18: All regressions pool across both information conditions (*incentive first* and *picture first*). Standard errors clustered by subject. Asterisks indicate the level of statistical significance by which objective and elicited posteriors differ. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

State	(1) Both	(2) Bad	(3) Good
Levels			
Treatment $(-3, +3)$	48.64 (1.62)	29.19 (2.07)	68.09 (2.15)
Treatment $(-0.5, +0.5)$	43.37 (1.61)	25.72 (1.99)	61.03 (2.26)
Treatment effect	5.26** (2.23)	3.47 (2.72)	7.06** (2.92)
Number of observations	1906	969	937
Number of subjects	953	735	719

Table C.19: Unconditional and state-dependent participation probabilities in the ancillary treatments. Standard errors clustered by subjects. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively. Asterisks are suppressed for levels.

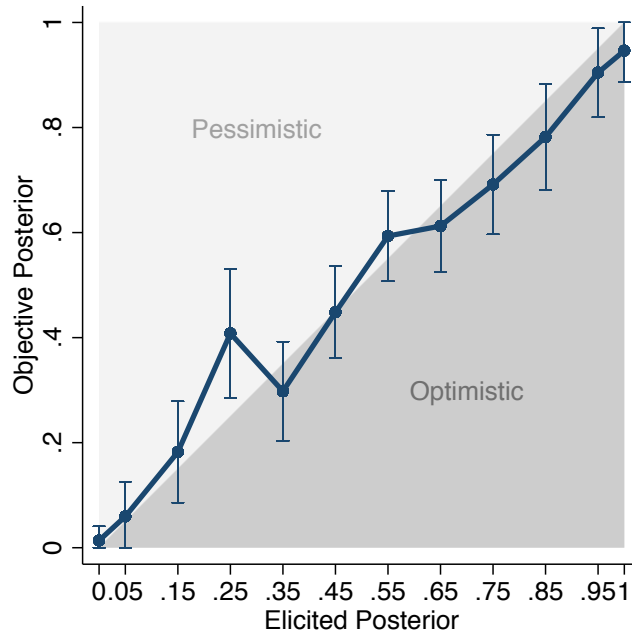


Figure C.9: Elicited and objective posteriors for ancillary treatments only. Whiskers indicate 95% confidence intervals. Standard errors clustered by subject.

D Experimental materials

Figure D.10 displays photographs of the insects used for the experiment in section 3. The following is a transcription of the videos used in that experiment. The videos are available at <https://youtu.>

[be/HiNnbYuuRcA](https://youtu.be/HiNnbYuuRcA) (“Why you may want to eat insects”) and <https://youtu.be/ii4YSGOEcRY> (“Why you may not want to eat insects”).

Transcription: Why You May Want to Eat Insects Five reasons you should consider eating insects. For your own personal health, and for the overall health of the planet, and, most importantly, for your pleasure, you should be eating more insects. This isn’t meant as a provocative, theoretical idea. Here are five very serious reasons why you should consider increasing your insect intake.

First, insects can be yummy. You’d think that insects would have a pungent, unusual aroma. But they are actually very tasty, and considered a delicacy in many parts of the world. Also, like tofu, they often take on the flavor of whatever they’re cooked with. That’s why we are on the verge of a real insectivorous moment in consumer culture. The Brooklyn startup Exo just started selling protein bars made from ground cricket flour, and the British company Ento sells sushi-like bento boxes with cricket-based foods. The restaurant Don Bugito in San Francisco’s Mission district offers creative insect-based foods inspired by Mexican pre-hispanic and contemporary cuisine. “I am trying to bring a solution into the food market which is introducing edible insects” [Monica Martinez, owner of Don Bugito]. New cookbooks are entering the market, such as Daniella Martin’s “Edible”, or van Huis et al.’s “The insect cookbook”. Don Bugito’s reviews on yelp are glowing. Most Americans need some courage to take a bite. But once they do, they are pleasantly surprised. Morgane M., from Sunnyvale, CA describes her experience: “I saw their stand at the Ferry Building farmers market and decided to take the plunge. I tried the chili-lime crickets and they were surprisingly good! For the curious-but-apprehensive: the chili-lime crickets taste like flavorful, super crunchy (almost flaky) chips. That’s it. If you’ve ever had super thin tortilla chips, you’ll have an idea what to expect.” Other people liked them even more. For example Nelson Q. from Las Vegas, NV: “This Pre-Hispanic Snackeria has made me a fan They had the most interesting menu items of the evening at Off The Grid ... Would I try insects again??? Yessir!...ALOHA!!! ” Rodney H. from San Francisco agrees: “It’s great! And the mealworms add kind of a nice, savory quality to it. You never would guess that you’re eating an insect.”

Second, insects are a highly nutritious protein source. “Insects are actually the most ... one of the most efficient proteins on the planet“ [Monica Martinez]. It turns out that pound for pound, insects provide much higher levels of protein compared to conventional meats like beef, chicken, and fish. While eggs consist to just 12% proteins, and beef jerky clocks in at 33%, a single pound of cricket flour has 65% protein. That’s twice as much as you get in beef jerky! Insects also have much higher levels of nutrients like calcium, iron, and zinc. They are also good sources of vitamin B12. That’s an essential vitamin that’s barely found in any plant-based foods and thus can be difficult for vegans to come by.

Third, our objection to eating insects is arbitrary. Your first reaction to this movie was probably a sense of dislike. But there’s nothing innate about that reaction. For one, billions of people already eat



Figure D.10: Insects eaten by subjects. A. House cricket (*acheta domesticus*) B. Mole cricket (*gryllotalpae*) C. Field cricket (*gryllus bimaculatus*) D. Mealworm (*tenebrio molitor*) F. Silkworm pupa (*bombyx mori*).

insects in Asia, Africa, and Latin America every day. More generally, the animals considered to be fit for consumption vary widely from culture to culture for arbitrary reasons. Most Americans consider the idea of eating horses or dogs repugnant, even though there's nothing substantial that differentiates horses from cows. Meanwhile, in India, eating cows is taboo, while eating goat is common. These random variations are the results of cultural beliefs that crystallize over generations. But luckily, these arbitrary taboos can be defeated over time. There was a time when raw fish – served as sushi – was seen as repugnant in mainstream US culture. Now it's ubiquitous. Soon, insects – which are closely related to shrimp – may be elegant hors d'oeuvres.

Fourth, insects are more sustainable than chicken, pork, or beef. “I think the biggest problem for United States right now is we eating too much cattle, too much meat” [Monica Martinez.] Insects are a serious solution to our increasingly pressing environmental problems. It takes 2000 gallons of water to produce a single pound of beef, and 800 gallons for one pound of pork. How much do you think is required for a pound of crickets? One single gallon! Producing a pound of beef also takes thirteen times more arable land than raising a pound of crickets. It needs twelve times as much feed, and produces 100 times as much greenhouse gases. These very handsome environmental benefits are why the UN has released a 200 page report on how eating insects could solve the world's hunger and environmental problems just two years ago. Needless to say, the UN strongly advocates for insects as a food source. And it's not just the UN. In 2011, the European Commission has offered a four million dollar prize to the group that comes up with the best idea for developing insects as a popular food.

Five, we already eat insects all the time. The majority of processed foods you buy have pieces of insect in them. The last jar of peanut butter you bought, for instance, may have had up to 50 insect fragments. A bar of chocolate can have about 60 fragments of various insect species. Some experts estimate that, in total, we eat about one or two pounds of insects each year with our food. These insects pose no health risks. The FDA does set limits, but they are simply set for aesthetic reasons in other words, so you don't actually see them mixed into your food.

To summarize, these are five very compelling reasons to give it a try. Five, we already eat insects all the time anyway. Four, insects are more sustainable and ethical than chicken, pork, or beef. Three, our objection to eating insects is completely arbitrary. Two, insects are a highly nutritious protein source. One. “Most of people react really, really positive” [Monica Martinez]. Insects can be very tasty!

Transcription: Why You May Not Want to Eat Insects Four reasons you may want to avoid eating insects. Reason 1. Some cultures eat insects. But to those of us who are not used to it, insects can be... well, see for yourself. [American tourist in China] “Welcome to eating crazy foods around the world with Mike. And we're in China. If I've learned one thing about China it's they will eat absolutely everything. So you have caterpillars and you have butterflies. The pupae is what the caterpillar turns into before it turns into a butterfly. ... they don't look very appealing at all. But ...

try everything once. So, up to the face. Hhh.” [Eats puppae.] “Not good. Ugh ... it ... it popped. It popped! It’s just ... it’s just too much for me.” [Throws remaining pupae into trash bin.] [Bear Grylls] “Whoa! Ready for this? Oh my goodness! Pfh! This one has been living in there a very, very long time. I’m not gonna need to eat for a week after this. Pfh.” [Eats live beetle larva.] “Argh! This actually ranks as one of the worst things I’ve ever, ever eaten!”

Reason 2. Insects have many body parts. Most of those parts we do not usually eat in other animals. Let’s see those parts... [Biology student] “Let’s take a closer look at some of the structures we see on this grasshopper. So the first thing I want to point out is that it has six legs. There are two pairs. Here is the pair of hindlegs. There’s a pair in the middle here, on the middle segment of the thorax. Ok, those are the midlegs. And then there’s another pair on the front here, those are the forelegs or prolegs. Ok? So there’s six altogether, all insects have six legs, or three pairs of legs, it’s characteristic of the class. Ok? So we also can see, right up here, there are a pair of wings. On each side of the body there are two wings. The forewing, k? – as in the one in front – and this is the hindwing down here, ok? So there are four wings on this animal. Other insects only have two, some have none. Now we’ll move up to the head. The first thing you’ll notice is this pair of long antennae. Ok, we’ve seen antennae in other animals. So, clearly, those are involved ... they have a sensory function. They’re usually involved in a tactile, or a touch sensory function. Some of them are used in chemoreception, which would be like a smell or taste. And speaking of sensory organs, we got one more here, which we would be remiss to not mention, uhm, which is the large compound eye here. So, I’ve made an incision on the dorsal surface of this grasshopper. Ok? And I’ve peeled back the exoskeleton. And before I go digging too much, uhm, it’s going to be difficult to see many structures, but on these individuals it’s very easy to see, uhm, all of these very large and pronounced little sort of tubular looking structures. There’s one right there. Those are all eggs.”

Reason 3. When you eat an insect, you eat ALL of it. In particular, its digestive system, including its stomach, intestine, rectum, anus, and whatever partially digested food is still in there. [Biology student] “Now, if we move on to the digestive system... there is a mouth, of course, we talked about that being down here, ok? The mouth opens into a small pharynx, ok? And then it basically opens up into this large, dark, thin-walled sack right here, ok? This is the crop. Ok, so this is basically a food storage pouch right in here. So ... getting to the stomach, that’s what we find next, this thin-walled, sort of darker colored sack right here, which I’ve just broken a little bit, that, uhm, is the stomach, all in here, ok? Below the stomach we find this slightly darker and a bit more muscular tube right here. That is the intestine. And the intestine opens into a short rectum and an anus.”

Reason 4. Edible insects are perfectly safe to eat. Nonetheless, we tend to associate insects with death and disease. Even if we know that eating some insects is harmless, this association is difficult to overcome. [Nature film maker] “Just a few days ago, one of those gaur was killed by a tiger in the night. This carcass is now probably about five days old, and, as you can see, absolutely riving with maggots of many different species.”

D.1 Instructions

D.1.1 Insect Experiment

This section reproduces the instructions for the *video* treatment. The instructions for the no information treatment are identical, except that no video is mentioned.

Welcome

This is a study of individual decision-making and behavior. Money earned will be paid to you in cash at the end of this experiment.

YOUR DECISIONS IN THIS EXPERIMENT

Your decisions in this experiment determine your payment for this study, and what kind and amount of food products you will eat. There is *no deception* in this experiment: Whenever we tell you that something will happen if you choose a given option, then that thing *will* happen if you choose that option. Whenever we tell you that the computer will do something with a specified chance, then to computer *will* do that thing with exactly that chance. Anything else would violate the IRB protocol under which we run this study (IRB34001).

WHAT YOU MAY AND MAY NOT DO

On your desk you find a pen, paper, and a calculator. You may make use of any of them if you wish. You are not required to do so.

This is a study of **individual** decision-making, which means you are **not** allowed to talk during the study. If you have any questions, please raise your hand and we will come and answer your questions privately.

Please do not use cell phones or other electronic devices until after the study is over. Also, do not browse the internet, or check your emails. If we do find you doing one of these things, the rules of the study require us to deduct \$20 from your payment. The only exception to this rule concerns the calculators that are provided on your desk.

Payment

This study has three parts. Parts 1 and 3 concern food items. Part 2 does not concern food items. Each part has several steps or rounds.

You will be paid at least \$15 if you complete this experiment (this includes your showup payment of \$5), regardless of your choices.

In addition, you receive payment for each part.

Part 1

1. You automatically receive \$20 if you complete part 1 *and* follow through with the decisions you make in this part.
2. Depending on your choices and luck, you may win additional money.

Part 2

You may win or lose up to \$10 from in this part. This depends both on your choices and on luck. Any earnings will be added to your payment from part 1. Any losses will be discounted from your payment in part 1.

Part 3

Depending on your choices and luck, you may earn additional money from this part.

At the end of this experiment, the computer will randomly select *exactly one* decision from part 1, *exactly one* decision from part 2, and *at most one* decision from part 3. Only these three decisions will determine your payment and consumption of food items in this study. Hence,

*In each part, you should make each decision
as if it is the only one that counts—because it might be!*

The payment for participation in this study will be paid to you in cash at the end of this study.

Decision Lists

All parts of this experiment involve decision lists similar to the one below. What X is will vary across different decision lists.

Do X. In exchange receive \$0.	<input type="radio"/>	Do not participate in the transaction on the left
Do X. In exchange receive \$2.	<input type="radio"/>	Do not participate in the transaction on the left
Do X. In exchange receive \$4.	<input type="radio"/>	Do not participate in the transaction on the left
Do X. In exchange receive \$6.	<input type="radio"/>	Do not participate in the transaction on the left
Do X. In exchange receive \$8.	<input type="radio"/>	Do not participate in the transaction on the left

Each line is a separate decision.

On each line, you are asked to select either the option on the right, or the option on the left.

If at the end of the study, the computer randomly chooses your payment from this study to be determined by a decision list, here's what will happen: The computer will randomly draw one line from that price list. Your payment will be determined according to the decision you made on that line. Your choices have *absolutely no* bearing on the line the computer may select.

Hence it is in your best interest to select on each line the option you genuinely prefer!

For instance, suppose that you filled in the decision list like this:

Do X. In exchange receive \$0.	<input type="radio"/>	<input checked="" type="radio"/>	Do not participate in the transaction on the left
Do X. In exchange receive \$2.	<input type="radio"/>	<input checked="" type="radio"/>	Do not participate in the transaction on the left
Do X. In exchange receive \$4.	<input checked="" type="radio"/>	<input type="radio"/>	Do not participate in the transaction on the left
Do X. In exchange receive \$6.	<input checked="" type="radio"/>	<input type="radio"/>	Do not participate in the transaction on the left
Do X. In exchange receive \$8.	<input checked="" type="radio"/>	<input type="radio"/>	Do not participate in the transaction on the left

Suppose that the computer randomly selects the decision on the first line to be carried out. On this line you selected the option on the right. Therefore, you will not do X, and you will neither earn or nor lose any additional money.

Suppose instead that the computer randomly selects the third line for payment. On that line, you selected the option on the left. Therefore, you will do X, and \$4 will be added to your earnings from this experiment.

Most people begin a decision list by choosing the option on the right and then switch to choosing the option on the left (as in the example above). For your convenience, once you click a given option, the computer automatically selects the option on the right on all lines above the line you clicked, and automatically selects the option on the left on all lines below the line you clicked.

Part 1 (involving food items)

Most of the decisions you will make in part 1 are about whether or not you want to consume some food item in exchange for a specified amount of money. Many of the food items may be unfamiliar and / or unappealing to you.

IMPORTANT:

*You will always be given the choice not to eat any food item at all! You will **never** be forced to eat any food item!*

All food items will be eaten in the small waiting room next to this room. At each time, there will be at most one participant in that room. Hence, except for the study staff, consumption of the food items occurs in private.

Recall: This study involves *no deception*. Whenever we tell you that something will happen if you choose a given option, then that thing *will* happen if you choose that option. In particular, if you decide to eat a food item in exchange for a specified amount of money (and that decision is selected to be carried out by the computer), then you will get that amount only if you completely eat the food item.

What happens if during the experiment, I decide to eat some food item for some price, and at the end of the experiment, when I should eat the item, I change my mind?

You are allowed to finish the study without eating the food item. In this case, because you did not follow through with your decisions, you will not receive the \$20 you would have received had you instead completed part 1 and followed through with your decision.

Example

Suppose you are given the offer

"Eat [specified food item]. In exchange, receive \$5."

Suppose your decision on this offer is randomly selected to be carried out at the end of this study.

If you **accept** the offer:

- If, at the end of this study, you **follow through** with your decision and eat the food item, you will receive the \$5 you were promised in the offer. In addition, you will receive \$20 for following through with your decision. Hence, your **total** payment from part 1 will be **\$25**.
- If, at the end of this study, you **decide not to eat** the food item, even though you have accepted the offer, then you will *not* receive the \$5 you were promised in the offer. You will also forfeit the \$20 you would have received for following through with your decision. Hence, your **total** payment from part 1 will be **\$0**.

If you **reject** the offer:

- You will receive the \$20 you are promised for following through with your decision. Because you rejected the offer, you will not eat a food item, and you will not receive any additional money for part 1. Hence, your **total** payment for part 1 will be **\$20**.

Main Decisions

Some participants will be presented with offers to eat specified food items in exchange for \$3. For instance, they will decide whether or not to accept offers such as this one:

Eat [specified food item]. In exchange, receive \$3.

Other participants will be presented with offers to eat specified food items in exchange for \$30. Those participants will decide whether or not to accept offers such as this one:

Eat [specified food item]. In exchange, receive \$30

All participants will be offered exactly the same food items, and in the same amount, regardless of how much they are paid. The only difference is that some participants will be offered a higher amount of money for eating any given food item.

There are five different food items in this study. For each of the food items you will make a decision such as the above. You will either be offered \$3 in exchange for eating the item in all these decisions, or you will be offered \$30 in exchange for eating the item in all of these decisions.

There is an 80% chance that at the end of this study, the computer selects one of these five main decisions to be carried out.

Part 1 consists of the following 8 steps.

Step 1: Learn how much money you will be offered in the main decisions

At the beginning of the experiment, you will click a button to make the computer randomly decide whether you will be offered \$30 in all five main decisions, or whether you will be offered \$3 in all five main decisions.

The computer will immediately tell you which one got selected.

Step 2: Video

The food items in this study may be unfamiliar to you. To help you with your decisions, you will choose videos to watch.

You will watch the video you selected before you make any decisions regarding food items.

Step 3: Decision Lists, round 1

You will fill in five decision lists similar to the one below, one for each of the five food items.

Eat [specified food item]. In exchange receive \$0.	<input type="radio"/>	Do not participate in the transaction on the left
Eat [specified food item]. In exchange receive \$2.	<input type="radio"/>	Do not participate in the transaction on the left
Eat [specified food item]. In exchange receive \$4.	<input type="radio"/>	Do not participate in the transaction on the left
Eat [specified food item]. In exchange receive \$6.	<input type="radio"/>	Do not participate in the transaction on the left
Eat [specified food item]. In exchange receive \$8.	<input type="radio"/>	Do not participate in the transaction on the left
...	<input type="radio"/>	...

The food items will be described in words. No further information will be given.

There is a 7 percent chance that at the end of this study, the computer selects one of the decisions you make in one of these five lists to be carried out

Step 4: Main decisions

It is at this stage that you will make the five main decisions that were described before. The food items will be described in words. No further information will be given. There is a **80% chance** that one of the decisions you make in this part will be carried out.

Step 5: Survey

You will answer some survey questions. Your answers do not affect your payment from this study.

Step 6: Food items handout

Each participant will be given five containers, each containing one of the five food items. You may closely inspect the food items.

Step 7: Decision Lists, round 2

You will fill in five decision lists as before, one for each of the five food items.

There is a **7 percent** chance that at the end of this study, the computer selects one of the decisions you make in one of these five lists to be carried out

Step 8: Additional decisions

You will make a number of additional decisions. You will learn what these are later. There is a **6 percent** chance that one of these decisions will be carried out for payment.

Summary of part 1

Let's recap. Here's how part 1 proceeds:

1. You learn how much money you will later be offered in exchange for eating food items (in step 4 below).
2. The food items in this study may be unfamiliar to you. To help you with your decisions, you may select choose videos to watch.
3. **Decision lists, round 1.** You will fill in five decision lists, one for each of five different food items. The food items will be described in words, but no further information will be given.

*There is a **7% chance** that one of these decisions will be carried out.*

4. **Main decisions.** You will make five decisions whether or not to participate a transaction such as this:

"Consume [specified food item]. In exchange, receive \$[payment]".

You will either be offered *payment* = \$30 in all five decisions, or you will be offered *payment* = \$3 for all five decisions. Again, the food items will be described in words, but no further information will be given.

*There is an **80% chance** that one of these five decisions will be carried out.*

5. You will complete a short survey.
6. Food items will be handed out, and you will be able to inspect them as closely as you wish.
7. **Decision lists, round 2.** You will fill in five decision lists, one for each of five different food items.

*There is a **7% chance** that one of these decisions will be carried out.*

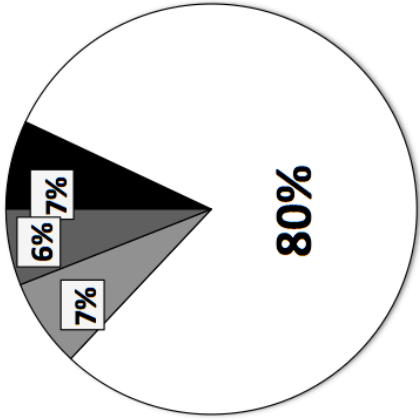
8. Additional decisions. These will be explained later.

*There is a **6% chance** that one of these decisions will be carried out.*

At the very end of this study, the computer will randomly select exactly one of your decisions from this part to be carried out, according to the chances specified above.

This chart illustrates the chances that a decision from any given step of part 1 will be chosen to be carried out:

Chance that a decision from a given part will be carried out



- First round of decision lists
- Main decision
- Second round of decision lists
- Additional Decision

Parts 2 and 3

The instructions for parts 2 and 3 of this study will be displayed on your screen right before those parts begin.

You now have 5 minutes to read through these instructions on your own pace.

If you have any questions about this study, please raise your hand. We will come by and answer them privately.

Recall

*This is a study of **individual** decision making. Hence, you are not allowed to talk.*

Please do not use cell phones or other electronic devices until after the study is over. Also, do not browse the internet, or check your emails. If we do find you doing one of these things, the rules of the study require us to deduct \$20 from your payment. The only exception to this rule concerns the calculators that are provided on your desk

Additional decisions: Estimate what other participants did

48 participants have previously completed this study.
(These participants participated on May 8 and 9, 2015)

Your task is to estimate the least amount of money for which the average participant would eat a given food item.

To be very precise: Suppose we had offered each of the previous participants just enough so (s)he would eat the food item.
How much money would we have spent, on average, per participant, for each food item?

Since all of the previous participants have completed the same decision lists as you, we know the correct answer to this question. We calculated the correct answer only using the choices participants had made BEFORE the food items had been handed out.

You will answer 15 questions in this stage.

Payment for this stage

There is a 6% chance that **one** of your decisions in this part will be carried out.

We told you in the beginning that you would receive \$20 for following through with your decisions. If this stage of the experiment you do not make any decisions regarding food items. Instead, you will get these \$20 if your estimate on the question that is selected for payment is perfect.

If your estimate is off, we will discount money from these \$20. Specifically, for each \$1 your estimate differs from the true amount, we will discount \$0.50 from your payment.

For instance, if you underestimate the true average by \$10, we will discount $10 * \$0.50 = \5 from your payment, and you will earn \$15 for part 1 of this experiment. If you overestimate the true average by \$15, we will discount $15 * \$0.50 = \7.50 from your payment, and you will earn \$12.50 for part 1 of this experiment.

If you have any questions, please raise your hand.

The following five questions are the same as those before.

However:

This time we are asking you only about those participants who were offered **\$30** in the main decisions.

That is, we are asking:

How much would the average participant **amongst those who were offered \$30 in the main decision** at least need to be paid to eat [specified food item].

The following five questions are the same as those before.

However:

This time we are asking you only about those participants who were offered **\$3** in the main decisions.

That is, we are asking:

How much would the average participant **amongst those who were offered \$3 in the main decision** at least need to be paid to eat [specified food item].

D.1.2 Online Experiment

WELCOME

This is a research study run by the department of economics at Stanford University.

This survey will take about 40 MINUTES to complete.

In addition to your HIT payment, you will get a bonus of \$3 for completing this study.

HOWEVER, depending on your choices, and on luck, you may get up to an additional \$3 (so that you earn, a total of \$9, including the payment for the HIT), or lose that bonus (so that you earn only the HIT payment of \$3)

Please pay FULL ATTENTION to the study.

Please take this study only once. If you have participated in this, or a similar study before, the survey won't let you continue.

By clicking the button below, you consent to participating in this research study.

Instructions

PLEASE READ CAREFULLY

You will have correctly to answer questions about these instructions to continue.

This study has 6 parts. Each part has two rounds.

At the end of the study, the computer will randomly select one part of this study, and one round within this part. Your earnings will be determined by that ONE decision you made in this part, and by luck.

The study ends with a short survey.

ABOUT LUCK

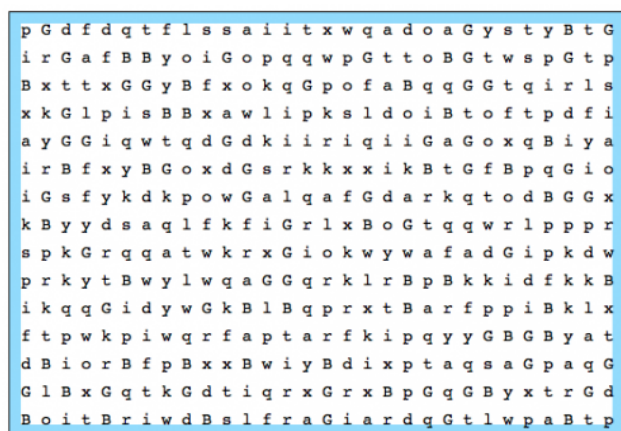
We guarantee, on behalf of Stanford University's Economics Department, that when we tell you that something will happen with some chance out of 100, it *will* happen with exactly that chance. Anything else would violate the IRB protocol under which we run this study (IRB26519).

In this study, you will see several pictures of scrambled letters. Each worker will see different pictures.

Because only one part of the study will determine your entire payment, you should make each decision as if it is the one that counts - it might be!

(Sometimes, the continue button will appear only after a few seconds.)

In each part of the study there will be a NEW picture consisting of many letters such as this one.



A picture can be GOOD or BAD. A picture is Good if it contains more letters G than letters B. Otherwise, it is Bad.

In each part, the computer randomly decides whether the picture in that round is Good or Bad, with equal chance. This does NOT depend on anything that happened before.

In each part, you will be given the option to bet on the picture.

If you bet, and the picture is Good (has more Gs than Bs), you will win money (added to your Bonus).

If you bet, and the picture is Bad (has more Bs than Gs), you will lose money (discounted from your Bonus).

A table like this will show you some information about the picture, and how much money you can win or lose.

	Good picture (chances 1 out of 2)	Bad picture (chances 1 out of 2)
Letters "G"	50	40
Letters "B"	40	50
If you bet on picture	WIN \$2	LOSE \$2

How much you can win and lose from betting is different in different parts.

Before you make your decision, you can get some idea whether the picture is Good or Bad. You will see the whole picture, as below; and can look at it for as long as you want.

```

p G d f d q t f l s s a i i t x w q a d o a G y s t y B t G
i r G a f B B y o i G o p q q w p G t t o B G t w s p G t p
B x t t x G G y B f x o k q G p o f a B q q G G t q i r l s
x k G l p i s B B x a w l i p k s l d o i B t o f t p d f i
a y G G i q w t q d G d k i i r i q i i G a G o x q B i y a
i r B f x y B G o x d G s r k k x x i k B t G f B p q G i o
i G s f y k d k p o w G a l q a f G d a r k q t o d B G G x
k B y y d s a q l f k f i G r l x B o G t q q w r l p p p r
s p k G r q q a t w k r x G i o k w y w a f a d G i p k d w
p r k y t B w y l w q a G G q r k l r B p B k k i d f k k B
i k q q G i d y w G k B l B q p r x t B a r f p p i B k l x
f t p w k p i w q r f a p t a r f k i p q y y G B G B y a t
d B i o r B f p B x x B w i y B d i x p t a q s a G p a q G
G l B x G q t k G d t i q r x G r x B p G q G B y x t r G d
B o i t B r i w d B s l f r a G i a r d q G t l w p a B t p

```

IMPORTANT

Whether you will win or lose money from betting on the picture depends on whether there are more Gs than Bs IN TOTAL, and NOT on how many Gs and Bs you happen to have seen.

After you are finished examining the picture, you decide whether or not to bet on the picture.

- If you take the bet:
 - if the picture is good, you win (added to your bonus)
 - if the picture is bad, you lose (discounted from your bonus)
- If you do not take the bet: You neither win nor lose any additional money.

Sometimes you will not know how much you can win from betting on a good picture, or how much you can lose from betting on a bad picture when you examine the picture; you will only learn this just before you decide whether to bet. In that case, there's an equal chance that from betting on a good picture, you'll win \$0.50 or you'll win \$3, and an equal chance that from betting on a bad picture, you'll lose \$0.50 or you'll lose \$3

Each time after you have decided whether or not to bet on the picture the survey will ask you how certain you are that the picture you have just seen was a good picture, in a question such as the following:

How sure are you about the yellow picture?											
It is...											
definitely good	most likely good	very likely good	quite likely good	fairly likely good	slightly likely good	slightly likely bad	fairly likely bad	quite likely bad	very likely bad	most likely bad	definitely bad
100%	90-99%	80-90%	70-80%	60-70%	50 - 60%	40 - 50%	30-40%	20-30%	10-20%	1-10%	0%
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

The number below each option is the percentage chance with which you think the picture was good.

There is a 1 in 5 chance that your earnings from this study may be determined by your answer to such a question. Depending on your answer, the picture you have seen, and luck, your bonus may raise or fall by \$0.50.

The payment procedure is designed such that it is in your best interest to give your best, genuine answer to this question.

If you believe, for example, that it is about 75% likely that you have seen a good picture, it is in your best interest to select "quite likely good (70 - 80%)". If you believe, for example that it is about 25% likely that you have seen a good picture (that is, you believe it is about 75% likely that you have seen a bad picture), then it is in your best interest to select "quite likely bad (20 - 30%)".

One of your decisions from this study will be randomly selected for payment. Thus, you will be paid for EITHER for your bet on a picture (with a 4 in 5 chance), OR for the answer you give to the questions explained above, but never for both.

*Read this if you would like to know more about the payment mechanism and about why it is in your best interest to answer this questions according to you true beliefs.
(We will not ask you test questions about the remaining material on this page.)*

The payment procedure works like this.

For most choices you can select, there is a range of chances (for example 50 - 60%). Your payment is determined by the number in middle of the range you select (for example 55%, if you select the range 50 - 60 %).

Suppose you select a choice for which the middle of the range is some number X . The computer will randomly and secretly draw another number Y between 0 and 100. If the number the computer randomly draws is the larger one, that is if $Y > X$, then you will win \$0.5 with chance Y in 100 (and lose \$0.5 if you don't win). If the number you stated is the larger one, that is, if $X > Y$, then you will win if the picture you have seen is good. So if X is your genuine belief that the picture you have seen was good, you will win with chance X or with chance Y , whichever of the two is larger.

Why is it in my best interest to answer this question according to my genuine beliefs?

Simply, the reason is that you lower your chance of winning if you state a chance that is lower than you genuinely believe, and you also lower your chance of winning if you state something that is higher than you genuinely believe. So the best you can do is state what you genuinely believe.

To see why, it's best to go through an example.

Here's why you lose from stating a chance that is higher than you genuinely think is true. For example, suppose you genuinely believe the chance that the picture is good 60%, but you reply that you think it is good with a higher chance, say 90%. Suppose the number Y that the computer draws is between 60% and 90%, let's say it is 80%. That is lower than what you told us (which is 90%), so you will *not* play the computers' bet. Instead, you will win if the picture is good, which you genuinely think only occurs with 60% chance. The computers' bet would have given you a higher, 80%, chance instead. Hence, you hurt your chance of winning by stating the picture was more likely good than you genuinely think.

And here's why you lose from stating a lower chance than you genuinely think is true. For example, suppose again you genuinely believe the chance that the picture is good 60%, but you reply that you think it is good with a lower chance, say 10%. Suppose the number Y that the computer draws is between 10% and 60%, let's say it is 30%. That is higher than what you told us (which is 10%), so you will play the computers' bet and win with chance 30%. That is lower than if you had instead received the bet on the picture, which, according to your genuine belief, has a 60% chance. Hence, you hurt your chance of winning by stating the picture was less likely good than you genuinely think.

Therefore, the best you can possibly do is to select exactly the answer that corresponds to your genuine beliefs.

If you have any questions about this payment mechanism, please send an email to sambuehl@stanford.edu.

To make sure you got all of this, check all statements below that are true. You can only continue if you tick all boxes correctly.

Use the back button on the bottom if you would like to revisit the instructions.

(Do not try random combinations, there are far too many possible combinations. If you feel you understand the instructions, but still cannot continue, or otherwise have a question, send an email to sambuehl@stanford.edu)

- | | |
|--|---|
| <input type="checkbox"/> Whether I win or lose depends on whether there are more Gs than Bs in the picture IN TOTAL, and not on how many Gs and Bs I happen to have seen | <input type="checkbox"/> I can study the picture for as long as I like before I make a decision |
| <input type="checkbox"/> If I think there is a 60% chance I have seen a bad picture, this means that I think there is a 40% chance that I have seen a good picture | <input type="checkbox"/> It is possible that I think the chance that a picture was good is 75% and the chance that the same picture was bad is 50% |
| <input type="checkbox"/> If I am paid for a part with a picture, I will be paid EITHER for the bet I take on that picture, OR for the my answer to the question how certain I am about the picture I have seen, but NOT for BOTH | <input type="checkbox"/> When I am asked about how certain I am about a picture, I will earn most from this study if I state something a little higher than I truly think |
| <input type="checkbox"/> When I am asked about how certain I am about a picture, I will earn most from this study if I state exactly what I truly think | <input type="checkbox"/> At the end of the study, the computer will randomly select one decision I made. I will be paid for that and only that decision. |
| <input type="checkbox"/> Whether I win or lose depends on whether I HAVE SEEN more Gs than Bs, and not on how many Bs and Gs there are in the picture in total. | <input type="checkbox"/> In each part, I will see the same picture |
| <input type="checkbox"/> When I am asked about how certain I am about a picture, I will earn most from this study if I state something a little lower than I truly think | |

Stanford

Now the study starts.

Your decisions count for real money.

<<

>>

	Good picture (chances 1 out of 2)	Bad picture (chances 1 out of 2)
Letters "G"	50	40
Letters "B"	40	50
If you bet on picture	<i>learn later</i>	<i>learn later</i>

Study the red picture for as long as you like, to learn whether you want to bet on the picture or not.



Click NEXT to make your decisions.

(You can NOT go back once you click NEXT)

	Good picture (chances 1 out of 2)	Bad picture (chances 1 out of 2)
Letters "G"	50	40
Letters "B"	40	50
If you bet on picture	WIN \$3	LOSE \$3

Make a decision

- ☐ Bet on the picture. WIN \$3 if the red picture is Good, LOSE \$3 if it is Bad
- ☐ Do not bet on the red picture

How sure are you about the red picture?

It is...

definitely good	most likely good	very likely good	quite likely good	fairly likely good	slightly likely good	slightly likely bad	fairly likely bad	quite likely bad	very likely bad	most likely bad	definitely bad
100%	90-99%	80-90%	70-80%	60-70%	50 - 60%	40 - 50%	30-40%	20-30%	10-20%	1-10%	0%
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

(The number below each option is the percentage chance with which you think that the red picture is good.)

If you like, you may now go back to change your decision on whether or not to bet on the picture.